

4.3.2.2 Cálculo dos Multiplicadores de Lagrange

Vejamos como calcular numericamente os multiplicadores de Lagrange. Associados a um ponto que foi calculado como sendo o ponto mínimo. Consideremos novamente a condição de estacionaridade do Lagrangeano usando notação matricial:

$$\nabla f - \mathbf{N}\lambda = 0$$

onde:

$$n_{ij} = \frac{\partial g_j}{\partial x_i} \quad j = 1, \dots, r \quad (\text{restrição ativas})$$

$$i = 1, \dots, n \quad (\text{número de variáveis de projeto})$$

Assim r é o número de incógnitas (λ - multiplicadores de Lagrange associados às restrições ativas) na equação matricial de estacionaridade e n o número de equações. Se $r=n$ a equação de estacionaridade apresenta solução única. Se $r < n$ (caso típico) o número de incógnitas é menor do que o número de equações. Nesse último caso, uma solução aproximada de λ pode ser obtida pelo método dos mínimos quadrados, como descrito abaixo:

$$u = \mathbf{N}\lambda - \nabla f \Rightarrow \text{Min } \|u\|^2 = (\mathbf{N}\lambda - \nabla f)^T (\mathbf{N}\lambda - \nabla f) \Rightarrow 2\mathbf{N}^T \mathbf{N}\lambda - 2\mathbf{N}^T \nabla f = 0 \Rightarrow \\ \Rightarrow \lambda = (\mathbf{N}^T \mathbf{N})^{-1} \mathbf{N}^T \nabla f$$

Assim, o que se obtém com essa formulação é o vetor λ que resulta no menor valor de u (não necessariamente zero).

O cálculo acima é usado em qualquer situação $r=n$ ou $r \neq n$. No entanto, se o ponto em questão satisfaz as equação de estacionaridade λ é a solução exata ($u=0$). Substituindo λ na equação $\nabla f - \mathbf{N}\lambda = 0$, obtém-se:

$$\mathbf{P}\nabla f = 0, \quad \text{onde } \mathbf{P} = \mathbf{I} - \mathbf{N}(\mathbf{N}^T \mathbf{N})^{-1} \mathbf{N}^T$$

e \mathbf{P} é chamada matriz de projeção.

Considere o exemplo abaixo onde deseja-se determinar se o ponto $(-2, -2, 4)$ é um mínimo local:

$$\begin{aligned} \text{Min } & f = x_1 + x_2 + x_3 \\ & x_1, x_2, x_3 \\ \text{tal que } & g_1 = 8 - x_1^2 - x_2^2 \geq 0 \\ & g_2 = x_3 - 4 \geq 0 \\ & g_3 = x_2 + 8 \geq 0 \end{aligned}$$

Solução:

Observa-se que nesse ponto as restrições g_1 e g_2 estão ativas ($g_1=g_2=0$) e a restrição g_3 está inativa ($g_3=6$), ou seja, $r=2$ e $n=3$. Assim, podemos construir a matriz \mathbf{N} :

$$\left. \begin{aligned} \frac{\partial g_1}{\partial x_1} = -2x_1 = 4; \quad \frac{\partial g_1}{\partial x_2} = -2x_2 = 4; \quad \frac{\partial g_1}{\partial x_3} = 0 \\ \frac{\partial g_2}{\partial x_1} = 0; \quad \frac{\partial g_2}{\partial x_2} = 0; \quad \frac{\partial g_2}{\partial x_3} = 1 \end{aligned} \right\} \Rightarrow \mathbf{N} = \begin{bmatrix} 4 & 0 \\ 4 & 0 \\ 0 & 1 \end{bmatrix}$$

e portanto utilizando o método dos mínimos quadrados acima, temos:

$$\frac{\partial f}{\partial x_1} = \frac{\partial f}{\partial x_2} = \frac{\partial f}{\partial x_3} = 1 \Rightarrow \nabla f = \begin{Bmatrix} 1 \\ 1 \\ 1 \end{Bmatrix}; \mathbf{N}^T \mathbf{N} = \begin{bmatrix} 32 & 0 \\ 0 & 1 \end{bmatrix}; \mathbf{N}^T \nabla f = \begin{Bmatrix} 8 \\ 1 \end{Bmatrix} \Rightarrow \lambda = (\mathbf{N}^T \mathbf{N})^{-1} \mathbf{N}^T \nabla f = \begin{Bmatrix} 1/4 \\ 1 \end{Bmatrix}$$

e calculando $\mathbf{P}\nabla f = 0$, temos:

$$\mathbf{P}\nabla f = [\mathbf{I} - \mathbf{N}(\mathbf{N}^T \mathbf{N})^{-1} \mathbf{N}^T] \nabla f = 0$$

Portanto, λ é a solução exata e todos os seus componentes são positivos. No entanto, como o número de restrições ativas (2) é diferente do número de variáveis (3) devemos obter a matriz Hessiana para verificar se as condições KKT são satisfeitas. Note que a equação $\nabla f - \mathbf{N}\lambda = 0$ é igual à:

$$\begin{cases} 1 - 4\lambda_1 = 0 \\ 1 - 4\lambda_1 = 0 \\ 1 - \lambda_2 = 0 \end{cases}$$

ou seja, as duas primeiras equações são linearmente dependentes, e portanto temos apenas 2 equações linearmente independentes. Se as 3 fossem linearmente independentes, o λ obtido não seria a solução exata e a condição de estacionaridade não seria satisfeita (e portanto KKT).

4.3.2.3 Definição de Problemas Convexos

A definição de problema convexo é muito importante em otimização, pois está relacionada com o conceito de mínimo local e mínimo global. Um problema é dito convexo se a função objetivo e o domínio viável correspondentes são convexas.

Uma função é definida matematicamente como convexa se dado dois pontos \mathbf{x}_1 e \mathbf{x}_2 , então:

$$f[\alpha \mathbf{x}_2 + (1 - \alpha) \mathbf{x}_1] \leq \alpha f(\mathbf{x}_2) + (1 - \alpha) f(\mathbf{x}_1), \quad 0 < \alpha < 1$$

Intuitivamente, uma função é convexa se é capaz de "reter" água, ou seja, tem a forma de uma tija. A negativa dessa função é chamada côncava, e finalmente existem as funções que não são nem convexas nem côncavas. A Fig.4.3.7 ilustra os três tipos de função.

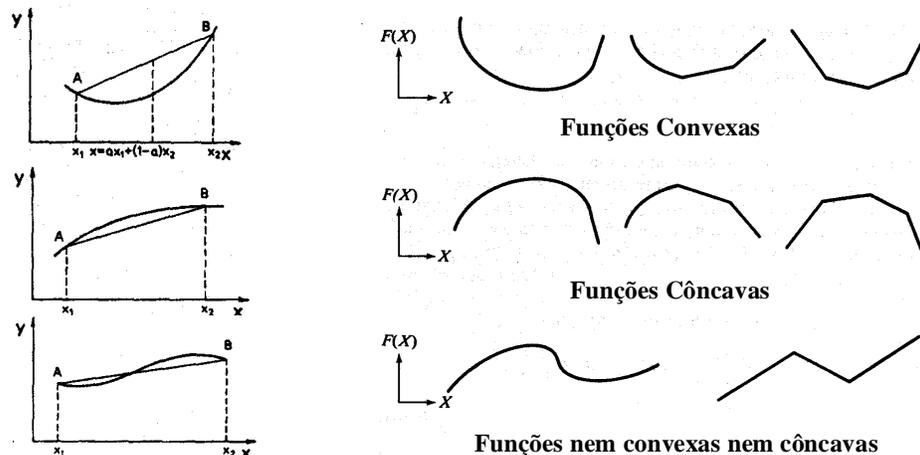


Fig.4.3.7: Funções convexas e côncavas.

Baseado na expressão matemática acima demonstra-se que se a função é convexa, a sua matrix Hessiana é positiva semi-definida, e se a função é côncava a sua matriz Hessiana será negativa semi-definida. Uma função linear é uma função convexa (sinal de igual na expressão acima).

Um domínio é definido matematicamente como convexo se dado dois pontos \mathbf{x}_1 e \mathbf{x}_2 , então:

$$\text{Seja } \mathbf{w} = \alpha \mathbf{x}_2 + (1 - \alpha) \mathbf{x}_1, \quad 0 < \alpha < 1 \text{ Se } \mathbf{w} \subset \text{domínio para } \forall \mathbf{x}_1, \mathbf{x}_2 \Rightarrow \text{domínio é convexo}$$

Fisicamente quer dizer que se escolhermos dois pontos quaisquer no interior do domínio e os unirmos por uma reta, então se todos os pontos dessa reta (entre os dois pontos) estiverem no interior do domínio, esse será considerado convexo. Caso contrário será um domínio côncavo. A Fig.4.3.8 ilustra esse conceito.

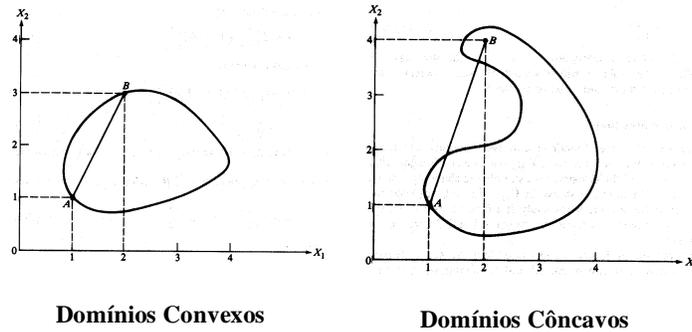


Fig.4.3.8: Verificação de domínios convexos e côncavos.

A Fig.4.3.9 mostra exemplos de domínios convexos e côncavos.

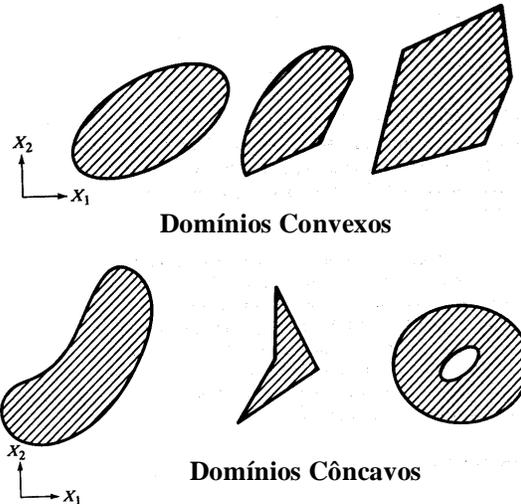


Fig.4.3.9: Exemplos de domínios convexos e côncavos.

Considerando que no problema de otimização o domínio viável é definido por um conjunto de equações, mediante a expressão matemática acima demonstra-se que esse domínio será convexo se todas as restrições de desigualdade são côncavas ou lineares (retas) e todas as restrições de igualdade são lineares.

Assim se a função objetivo e o domínio viável são convexos, o problema é dito convexo. A Fig.4.3.10 ilustra um problema convexo.

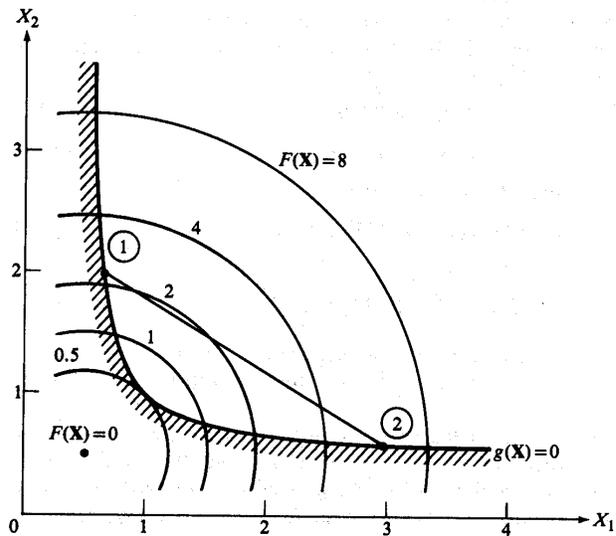


Fig.4.3.10: Exemplo de problema convexo.

A grande vantagem de um problema ser convexo é que apresenta um mínimo global. Ou seja, se provarmos que o problema é convexo, então a solução ótima obtida, dado que as condições KKT sejam satisfeitas, necessariamente será um mínimo global. Dessa forma, para problemas convexos as condições KKT também são condições suficientes de optimalidade, uma vez que o problema somente apresenta um mínimo global. No entanto nem sempre é fácil identificar se um problema é ou não convexo. Alguns problemas clássicos em otimização estrutural foram matematicamente provados serem convexos, como a minimização da flexibilidade, enquanto outros como maximização da frequência de ressonância e carga de flambagem não são problemas convexos.

Infelizmente a grande maioria dos problema não são convexos. No entanto, existe um método de solução chamado programação linear seqüencial em que um problema não-convexo é aproximado por uma seqüência de subproblemas convexos.

Vejam os um exemplo de problema convexo. Determine se o problema de otimização abaixo é convexo ou não:

$$\begin{array}{ll}
 \text{Min} & f = 3x_1 + \sqrt{3}x_2 \\
 x_1, x_2 & \\
 \text{tal que} & g_1 = 3 - \frac{18}{x_1} - \frac{6\sqrt{3}}{x_2} \geq 0 \\
 & g_2 = x_1 - 5,73 \geq 0 \\
 & g_3 = x_2 - 7,17 \geq 0
 \end{array}$$

Solução:

Note que a função objetivo é linear e portanto é convexa. Com relação ao domínio viável temos duas restrições de desigualdade lineares (retas), resta analisar se a restrição g_1 de desigualdade é côncava. Para isso devemos calcular a sua matriz Hessiana:

$$\mathbf{H}_1 = \begin{bmatrix} -36/x_1^3 & 0 \\ 0 & -12\sqrt{3}/x_2^3 \end{bmatrix}$$

Assim, considerando que x_1 e x_2 serão positivos (ver restrições g_2 e g_3 acima), a matriz acima é negativa definida, e portanto a restrição g_1 é côncava. Assim o domínio viável é convexo e o problema é convexo. Isso significa que a solução ótima encontrada, dado que as condições KKT sejam satisfeitas, será um mínimo global do problema.

4.3.2.4 Ponto Sela ("Saddle Point")

Dizemos que uma função $F(x,y)$ possui um ponto sela ("Saddle Point") se:

$$F(\mathbf{x}^*, \mathbf{y}) \leq F(\mathbf{x}^*, \mathbf{y}^*) \leq F(\mathbf{x}, \mathbf{y}^*)$$

onde: $F(\mathbf{x}^*, \mathbf{y}^*) \left\{ \begin{array}{l} \text{mínimo em } \mathbf{x} \\ \text{máximo em } \mathbf{y} \end{array} \right.$

Assim $F(\mathbf{x}, \mathbf{y}^*)$ apresenta um mínimo em \mathbf{x}^* e $F(\mathbf{x}^*, \mathbf{y})$ apresenta um máximo em \mathbf{y}^* . A Fig.4.3.11 ilustra o conceito de ponto sela $(\mathbf{x}^*, \mathbf{y}^*)$ para uma função bidimensional:

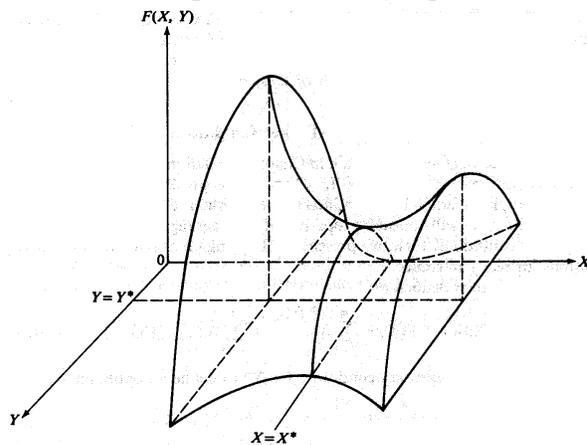


Fig.4.3.11: Ponto Sela.

Pelas desigualdades acima, conclue-se que:

$$\begin{cases} \nabla_y F(\mathbf{x}^*, \mathbf{y}) \leq 0 \\ \nabla_x F(\mathbf{x}, \mathbf{y}^*) \geq 0 \end{cases} \text{ e no ponto sela : } \begin{cases} \nabla_y F(\mathbf{x}^*, \mathbf{y}^*) = 0 \\ \nabla_x F(\mathbf{x}^*, \mathbf{y}^*) = 0 \end{cases}$$

No caso particular da função Lagrangeana $L(\mathbf{x}, \lambda)$, é demonstrado que $L(\mathbf{x}^*, \lambda^*)$ é um ponto sela, representando um mínimo com relação à \mathbf{x} e um máximo com relação à λ . Assim:

$$\nabla_x L(\mathbf{x}^*, \lambda^*) = \nabla_\lambda L(\mathbf{x}^*, \lambda^*) = 0$$

e:

$$\begin{cases} \nabla_\lambda L(\mathbf{x}^*, \lambda) \leq 0 \\ \nabla_x L(\mathbf{x}, \lambda^*) \geq 0 \end{cases}$$

Sabendo que $(\mathbf{x}^*, \lambda^*)$ define um ponto sela para a função Lagrangeana podemos definir uma forma alternativa de encontrar o ponto ótimo do problema.

A função Lagrangeana pode ser definida em termos somente de λ :

$$L(\lambda) = \min_x L(\mathbf{x}, \lambda)$$

Como $L(\lambda)$ é máximo em relação à λ no ponto ótimo, para obter o ponto ótimo queremos maximizar $L(\lambda)$:

$$\max_{\lambda} L(\lambda) = \max_{\lambda} \min_x L(\mathbf{x}, \lambda)$$

Analogamente, o mesmo ponto ótimo pode ser obtido pelo outro caminho (tradicional):

$$L(\mathbf{x}) = \max_{\lambda} L(\mathbf{x}, \lambda) \Rightarrow \min_x L(\mathbf{x}) = \min_x \max_{\lambda} L(\mathbf{x}, \lambda)$$

Assim:

$$\boxed{\max_{\lambda} \min_x L(\mathbf{x}, \lambda) \equiv \min_x \max_{\lambda} L(\mathbf{x}, \lambda)}$$

Assim trata-se de um problema do tipo min-max.

Vejamos um exemplo de obter o ponto ótimo usando o conceito:

$$\max_{\lambda} L(\lambda) = \max_{\lambda} \min_x L(\mathbf{x}, \lambda)$$

Considere o problema abaixo e obtenha o seu ponto ótimo:

$$\boxed{\begin{array}{l} \text{Min} \quad \frac{1}{x} \\ \text{tal que } x - 1 \leq 0 \\ x \geq 0 \end{array}}$$

Solução:

O Lagrangeano do problema vale: $L(x, \lambda) = \frac{1}{x} + \lambda(x - 1)$

Calculando o mínimo em x : $\nabla_x L(x, \lambda) = 0 \Rightarrow -\frac{1}{x^2} + \lambda = 0 \Rightarrow$

$$\Rightarrow x = \pm \frac{1}{\sqrt{\lambda}} \text{ (mas } x \geq 0) \Rightarrow x = \frac{1}{\sqrt{\lambda}} \Rightarrow L(\lambda) = \sqrt{\lambda} + \lambda \left(\frac{1}{\sqrt{\lambda}} - 1 \right) = 2\sqrt{\lambda} - \lambda$$

Calculando o máximo em λ : $\nabla_{\lambda} L(\lambda) = \frac{1}{\sqrt{\lambda}} - 1 \Rightarrow \lambda^* = 1 \Rightarrow x^* = 1$

$$\therefore (x^*, \lambda^*) = (1, 1) \text{ (Ponto Sela)}$$

A Fig.4.3.11.1 ilustra a função Lagrangeano do problema acima mostrando que a solução ótima é um ponto sela.

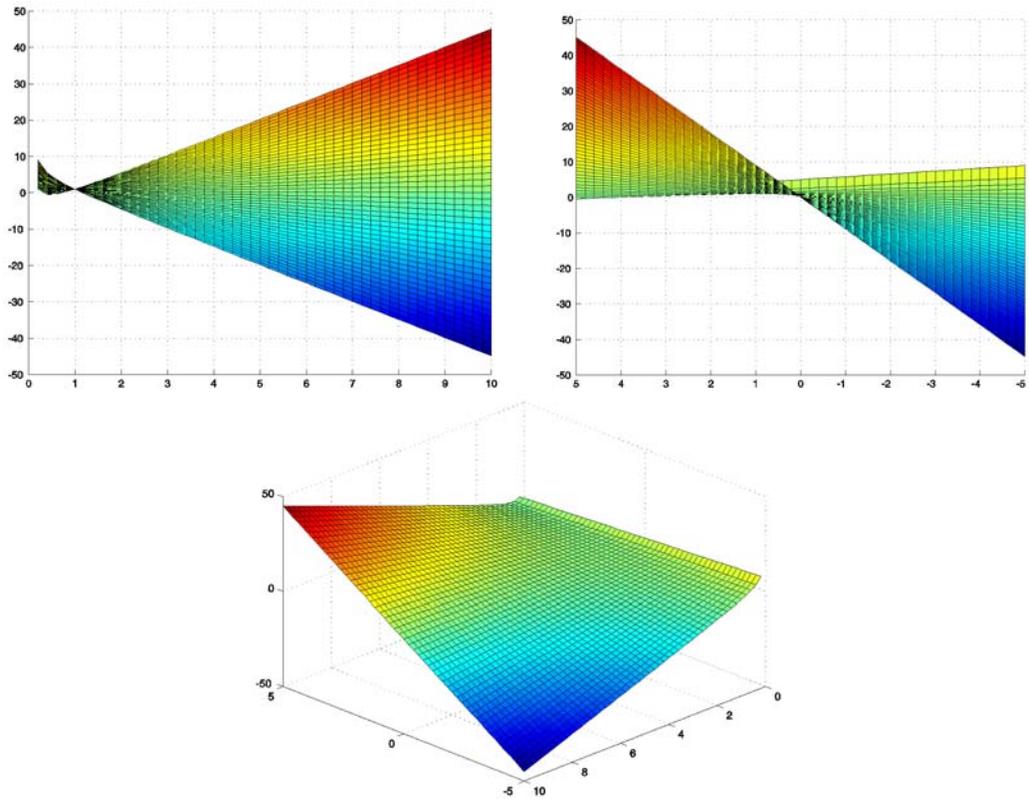


Fig.4.3.11.1: Função Lagrangeana do problema.

Os gradientes devem ser contínuos em \mathbf{x} e λ , caso contrário, a abordagem de solução para o problema min-max descrita na seção 3.3 deve ser utilizada.

4.3.2.5 Dualidade

Vamos definir inicialmente o conceito de problema "Primal" e problema "Dual" considerando um problema de programação linear. O problema na sua forma Primal consiste na formulação que foi apresentada anteriormente:

$$\begin{array}{l} \text{Min } f_p = \mathbf{c}^T \mathbf{x} \\ \text{tal que } \mathbf{Ax} \leq \mathbf{b} \\ \mathbf{x} \geq \mathbf{0} \end{array}$$

Vários métodos estão sendo estudados para a solução desse problema sem se preocupar com o valor dos multiplicadores de Lagrange. A princípio os multiplicadores de Lagrange não são calculados porque essa informação não é utilizada. No entanto, se os λ^* forem conhecidos com antecedência, o problema de otimização pode ser resolvido como um problema de minimização sem restrição usando os métodos que serão descritos na seção 4.3.2.8.

Vejam agora a formulação do problema Dual equivalente ao problema Primal apresentado acima:

$$\begin{array}{l} \text{Max } f_D = -\lambda^T \mathbf{b} \\ \lambda \\ \text{tal que } \mathbf{A}^T \lambda \geq -\mathbf{c} \\ \lambda \geq \mathbf{0} \end{array}$$

onde os multiplicadores de Lagrange são denominados *variáveis duais*. Como se sabe, após resolver o problema Primal, podemos recuperar as variáveis duais λ^* . Da mesma forma, deve ser possível em alguns casos, resolvendo o problema dual, recuperar o valor ótimo das variáveis primais \mathbf{x}^* . A motivação para essa abordagem é que na maioria dos problemas de otimização somente algumas restrições são críticas no ponto ótimo, assim somente alguns λ^* são não-zero. As variáveis duais (λ^*) associadas às restrições de igualdade serão não-zero, e portanto essas restrições são facilmente incluídas na otimização.

Além disso, sabemos que como o ótimo do Lagrangeano (\mathbf{x}^*, λ^*) é um ponto sela, as funções objetivo dual e primal devem ter o mesmo valor. Assim, se compararmos o ótimo da função objetivo dual com a primal, o grau de igualdade entre elas é uma medida da precisão com que o problema primal foi resolvido, ou seja, o problema dual sempre fornece um limite inferior no valor ótimo do problema primal (ver Fig.4.3.11).

Métodos duais tem sido extensivamente usados em programação linear pra aumentar a eficiência da otimização. Na programação não-linear, esta abordagem não é em geral utilizada porque nem sempre as variáveis primais podem ser manipuladas de forma conveniente no problema dual.

Na programação linear a vantagem de se utilizar a dualidade é clara. Se considerarmos que o problema primal possui m restrições e n variáveis de projeto (primais), onde $m > n$, ao convertê-lo num problema dual, teremos n restrições e m variáveis de projeto. Assim, no problema dual teremos mais variáveis de projeto do que restrições o que computacionalmente é mais eficiente. Vamos demonstrar matematicamente a equivalência dos problemas primal e dual (conceito da dualidade) para a programação linear. Ambas as formulações foram apresentadas acima.

Escrevendo o lagrangeano do problema primal e calculando o gradiente com relação a \mathbf{x} (pois queremos inicialmente minimizar $L(\mathbf{x}, \lambda)$ em \mathbf{x}), temos:

$$L(\mathbf{x}, \lambda) = \mathbf{c}^T \mathbf{x} + \lambda^T \{\mathbf{A}\mathbf{x} - \mathbf{b}\} \Rightarrow \nabla_{\mathbf{x}} L(\mathbf{x}, \lambda) = \mathbf{c} + \mathbf{A}^T \lambda \geq \mathbf{0}$$

onde a desigualdade vem do conceito de ponto sela apresentado na seção 4.3.2.4. Mas no ponto ótimo (\mathbf{x}^*) devido a condição de estacionaridade devemos ter $\nabla_{\mathbf{x}} L(\mathbf{x}, \lambda) = \mathbf{0}$. Assim, multiplicando a equação acima por \mathbf{x} e impondo a igualdade:

$$\mathbf{c}^T \mathbf{x}^* + \lambda^T \mathbf{A} \mathbf{x}^* = 0$$

Substituindo em $L(\mathbf{x}, \lambda)$, obtemos $L(\lambda) = -\lambda^T \mathbf{b}$, que representa a função objetivo dual a ser maximizada. A equação $\mathbf{c} + \mathbf{A}^T \lambda \geq \mathbf{0}$ é a restrição do problema dual, além de $\lambda_j \geq 0$. Note que se tivéssemos $f_p = \mathbf{c}^T \mathbf{x} + F_0$ a constante F_0 seria simplesmente adicionada à $L(\lambda)$.

Assim formula-se o problema dual apresentado acima. Esse problema é resolvido como um problema comum de programação linear e sabemos que as funções objetivo do problema dual e primal são iguais no ponto ótimo, ou seja:

$$f(\mathbf{x}^*) = L(\lambda^*)$$

Para obter \mathbf{x}^* (variáveis primais), devemos usar a informação que se $\lambda_j \neq 0$ a j -ésima restrição é ativa no problema primal, e se $\lambda_j = 0$ a j -ésima restrição é inativa no problema primal.

Vejamos um exemplo do uso da dualidade na solução de problemas de programação linear. Considere o problema de programação linear abaixo:

$$\begin{array}{l} \text{Min } f = -4x_1 - x_2 + 50 \\ x_1, x_2 \\ \text{tal que} \end{array} \begin{array}{l} x_1 - x_2 \leq 2 \\ x_1 + 2x_2 \leq 8 \\ x_1 + x_2 \leq 10 \\ -5x_1 + x_2 \leq 5 \\ x_1 \geq 0; \quad x_2 \geq 0 \end{array}$$

Solução:

A matriz **A** e os vetores **b**, **c** são iguais à:

$$\mathbf{c} = \begin{Bmatrix} -4 \\ -1 \end{Bmatrix}; \mathbf{A} = \begin{Bmatrix} 1 & -1 \\ 1 & 2 \\ 1 & 1 \\ -5 & 1 \end{Bmatrix}; \mathbf{b} = \begin{Bmatrix} 2 \\ 8 \\ 10 \\ 5 \end{Bmatrix}$$

Assim, convertendo no problema dual equivalente usando a teoria acima, obtemos:

$$\begin{array}{l} \text{Max } L = -2\lambda_1 - 8\lambda_2 - 10\lambda_3 - 5\lambda_4 + 50 \\ \lambda_1, \lambda_2, \lambda_3, \lambda_4 \\ \text{tal que} \end{array} \begin{array}{l} \lambda_1 + \lambda_2 + \lambda_3 - 5\lambda_4 \geq 4 \\ -\lambda_1 + 2\lambda_2 + \lambda_3 + \lambda_4 \geq 1 \\ \lambda_i \geq 0; \quad i = 1, 4 \end{array}$$

Note que temos menos restrições e mais variáveis. Resolvendo o problema:

$$\lambda^* = \begin{Bmatrix} 7/3 \\ 5/3 \\ 0 \\ 0 \end{Bmatrix} \Rightarrow L(\lambda^*) = 32$$

Agora, no cálculo das variáveis primais utilizamos a equação de estacionaridade de $L(\mathbf{x}, \lambda)$ que é igual a equação $\mathbf{c} + \mathbf{A}^T \lambda \geq 0$ juntamente com as equações de restrições ativas no problema primal ($\lambda_j \neq 0$):

$$\left\{ \begin{array}{l} \nabla_{\mathbf{x}} L(\mathbf{x}^*, \lambda^*) = \begin{Bmatrix} -4 \\ -1 \end{Bmatrix} + \begin{bmatrix} 1 & 1 & 1 & -5 \\ -1 & 2 & 1 & 1 \end{bmatrix} \begin{Bmatrix} 7/3 \\ 5/3 \\ 0 \\ 0 \end{Bmatrix} = \begin{Bmatrix} 0 \\ 0 \end{Bmatrix} \Rightarrow \mathbf{x}^* = \begin{Bmatrix} 4 \\ 2 \end{Bmatrix} \Rightarrow f(\mathbf{x}^*) = 32 \\ x_1^* - x_2^* = 2; \quad x_1^* + 2x_2^* = 8 \end{array} \right.$$

Assim, como era esperado $f(\mathbf{x}^*) = L(\lambda^*)$.

4.3.2.6 Programação Quadrática

Um problema de otimização não-linear que também pode ser resolvido usando o conceito de dualidade é um problema de programação quadrática (PQ). Supondo n variáveis de projeto e m restrições, um problema de programação quadrática é definido como:

$$\begin{array}{l} \text{Min}_{\mathbf{x}} F(\mathbf{x}) = \frac{1}{2} \mathbf{x}^T \mathbf{A} \mathbf{x} + \mathbf{b} \mathbf{x} \\ \text{tal que } \mathbf{B} \mathbf{x} \leq \mathbf{c} \\ \mathbf{x} \geq \mathbf{0} \end{array}$$

onde \mathbf{A} é a matriz Hessiana.

Partindo do Lagrangeano do problema e repetindo a dedução acima realizada para a programação linear temos:

$$\begin{aligned} L(\mathbf{x}, \lambda) &= \frac{1}{2} \mathbf{x}^T \mathbf{A} \mathbf{x} + \mathbf{b} \mathbf{x} + \lambda^T (\mathbf{B} \mathbf{x} - \mathbf{c}) \Rightarrow \nabla_{\mathbf{x}} L(\mathbf{x}, \lambda) = \mathbf{A} \mathbf{x} + \mathbf{b} + \mathbf{B}^T \lambda = \mathbf{0} \Rightarrow \\ &\Rightarrow \mathbf{x} = -\mathbf{A}^{-1} (\mathbf{b} + \mathbf{B}^T \lambda) \Rightarrow L(\lambda) = -\frac{1}{2} \lambda^T \mathbf{D} \lambda - \mathbf{d} \cdot \lambda - \frac{1}{2} \mathbf{b}^T \mathbf{A}^{-1} \mathbf{b} \end{aligned}$$

$$\text{onde: } \mathbf{D} = \mathbf{B} \mathbf{A}^{-1} \mathbf{B}^T \text{ e } \mathbf{d} = \mathbf{c} + \mathbf{B} \mathbf{A}^{-1} \mathbf{b}$$

Note que o termo $\frac{1}{2} \mathbf{b}^T \mathbf{A}^{-1} \mathbf{b}$ é uma constante. Assim, o problema dual equivalente é formulado como:

$$\begin{array}{l} \text{Max}_{\lambda} L(\lambda) = -\frac{1}{2} \lambda^T \mathbf{D} \lambda - \mathbf{d} \cdot \lambda \\ (\lambda \geq \mathbf{0}) \end{array}$$

onde temos m variáveis de projeto.

Dessa forma, converteu-se o problema inicial de programação quadrática com m restrições de desigualdade e n variáveis de projeto em um problema com m variáveis de projeto com simplesmente restrições do tipo $\lambda_j \geq 0$.

A solução do problema de programação quadrática (PQ) será usada na implementação da programação quadrática seqüencial descrita na seção 4.5.

No aplicativo SCILAB a rotina que resolve um problema de PQ é denominada QUAPRO. Essa rotina resolve um problema de PQ da forma:

$$\begin{aligned} \text{Min}_x \quad & f = \frac{1}{2} \mathbf{x}^T \mathbf{Q} \mathbf{x} + \mathbf{p}^T \mathbf{x} \\ \text{tal que} \quad & \mathbf{C}_1 \mathbf{x} = \mathbf{b}_1 \\ & \quad \quad \quad \downarrow (m_i \times n) \\ & \mathbf{C}_2 \mathbf{x} \leq \mathbf{b}_2 \\ & \quad \quad \quad \downarrow (m_d \times n) \\ & \mathbf{c}_i \leq \mathbf{x} \leq \mathbf{c}_s \end{aligned}$$

$$\begin{aligned} \mathbf{C} &= [\mathbf{C}_1, \mathbf{C}_2] \\ & (m_i + m_d) \times n \\ \mathbf{b} &= [\mathbf{b}_1, \mathbf{b}_2] \\ & 1 \times (m_i + m_d) \end{aligned}$$

e \mathbf{Q} é uma matriz simétrica $n \times n$.

A sua sintaxe é dada por:

$$[\mathbf{x}, \text{lagr}, f] = \text{quapro}(\mathbf{Q}, \mathbf{p}, \mathbf{C}, \mathbf{b}, \mathbf{c}_i, \mathbf{c}_s, m_i, x_0)$$

onde:

- \mathbf{x} - vetor da solução ótima;
- lagr - vetor de multiplicadores de Lagrange $(n+m_i+m_d)$. Se o valor de um multiplicador é zero, a restrição correspondente está inativa, caso contrário está ativa na ponto ótimo;
- f - valor ótimo da função;
- m_i - número de restrições de igualdade;
- x_0 - chute inicial de \mathbf{x} ; ou “ $x_0=v$ ” o cálculo é iniciado num vértice do domínio; ou “ $x_0=g$ ” o valor inicial é arbitrário;

4.3.2.7 Métodos Diretos

Vejam agora alguns métodos para a solução de problemas de otimização com restrição. Como já mencionado, esses métodos são classificados em métodos diretos e indiretos. Os métodos diretos verificam as restrições após seguirem o procedimento do algoritmo de otimização sem restrição. Entre esses métodos temos o Método das Projeções do Gradientes e dos Gradientes Reduzidos e o Método das Direções Viáveis.

a) Método dos Gradientes Reduzidos e Projeção do Gradientes

Inicialmente consideremos um problema de otimização não-linear com restrições lineares:

$$\begin{aligned} \text{Min}_x \quad & f(\mathbf{x}) \\ \text{tal que} \quad & \mathbf{g}_j = \mathbf{a}_j^T \mathbf{x} - \mathbf{b}_j \geq 0 \\ & j = 1, \dots, n_g \end{aligned}$$

Considerando r restrições ativas podemos escrever as equações de restrição como:

$$\mathbf{g}_a = \mathbf{N}^T \mathbf{x} - \mathbf{b} = 0 \quad \text{onde: } n_{ij} = a_{ij}$$

A hipótese básica do método é que \mathbf{x} se encontra no subespaço tangente às restrições ativas, dessa forma o método caminha na fronteira das restrições. Assim, se:

$$\mathbf{x}_{i+1} = \mathbf{x}_i + \alpha \mathbf{s}$$

e ambos \mathbf{x}_{i+1} e \mathbf{x}_i satisfazem $\mathbf{g}_a = \mathbf{N}^T \mathbf{x} - \mathbf{b} = 0$ então:

$$\mathbf{N}^T \mathbf{x}_{i+1} = \mathbf{N}^T \mathbf{x}_i + \alpha \mathbf{N}^T \mathbf{s} \Rightarrow \mathbf{b} = \mathbf{b} + \alpha \mathbf{N}^T \mathbf{s} \Rightarrow \mathbf{N}^T \mathbf{s} = 0$$

A direção \mathbf{s} é obtida através da solução do problema de otimização:

$$\begin{array}{l} \text{Min } \mathbf{s}^T \nabla f \\ \mathbf{s} \\ \text{tal que } \mathbf{N}^T \mathbf{s} = 0 \\ \mathbf{s}^T \mathbf{s} = 1 \end{array}$$

Considerando o Lagrangeano do problema e impondo a condição de estacionaridade, obtemos \mathbf{s} :

$$\begin{aligned} L = \mathbf{s}^T \nabla f - \mathbf{s}^T \mathbf{N} \lambda - \mu (\mathbf{s}^T \mathbf{s} - 1) &\Rightarrow \frac{\partial L}{\partial \mathbf{s}} = \nabla f - \mathbf{N} \lambda - 2\mu \mathbf{s} = 0 \Rightarrow \\ \Rightarrow \mathbf{N}^T \nabla f - \mathbf{N}^T \mathbf{N} \lambda = 0 &\Rightarrow \lambda = (\mathbf{N}^T \mathbf{N})^{-1} \mathbf{N}^T \nabla f \Rightarrow \\ \Rightarrow \mathbf{s} = \frac{1}{2\mu} [\mathbf{I} - \mathbf{N}(\mathbf{N}^T \mathbf{N})^{-1} \mathbf{N}^T] \nabla f &= \frac{1}{2\mu} \mathbf{P} \nabla f \end{aligned}$$

onde a matriz \mathbf{P} é denominada matriz de projeção, ou seja, para \mathbf{w} arbitrário \mathbf{Pw} é o subespaço tangente das restrições ativas, pois:

$$\mathbf{N}^T \mathbf{Pw} = \mathbf{N}^T [\mathbf{I} - \mathbf{N}(\mathbf{N}^T \mathbf{N})^{-1} \mathbf{N}^T] \mathbf{w} = 0$$

Após calcular \mathbf{s} , o método segue o fluxograma da Fig.4.3.12.

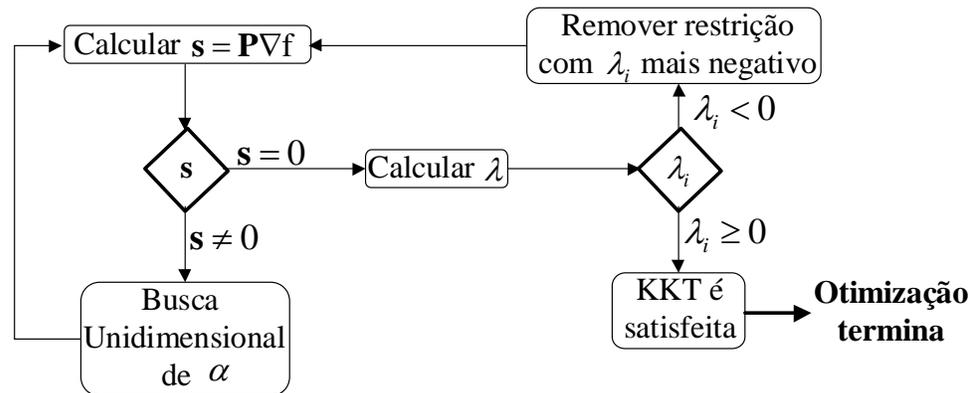


Fig.4.3.12: Fluxograma do método de gradientes reduzidos e projeção dos gradientes.

Note que se $\mathbf{s} = 0$, então as condições KKT podem ser satisfeitas. Assim, calculam-se os multiplicadores de Lagrange. Se todos $\lambda_j \geq 0$ as condições KKT são satisfeitas. Se algum $\lambda_j < 0$ é uma indicação de que embora não se possa mais obter progresso com o conjunto de restrições ativas corrente, ainda pode-se prosseguir removendo algumas das restrições associadas aos $\lambda_j < 0$. Em geral, remove-se a restrição associada ao λ_j mais negativo. Se $\mathbf{s} \neq 0$ realiza-se uma busca

unidimensional. Se s permanece igual à zero e ainda há $\lambda_j < 0$ remove-se outra restrição, até que todos $\lambda_j \geq 0$ e assim por diante.

Após cada busca unidimensional de α é necessário verificar o valor limite de α definido pelas restrições inativas, ou seja, o α obtido não pode desrespeitar as restrições que estavam inativas (nem as ativas lógico). Assim substituindo $\mathbf{x}_{i+1} = \mathbf{x}_i + \alpha \mathbf{s}$ em $\mathbf{g} = \mathbf{a}_j^T \mathbf{x} - b_j \geq 0$, temos:

$$g_j = \mathbf{a}_j^T (\mathbf{x}_i + \alpha \mathbf{s}) - b_j \geq 0 \Rightarrow \alpha \leq -\frac{(\mathbf{a}_j^T \mathbf{x}_i - b_j)}{\mathbf{a}_j^T \mathbf{s}} = -\frac{g_j(\mathbf{x}_i)}{\mathbf{a}_j^T \mathbf{s}}$$

que é válido se $\mathbf{a}_j^T \mathbf{s} < 0$, caso contrário não há valor limite superior de α devido à j -ésima restrição. O valor limite de α vale:

$$\bar{\alpha} = \min_{\substack{\alpha_j > 0 \\ \text{restr. ativas}}} \alpha_j$$

Vamos considerar agora restrições não-lineares. Essencialmente, essas restrições são linearizadas em torno de \mathbf{x}_i . No entanto, após realizar a busca unidimensional o valor de \mathbf{x} é movido pra fora da fronteira das restrições como ilustrado na Fig.4.3.13. Isso ocorre, porque movemos na direção tangente às restrições. Assim é necessário realizar uma correção após a busca unidimensional.

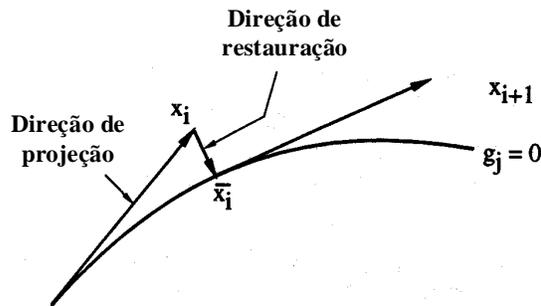


Fig.4.3.13: Direções de projeção e restauração.

Para obter a correção se utiliza a aproximação linear:

$$g_j \approx g_j(\mathbf{x}_i) + \nabla g_j^T (\bar{\mathbf{x}}_i - \mathbf{x}_i)$$

onde queremos $\bar{\mathbf{x}}_i - \mathbf{x}_i$ que reduza g_j a zero. Verifica-se que

$$\bar{\mathbf{x}}_i - \mathbf{x}_i = -\mathbf{N}(\mathbf{N}^T \mathbf{N})^{-1} \mathbf{g}_a(\mathbf{x}_i)$$

é a correção desejada, onde \mathbf{g}_a representa as restrições ativas (lembre-se que $\nabla g_j^T = \mathbf{N}^T$). Além dessa correção, a não-linearidade das restrições exige o cálculo de \mathbf{N} em cada ponto. Isso complica a determinação do valor limite de α que garante que as restrições inativas não serão desrespeitadas. Assim, o valor de α é determinado especificando uma redução desejada γ na função objetivo, ou seja:

$$f(\mathbf{x}_i) - f(\mathbf{x}_{i+1}) \cong \nabla f^T (\mathbf{x}_i - \mathbf{x}_{i+1}) \approx \gamma f(\mathbf{x}_i) \Rightarrow \alpha^* = -\frac{\gamma f(\mathbf{x}_i)}{\mathbf{s}^T \nabla f}$$

onde foi usado a equação $\mathbf{x}_{i+1} = \mathbf{x}_i + \alpha \mathbf{s}$. Dessa forma \mathbf{x}_{i+1} vale:

$$\mathbf{x}_{i+1} = \mathbf{x}_i + \alpha^* \mathbf{s} - \mathbf{N}(\mathbf{N}^T \mathbf{N})^{-1} \mathbf{g}_a$$

Vejam os um exemplo. Considere a solução do problema abaixo adotando como ponto inicial $\mathbf{x}_0 = (11,61;7,17) \Rightarrow f = 47,25$:

$$\begin{array}{l} \text{Min} \quad f = 3x_1 + \sqrt{3}x_2 \\ x_1, x_2 \\ \text{tal que} \quad g_1 = 3 - \frac{18}{x_1} - \frac{6\sqrt{3}}{x_2} \geq 0 \\ \quad \quad \quad g_2 = x_1 - 5.73 \geq 0 \\ \quad \quad \quad g_3 = x_2 - 7.17 \geq 0 \end{array}$$

Solução:

Nesse ponto temos duas restrições ativas: $g_1 = 0, g_3 = 0$. Assim, calculando-se os gradientes de f, g_1 e g_3 , temos:

$$\begin{aligned} \nabla f &= \begin{Bmatrix} 3 \\ \sqrt{3} \end{Bmatrix}; \quad \nabla g_1 = \begin{Bmatrix} 0,1335 \\ 0,2021 \end{Bmatrix}; \quad \nabla g_3 = \begin{Bmatrix} 0 \\ 1 \end{Bmatrix}; \quad \mathbf{N} = \begin{bmatrix} 0,1335 & 0 \\ 0,2021 & 1 \end{bmatrix} \Rightarrow \\ \Rightarrow \mathbf{P} = \mathbf{0} \Rightarrow \mathbf{s} = \mathbf{0} \Rightarrow \lambda &= \begin{Bmatrix} 22,47 \\ -2,798 \end{Bmatrix} \end{aligned}$$

onde foi usada a equação $\lambda = (\mathbf{N}^T \mathbf{N})^{-1} \mathbf{N}^T \nabla f$ para o cálculo dos multiplicadores de Lagrange. Como g_3 apresenta $\lambda_3 < 0$ deve ser eliminado do conjunto de restrições ativas. Então:

$$\mathbf{N} = \begin{bmatrix} 0,1335 \\ 0,2021 \end{bmatrix} \Rightarrow \mathbf{P} = \begin{bmatrix} 0,6962 & -0,46 \\ -0,46 & 0,3036 \end{bmatrix} \Rightarrow \mathbf{s} = -\mathbf{P} \nabla f = \begin{Bmatrix} -1,29 \\ 0,854 \end{Bmatrix}$$

Para 5% de redução na função objetivo ($\gamma=0,05$): $\alpha^* = \frac{0,05 \times 47,25}{\begin{bmatrix} -1,29 & 0,854 \end{bmatrix} \begin{Bmatrix} 3 \\ \sqrt{3} \end{Bmatrix}} = 0,988$

Como nenhuma das restrições estavam violadas, não há necessidade de se implementar a correção. Assim, continuando:

$$\mathbf{x}_1 = \mathbf{x}_0 + \alpha^* \mathbf{s} = \begin{Bmatrix} 11,61 \\ 7,17 \end{Bmatrix} + 0,988 \begin{Bmatrix} -1,29 \\ 0,854 \end{Bmatrix} = \begin{Bmatrix} 10,34 \\ 8,01 \end{Bmatrix} \Rightarrow f(\mathbf{x}_1) = 44,89; \quad g_1(\mathbf{x}_1) = -0,0382$$

A restrição g_1 está ativa (sendo violada). Portanto:

$$\mathbf{N} = \nabla g_1 = \begin{Bmatrix} 0,1684 \\ 0,1620 \end{Bmatrix} \Rightarrow \mathbf{P} = \begin{bmatrix} 0,4806 & -0,4996 \\ -0,4996 & 0,5194 \end{bmatrix} \Rightarrow \mathbf{s} = -\mathbf{P} \nabla f = \begin{Bmatrix} -0,5764 \\ 0,5991 \end{Bmatrix};$$

Devido à violação da restrição o coeficiente de redução é diminuído para 2,5% ($\gamma=0,025$):

$$\gamma = 0,025 \Rightarrow \alpha = -\frac{0,025 \times 44,89}{\begin{bmatrix} -0,5764 & 0,5991 \end{bmatrix} \begin{Bmatrix} 3 \\ \sqrt{3} \end{Bmatrix}} = 1,62$$

e também devido à violação ($g_a=-0,0382$) há necessidade da correção:

$$\mathbf{g}_a = -0,0382 \Rightarrow -\mathbf{N}(\mathbf{N}^T\mathbf{N})^{-1} \mathbf{g}_a = \begin{Bmatrix} 0,118 \\ 0,113 \end{Bmatrix} \Rightarrow \mathbf{x}_2 = \mathbf{x}_1 + \alpha^* \mathbf{s} - \mathbf{N}(\mathbf{N}^T\mathbf{N})^{-1} \mathbf{g}_a = \begin{Bmatrix} 10,34 \\ 8,01 \end{Bmatrix} +$$

$$-1,62 \begin{Bmatrix} 0,576 \\ -0,599 \end{Bmatrix} + \begin{Bmatrix} 0,118 \\ 0,113 \end{Bmatrix} = \begin{Bmatrix} 9,52 \\ 9,10 \end{Bmatrix} \Rightarrow f(\mathbf{x}_2) = 44,32; g_1(\mathbf{x}_2) = -0,0328$$

A solução ótima vale: $\mathbf{x}^* = \begin{Bmatrix} 9,464 \\ 9,464 \end{Bmatrix} \Rightarrow f(\mathbf{x}^*) = 44,78$.

b) Método das Direções Admissíveis

Ao contrário do método anterior, o método das direções viáveis procura ficar distante das fronteiras das restrições. No entanto, a sua iteração se inicia na fronteira do domínio viável. Se nenhuma restrição está ativa então métodos de minimização sem restrição são usados para obter uma direção.

Essencialmente esse método sempre procura caminhar numa direção útil (que melhora a função objetivo) e viável (respeita as restrições) como ilustrado na Fig.4.3.14.

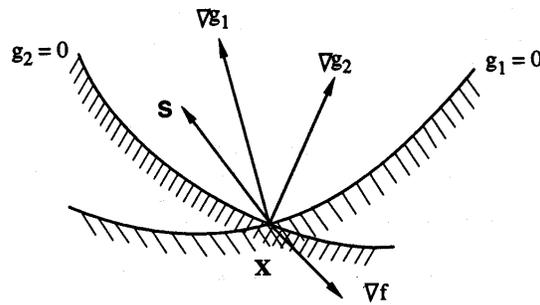


Fig.4.3.14: Seleção das direções de busca usando o método das direções viáveis.

Uma direção é viável se satisfaz:

$$\mathbf{s}^T \nabla g_j > 0$$

e uma direção é útil se satisfaz:

$$\mathbf{s}^T \nabla f < 0$$

Assim a direção \mathbf{s} é obtida através da solução do problema de otimização de programação linear:

$$\begin{array}{ll} \text{Min} & \beta \\ \text{s. } \beta & -\mathbf{s}^T \nabla g_j + \theta_j \beta \leq 0 \\ \text{tal que} & \mathbf{s}^T \nabla f + \beta \leq 0 \quad \theta_j \geq 0 \\ & |s_i| \leq 1 \end{array}$$

Os coeficientes θ_j determinam o quanto \mathbf{x} irá se afastar da fronteira das restrições. Um valor de $\theta_j=0$ resultará num movimento tangente à fronteira das restrições, sendo apropriado para restrições lineares. Um valor grande de θ_j resultará num grande ângulo entre a fronteira da restrição e a direção \mathbf{s} , sendo apropriado para uma restrição altamente não-linear.

O problema acima é resolvido usando o método Simplex. Se $\beta_{\max} > 0$ uma direção viável e útil foi encontrada. Se $\beta_{\max} = 0$ pode ser mostrado que as condições KKT são satisfeitas.

Vejam um exemplo. Considere a solução do problema abaixo, adotando como ponto inicial $\mathbf{x}_0 = (11,61; 7,17) \Rightarrow f = 47,25$:

$$\begin{aligned} \text{Min} \quad & f = 3x_1 + \sqrt{3}x_2 \\ & x_1, x_2 \\ \text{tal que} \quad & g_1 = 3 - \frac{18}{x_1} - \frac{6\sqrt{3}}{x_2} \geq 0 \\ & g_2 = x_1 - 5.73 \geq 0 \\ & g_3 = x_2 - 7.17 \geq 0 \end{aligned}$$

Solução:

Nesse ponto temos duas restrições ativas: $g_1 = 0, g_3 = 0$. Assim, calculando-se os gradientes de f, g_1 e g_3 , temos:

$$\nabla f = \begin{Bmatrix} 3 \\ \sqrt{3} \end{Bmatrix}; \quad \nabla g_1 = \begin{Bmatrix} 0,1335 \\ 0,2021 \end{Bmatrix}; \quad \nabla g_3 = \begin{Bmatrix} 0 \\ 1 \end{Bmatrix}$$

Selecionando-se $\theta_1 = \theta_2 = 1$, temos:

$$\begin{aligned} \text{Min} \quad & \beta \\ & s_1, s_2, \beta \\ \text{tal que} \quad & -0,1335s_1 - 0,2021s_2 + \beta \leq 0 \\ & -s_2 + \beta \leq 0 \\ & 3s_1 + \sqrt{3}s_2 + \beta \leq 0 \\ & -1 \leq s_1 \leq 1 \\ & -1 \leq s_2 \leq 1 \end{aligned}$$

Cuja solução vale: $s_1 = -0,6172; s_2 = 1$

Realizando uma busca unidimensional, temos:

$$\mathbf{x}_1 = \begin{Bmatrix} 11,61 \\ 7,17 \end{Bmatrix} + \alpha \begin{Bmatrix} -0,6172 \\ 1 \end{Bmatrix}$$

Como a função objetivo é linear o valor de α será limitado pelas restrições. A exigência que g_2 não seja violado nos leva à $\alpha = 9,527, x_1 = 5,73, x_2 = 16,7$ o que resulta na violação de g_1 . Verifica-se que para $\alpha < 5,385 (x_1 = 8,29; x_2 = 12,56 \Rightarrow f = 46,62)$, g_1 nem g_2 são violados. A Fig.4.3.15 ilustra o que ocorre.

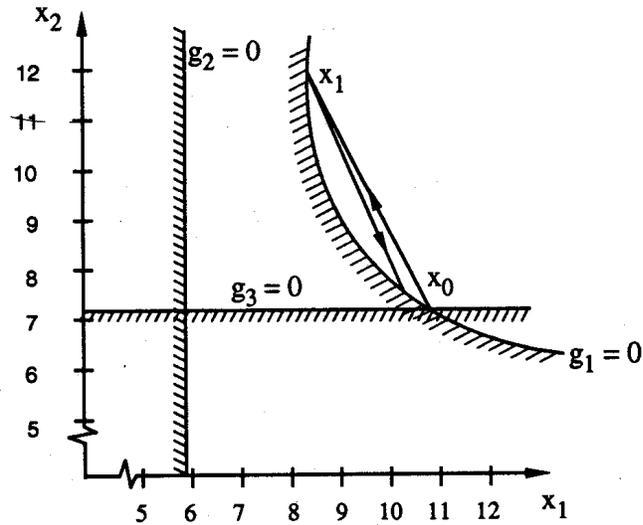


Fig.4.3.15: Direções viáveis.

Na próxima iteração temos apenas uma restrição ativa (g_1). Assim, calculando-se os gradientes:

$$\nabla g_1 = \begin{Bmatrix} -\frac{18}{x_1^2} \\ 6\sqrt{3} \\ -\frac{x_2^2}{x_1^2} \end{Bmatrix} = \begin{Bmatrix} 0,2619 \\ 0,0659 \end{Bmatrix}; \quad \nabla f = \begin{Bmatrix} 3 \\ 3\sqrt{3} \end{Bmatrix}$$

Selecionando-se os mesmos valores de θ_1 e θ_2 , obtém-se o problema de otimização:

$$\begin{array}{l} \text{Min } \beta \\ s_1, s_2, \beta \\ \text{tal que } \end{array} \begin{array}{l} -0,2619s_1 - 0,0659s_2 + \beta \leq 0 \\ 3s_1 + \sqrt{3}s_2 + \beta \leq 0 \\ -1 \leq s_1 \leq 1 \\ -1 \leq s_2 \leq 1 \end{array}$$

Cuja solução vale: $s_1 = 0,5512$; $s_2 = -1$

Realizando uma busca unidimensional, temos:

$$\mathbf{x}_2 = \begin{Bmatrix} 8,29 \\ 12,56 \end{Bmatrix} + \alpha \begin{Bmatrix} 0,5512 \\ -1 \end{Bmatrix}$$

onde realizando a análise anterior α é limitado em $\alpha < 4,957$. Assim:

$$\alpha = 4,957 \Rightarrow \mathbf{x}_2 = \begin{Bmatrix} 11,02 \\ 7,60 \end{Bmatrix} \Rightarrow f(\mathbf{x}_2) = 46,22$$

A solução ótima vale: $\mathbf{x}^* = \begin{Bmatrix} 9,464 \\ 9,464 \end{Bmatrix} \Rightarrow f(\mathbf{x}^*) = 44,78$.

4.3.2.8 Métodos Indiretos

Esses métodos transformam um problema de otimização com restrição num problema de otimização sem restrição. As restrições são substituídas por penalizações que dependem do grau com que a restrição é violada. As penalizações devem ser altas o suficiente de forma que as restrições sejam ligeiramente violadas. No entanto a imposição de altas penalizações logo de início resultam em problemas numéricos, assim se utiliza uma abordagem gradual em que se inicia com pequenas penalizações que vão sendo aumentadas gradualmente. Entre esses métodos temos o método das Penalizações e o método de Lagrange Aumentado.

a) Método de Penalização Exterior

Considere o problema de otimização abaixo:

$$\begin{aligned} & \text{Min}_{\mathbf{x}} \quad f(\mathbf{x}) \\ & \text{tal que} \quad h_i(\mathbf{x}) = 0, \\ & \quad \quad \quad i = 1, \dots, n_e \\ & \quad \quad \quad g_j(\mathbf{x}) \geq 0, \\ & \quad \quad \quad j = 1, \dots, n_g \end{aligned}$$

No método de penalização exterior, esse problema é transformado no seguinte problema de otimização sem restrição:

$$\begin{aligned} & \text{Min}_{\mathbf{x}} \quad \phi(\mathbf{x}, r) = f(\mathbf{x}) + r \sum_{i=1}^{n_e} h_i^2(\mathbf{x}) + r \sum_{j=1}^{n_g} \langle -g_j \rangle^2 \\ & \quad \quad \quad r = r_1, r_2, \dots, r_i \rightarrow \infty \end{aligned}$$

onde $\langle a \rangle = \max(a, 0)$ e o valor de r_i aumenta gradualmente com as iterações. Como exemplo numérico, considere o problema abaixo:

$$\begin{aligned} & \text{Min}_{x} \quad \frac{x}{2} \\ & \text{tal que} \quad x - 4 \geq 0 \end{aligned}$$

que é transformado no problema abaixo:

$$\text{Min}_{x} \quad \frac{x}{2} + r \langle 4 - x \rangle^2$$

onde $r = 0, 0,1, 0,5, 2,5, \dots \rightarrow \infty$. A Fig.4.3.16 mostra a plotagem da função $\phi(x)$ e ilustra o que acontece a medida que aumentamos o valor de r .

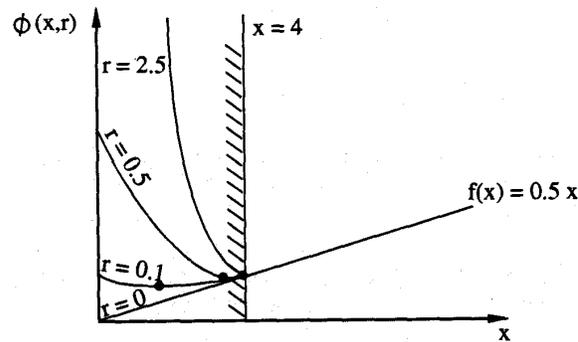


Fig.4.3.16: Função de penalização exterior.

Nota-se que a curva $\phi(x)$ apresenta o seu ponto mínimo junto a fronteira da restrição, reproduzindo assim o resultado do problema de otimização inicial. Para um alto valor de r a função $\phi(x)$ apresenta uma curvatura acentuada no ponto mínimo. Essa curvatura acentuada causa um mau condicionamento numérico se tentarmos logo de início calcular o mínimo de $\phi(x)$ com um alto valor de r . Aumentando o valor de r gradualmente, de forma que para cada valor de r o ponto ótimo calculado na iteração anterior é usado como ponto de partida nas iterações para o cálculo do ponto ótimo do valor de r corrente, contorna-se o problema do condicionamento numérico.

O valor de r é aumentado até que se atinja a convergência da solução ótima. Existem vários critérios de convergência para a solução ótima, vejamos alguns deles:

- $\| \mathbf{x}^* - \mathbf{a} \| \leq \varepsilon_1$ onde $\mathbf{a} = \frac{c\mathbf{x}^*(r_{i-1}) - \mathbf{x}^*(r_i)}{c-1}$ e $c = \frac{r_{i-1}}{r_i}$. r_i e r_{i-1} são os dois últimos valores de r ;
- $\left| \frac{\phi - f}{f} \right| \leq \varepsilon_2$, uma vez que $\phi \rightarrow f$ a medida que $r_i \rightarrow \infty$;
- $\left| \frac{f^*(r_i) - f^*(r_{i-1})}{f^*(r_i)} \right| \leq \varepsilon_3$, que é baseado na mudança do valor mínimo da função objetivo.

Vejamos um exemplo de solução de um problema no espaço bidimensional. Seja o problema:

$$\begin{aligned} \text{Min} \quad & f = x_1^2 + 10x_2^2 \\ & x_1, x_2 \\ \text{tal que} \quad & x_1 + x_2 = 4 \end{aligned}$$

Solução:

O problema transformado fica:

$$\text{Min}_{x_1, x_2} \phi = x_1^2 + 10x_2^2 + r(4 - x_1 - x_2)^2$$

Como se trata de um problema de minimização sem restrição, devemos impor a condição de estacionaridade para obter a solução ótima:

$$\nabla \phi = \begin{Bmatrix} 2x_1(1+r) + 2rx_2 - 8r \\ 2x_2(10+r) + 2rx_1 - 8r \end{Bmatrix} \Rightarrow \nabla \phi = 0$$

isolando-se os valores de x_1 e x_2 , temos: $\mathbf{x}^* = \begin{Bmatrix} \frac{40r}{10+11r} \\ \frac{4r}{10+11r} \end{Bmatrix}$

A tabela abaixo mostra a solução \mathbf{x}^* para valores crescentes de r .

Tabela 4.3.2: Soluções para valores crescentes de r .

r	x_1	x_2	f	ϕ
1	1,905	0,1905	3,992	7,619
10	3,333	0,3333	12,220	13,333
100	3,604	0,3604	14,288	14,144
1000	3,633	0,3633	14,518	14,532

A medida que r aumenta a solução converge para a solução ótima dada por:

$$\mathbf{x}^* = \begin{Bmatrix} 3,636 \\ 0,3636 \end{Bmatrix} \Rightarrow f^* = 14,54$$

O mau condicionamento de se utilizar um valor de r grande logo no início pode ser medido através da matriz Hessiana de $\phi(x_1, x_2)$, dada por:

$$\mathbf{H} = \begin{bmatrix} 2+2r & 2r \\ 2r & 20+2r \end{bmatrix}$$

Assim, para grandes valores de r todos os componentes ficam aproximadamente iguais à $2r$, tornando a matriz mal condicionada.

Vejamos agora a interpretação gráfica da solução de um problema bidimensional. A Fig.4.3.17 mostra a região do domínio viável e as curvas de nível da função objetivo do problema abaixo.

Min $f(\mathbf{x}) = x_1 + x_2$
 x_1, x_2
tal que $g_1(\mathbf{x}) = -2 + x_1 - 2x_2 \leq 0$
 $g_2(\mathbf{x}) = 8 - 6x_1 + x_1^2 - x_2 \leq 0$

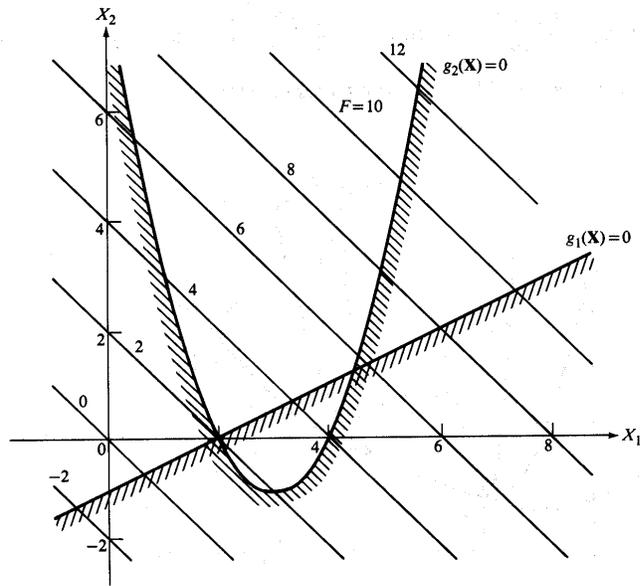


Fig.4.3.17: Região do domínio viável.

A Fig.4.3.18 ilustra a função $\phi(x_1, x_2)$ para $r=0,05$.

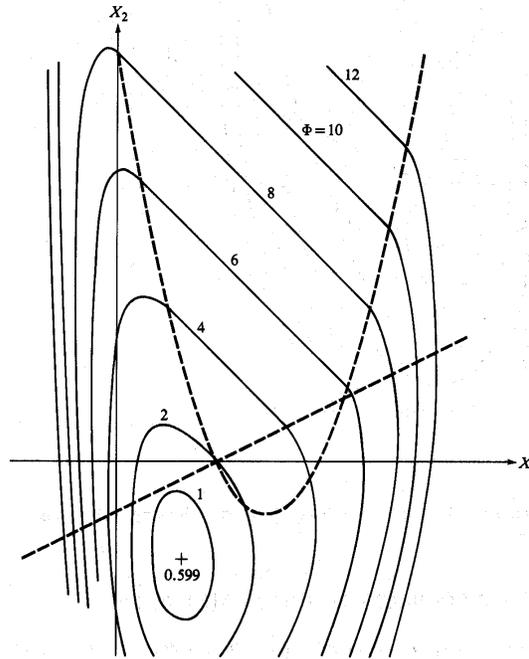


Fig.4.3.18: Função $\phi(x_1, x_2)$ para $r=0,05$.

A Fig.4.3.19 ilustra a função $\phi(x_1, x_2)$ para $r=0,1$.

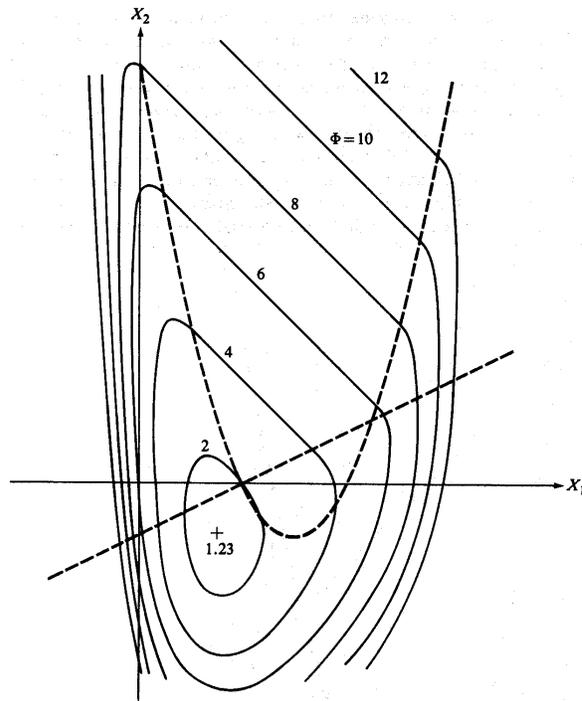


Fig.4.3.19: Função $\phi(x_1, x_2)$ para $r=0,1$.

A Fig.4.3.19 ilustra a função $\phi(x_1, x_2)$ para $r=1,0$.

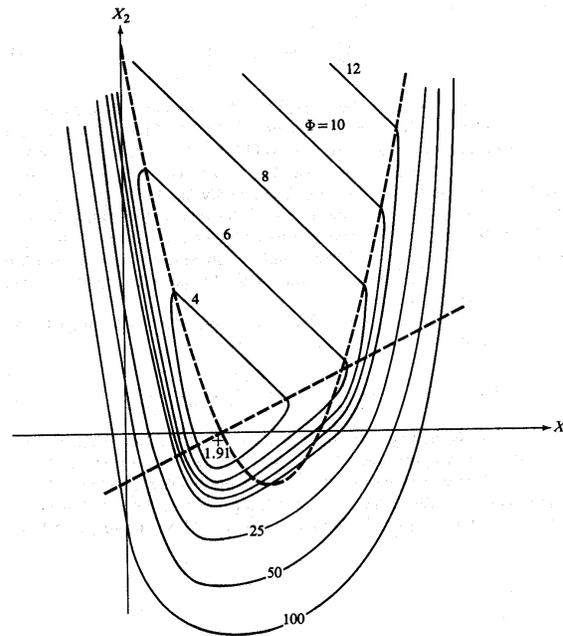


Fig.4.3.19: Função $\phi(x_1, x_2)$ para $r=1,0$.

Nas figuras acima podemos verificar que a medida que aumentamos r o ponto mínimo de ϕ se aproxima do ponto mínimo real.

b) Método de Penalização Interior

A grande desvantagem do método anterior é que as restrições somente contribuem para a função objetivo quando são violadas. Dessa forma se a minimização termina antes que r atinja um valor alto (por exemplo, por falta de espaço em disco), o resultado será inútil, pois não respeita as restrições.

Quando somente restrições de desigualdade estão presentes, é possível definir uma função de penalização interior que mantém a solução das iterações no interior do domínio de projeto. Assim, o problema abaixo:

$$\begin{aligned} & \text{Min}_{\mathbf{x}} \quad f(\mathbf{x}) \\ & \text{tal que } g_j(\mathbf{x}) \geq 0, \\ & \quad \quad \quad j = 1, \dots, n_g \end{aligned}$$

pode ser transformado no seguinte problema sem restrição:

$$\begin{aligned} & \text{Min}_{\mathbf{x}} \quad \phi(\mathbf{x}, r) = f(\mathbf{x}) + r \sum_{j=1}^{n_g} \frac{1}{g_j(\mathbf{x})} \\ & \quad \quad \quad r = r_1, r_2, \dots, r_i \rightarrow 0, \quad r_i > 0 \end{aligned}$$

Vejamos novamente o exemplo anterior:

$$\begin{aligned} & \text{Min}_{\mathbf{x}} \quad \frac{x}{2} \\ & \text{tal que } x - 4 \geq 0 \end{aligned}$$

que agora é transformado no problema abaixo:

$$\text{Min}_{\mathbf{x}} \quad \frac{x}{2} + \frac{r}{x-4}$$

onde $r = 10, 1, 0,1, \dots \rightarrow 0$. A Fig.4.3.20 mostra a plotagem da função $\phi(x)$ e ilustra o que acontece a medida que diminuimos o valor de r .

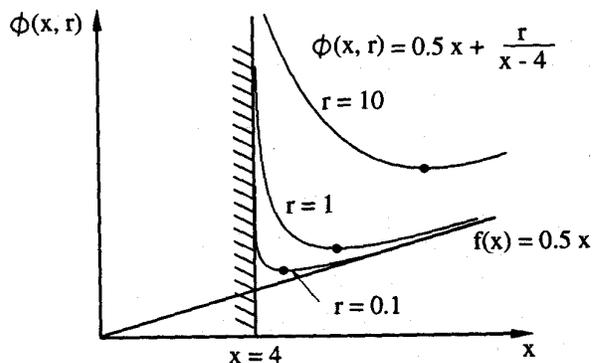


Fig.4.3.20: Função penalizadora interior.

Assim, observamos que a medida que diminuimos o valor de r a função se aproxima da barreira da restrição.

Uma outra opção de função penalizadora é dada por:

$$\phi(\mathbf{x}, r) = f(\mathbf{x}) - r \sum_{j=1}^{n_g} \log(g_j(\mathbf{x}))$$

$$r = r_1, r_2, \dots, r_i \rightarrow 0, \quad r_i > 0$$

Também é possível incluir restrições de igualdade através da expressão:

$$\phi(\mathbf{x}, r) = f(\mathbf{x}) + r \sum_{j=1}^{n_g} \frac{1}{g_j(\mathbf{x})} + r^{-1/2} \sum_{i=1}^{n_e} h_i^2(\mathbf{x})$$

$$r = r_1, r_2, \dots, r_i \rightarrow 0, \quad r_i > 0$$

Entretanto apresenta como desvantagem o fato de que exige um ponto de partida dentro do domínio viável, o que em geral é difícil de se obter.

c) Método de Penalização Interior Estendido

Ambos os métodos anteriores apresentam vantagens e desvantagens. O método de penalização exterior não exige que o ponto de partida esteja no interior do domínio viável, mas fornece um resultado inútil caso a minimização pare antes que r atinja valores altos. Já o método de penalização interior não apresenta esse problema, mas exige que o ponto de partida esteja dentro do domínio viável o que é difícil de se obter. Assim o método de penalização interior estendido combina a vantagem de ambos os métodos.

Considere o problema de otimização abaixo (com somente restrições de desigualdade para facilitar):

$$\begin{array}{l} \text{Min} \quad f(\mathbf{x}) \\ \mathbf{x} \\ \text{tal que } g_j(\mathbf{x}) \geq 0, \\ \quad \quad \quad j = 1, \dots, n_g \end{array}$$

Esse problema pode ser transformado no seguinte problema sem restrição:

$$\begin{array}{l} \text{Min} \quad \phi(\mathbf{x}, r) = f(\mathbf{x}) + r \sum_{j=1}^{n_g} p(g_j) \\ \mathbf{x} \\ \text{e: } p(g_j) = \begin{cases} \frac{1}{g_j}, & \text{se } g_j \geq g_0 \\ \frac{1}{g_0} \left[3 - 3 \left(\frac{g_j}{g_0} \right) + \left(\frac{g_j}{g_0} \right)^2 \right], & \text{se } g_j < g_0 \end{cases} \\ \text{onde: } g_0 = cr^{1/2} \text{ e } c \text{ é uma constante} \\ r = r_1, r_2, \dots, r_i \rightarrow 0, \quad r_i > 0 \end{array}$$

Considerando o mesmo exemplo anterior:

$$\begin{array}{l} \text{Min} \quad \frac{x}{2} \\ \mathbf{x} \\ \text{tal que } x - 4 \geq 0 \end{array}$$

obtemos o problema:

$$\begin{aligned} & \text{Min}_x \quad \frac{x}{2} + rp(g) \\ & r = 10, 1, 0,1, \dots \rightarrow 0 \end{aligned}$$

$$\text{onde: } p(g) = \begin{cases} \frac{1}{x-4}, & \text{se } x - 4 \geq g_0 \\ \frac{1}{g_0} \left[3 - 3 \left(\frac{x-4}{g_0} \right) + \left(\frac{x-4}{g_0} \right)^2 \right], & \text{se } x - 4 < g_0 \end{cases}$$

A Fig.4.3.21 ilustra a função $\phi(x)$.

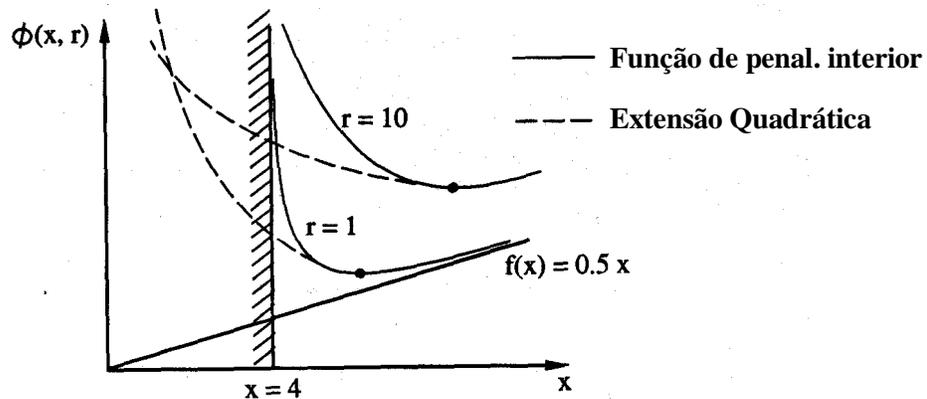


Fig.4.3.21: Função de penalização interior estendida.

Uma vez obtido o problema de otimização sem restrição por qualquer um dos métodos acima, os termos de penalização fazem com que a função $\phi(\mathbf{x})$ tenha uma alta curvatura próxima ao contorno da restrição, mesmo que a curvatura da restrição e função objetivo sejam pequenas. Como o conceito de curvatura envolve as derivadas segundas da função, esse efeito permite obter aproximações com baixo custo computacional da matriz Hessiana de $\phi(\mathbf{x})$. Isso permite que se possa usar o método de Newton na solução do problema sem ter um alto custo no cálculo das derivadas segundas. Essa abordagem é mais atrativa do que usar o método de Quasi-Newton (onde a matriz Hessiana é aproximada utilizando-se derivadas primeiras) porque no método de Newton o número de iterações é independente do número de variáveis de projeto, enquanto no método de Quasi-Newton o número de iterações é proporcional ao número de variáveis de projeto (ver seção 4.3.1.3).

Assim, considerando o método de penalização exterior com restrições de igualdade:

$$\phi(\mathbf{x}, r) = f(\mathbf{x}) + r \sum_{i=1}^{n_c} h_i^2(\mathbf{x})$$

A segunda derivada de $\phi(\mathbf{x})$ é dada por:

$$\frac{\partial^2 \phi}{\partial x_k \partial x_l} = \frac{\partial^2 f}{\partial x_k \partial x_l} + r \sum_{i=1}^{n_c} 2 \left(\frac{\partial h_i}{\partial x_k} \frac{\partial h_i}{\partial x_l} + h_i \frac{\partial^2 h_i}{\partial x_k \partial x_l} \right)$$

O termo $\frac{\partial^2 f}{\partial x_k \partial x_1}$ pode ser desprezado em relação aos altos valores de r nos últimos estágios da otimização e no termo $h_i \frac{\partial^2 h_i}{\partial x_k \partial x_1}$, h_i é próximo de zero, especialmente nos últimos estágios da otimização, e portanto também pode ser desprezado. Dessa forma:

$$\frac{\partial^2 \phi}{\partial x_k \partial x_1} \cong r \sum_{i=1}^{n_e} 2 \left(\frac{\partial h_i}{\partial x_k} \frac{\partial h_i}{\partial x_1} \right)$$

Para o método de penalização interior a situação é similar. Considerando a função:

$$\phi(\mathbf{x}, r) = f(\mathbf{x}) + r \sum_{i=1}^{n_g} \frac{1}{g_j(\mathbf{x})}$$

A segunda derivada de $\phi(\mathbf{x})$ é dada por:

$$\frac{\partial^2 \phi}{\partial x_k \partial x_1} = \frac{\partial^2 f}{\partial x_k \partial x_1} + r \sum_{j=1}^{n_g} \frac{1}{g_j^3} \left(2 \frac{\partial g_j}{\partial x_k} \frac{\partial g_j}{\partial x_1} - g_j \frac{\partial^2 g_j}{\partial x_k \partial x_1} \right)$$

O termo $\frac{\partial^2 f}{\partial x_k \partial x_1}$ pode ser desprezado em relação aos demais termos, pois $r/g_j^3 \rightarrow \infty$ nos últimos estágios da otimização quando r tende a zero e no termo $g_j \frac{\partial^2 g_j}{\partial x_k \partial x_1}$, g_j é próximo de zero, especialmente nos últimos estágios da otimização e portanto esse termo também pode ser desprezado em relação ao termo de primeiras derivadas. Dessa forma:

$$\frac{\partial^2 \phi}{\partial x_k \partial x_1} \cong r \sum_{j=1}^{n_g} \frac{2}{g_j^3} \left(\frac{\partial g_j}{\partial x_k} \frac{\partial g_j}{\partial x_1} \right)$$

Para ter um idéia da eficiência computacional que isso representa consideremos o problema abaixo em que se deseja minimizar a massa da asa de um avião tendo restrições de tensões mecânicas e deslocamento:

Min \mathbf{x} Massa
 tal que (restrições de deslocamentos e tensões mecânicas)

A Fig.4.3.22 ilustra o problema.

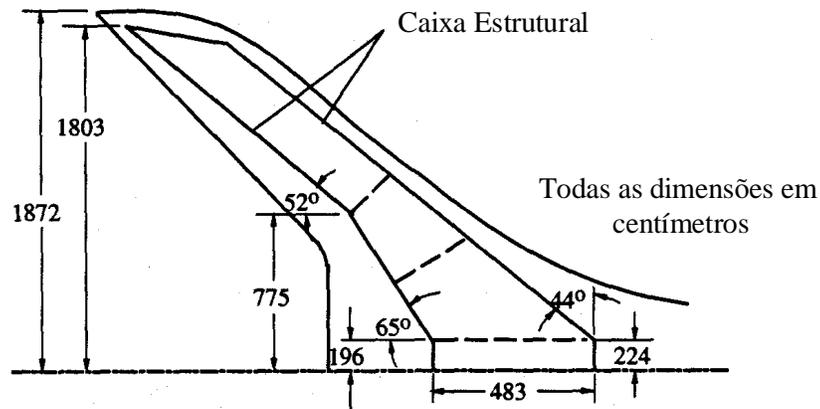


Fig.4.3.22: Asa de avião a ser otimizada.

A asa foi modelada através de um modelo de elementos finitos com 67 nós e 290 elementos. O número de variáveis de projeto (dimensões da asa) foi variado de 13 à 146. A aproximação apresentada acima foi utilizada juntamente com o método de Newton na solução do problema. A tabela 4.3.3 mostra o efeito do número de variáveis de projeto no número de iterações.

Tabela 4.3.3: Resultados da Otimização

Nº de variáveis de projeto	Tempo de CPU (s) (CDC6600)	Total nº de iterações na otimização	Total nº de análises	Massa Final (Kg)
13	142	4	21	887,3
25	217	4	19	869,1
32	293	5	22	661,7
50	460	5	25	658,2
74	777	5	28	648,6
146	1708	5	26	513,0

Note que o número de iterações é aproximadamente constante (em torno de 5) usando esse método. Com o método de Quasi-Newton, esse número seria igual ou próximo ao número de variáveis de projeto.

d) Método de Lagrange Aumentado

O método de Lagrange Aumentado combina o uso de multiplicadores de Lagrange com as funções penalizadoras. Quando somente multiplicadores de Lagrange são empregados o ponto é um ponto estacionário (ponto sela - ver seção 4.3.2.4) ao invés de um mínimo da função Lagrangeana. Quando somente funções penalizadoras são empregadas há um mínimo, mas também problemas de mau condicionamento para encontrar o mínimo. Combinando ambos os métodos obtém-se um problema sem restrição onde a função a ser minimizada não apresenta mau condicionamento.

Assim, considerando o problema:

$$\begin{aligned} \text{Min}_{\mathbf{x}} \quad & f(\mathbf{x}) \\ \text{tal que} \quad & h_j(\mathbf{x}) = 0, \\ & j = 1, \dots, n_g \end{aligned}$$

A função Lagrange aumentada fica:

$$\text{Min}_{\mathbf{x}} L(\mathbf{x}, \lambda, r) = f(\mathbf{x}) - \sum_{j=1}^{n_g} \lambda_j h_j(\mathbf{x}) + r \sum_{j=1}^{n_g} h_j^2(\mathbf{x})$$

Se $\lambda_j = 0$ obtemos o método de penalização exterior comum. Por outro lado, se os valores corretos, ou bem estimados, de λ_j são usados, obtém-se a solução ótima para qualquer valor de r , não necessariamente alto e portanto não há problema de condicionamento numérico. No entanto não sabemos os valores corretos dos multiplicadores de Lagrange, assim a idéia do método de Lagrange Aumentado é estimar esses multiplicadores. Se a estimativa for boa, podemos obter o ponto ótimo sem usar altos valores de r . O valor de r somente precisa ser grande o suficiente de forma que o Lagrangeano Aumentado (L) tenha um mínimo ao invés de um ponto estacionário no ponto ótimo. Para obter uma estimativa dos multiplicadores de Lagrange consideremos a condição de estacionaridade de L acima:

$$\frac{\partial L}{\partial x_i} = \frac{\partial f}{\partial x_i} - \sum_{j=1}^{n_g} (\lambda_j - 2rh_j) \frac{\partial h_j}{\partial x_i} = 0$$

Com os multiplicadores corretos λ_j^* :

$$\frac{\partial f}{\partial x_i} - \sum_{j=1}^{n_g} \lambda_j^* \frac{\partial h_j}{\partial x_i} = 0$$

Comparando as equações espera-se que a medida que o mínimo se aproxima:

$$\lambda_j - 2rh_j \rightarrow \lambda_j^*$$

Assim, foi sugerida a seguinte regra de iteração:

$$\lambda_j^{(k+1)} = \lambda_j^{(k)} - 2r^{(k)} h_j^{(k)}$$

onde k é o número da iteração.

Como exemplo considere a solução do problema de otimização a seguir:

$$\begin{aligned} \text{Min}_{\mathbf{x}} \quad & f(\mathbf{x}) = x_1^2 + 10x_2^2 \\ \text{tal que} \quad & h(\mathbf{x}) = x_1 + x_2 - 4 = 0 \end{aligned}$$

Solução:

O Lagrange Aumentado vale: $f(\mathbf{x}) = x_1^2 + 10x_2^2 - \lambda(x_1 + x_2 - 4) + r(x_1 + x_2 - 4)^2$

As condições de estacionaridade são:

$$\left. \begin{aligned} 2x_1 - \lambda + 2r(x_1 + x_2 - 4) &= 0 \\ 20x_2 - \lambda + 2r(x_1 + x_2 - 4) &= 0 \end{aligned} \right\} \Rightarrow x_1 = 10x_2 = \frac{5\lambda + 40r}{10 + 11r}$$

Assim, iniciando com $r_0 = 1$ e $\lambda = 0 \Rightarrow \mathbf{x}^{(1)} = (1,905; 0,1905)^T$ e $h = -1,905 \Rightarrow \lambda^{(1)} = -2 \times 1 \times (-1,905) = 3,81$

Repetindo a otimização com $r^{(1)} = 10$, $\lambda^{(1)} = 3,81 \Rightarrow \mathbf{x}^{(2)} = (3,492; 0,3492)^T$, $h = -0,1587$

Para o mesmo valor de r usando penalização exterior foi obtido $\mathbf{x}^{(2)} = (3,333; 0,3333)^T$, enquanto que o ótimo vale $\mathbf{x}^* = (3,636; 0,3636)^T$. Portanto estamos mais próximos do ótimo.

A estimativa do multiplicador vale: $\lambda^{(2)} = 3,81 - 2 \times 10 \times (-0,1587) = 6,984$

Continuando com $r^{(3)} = 10$, $\lambda^{(3)} = 6,984 \Rightarrow \mathbf{x}^{(3)} = (3,624; 0,3624)^T$, $h = -0,0136$.

Mostrando que uma boa convergência pode ser obtida sem aumentar r . Para o caso em que há restrições de desigualdade:

$$\begin{array}{l} \text{Min } f(\mathbf{x}) \\ \mathbf{x} \\ \text{tal que } g_j(\mathbf{x}) \geq 0, \\ j = 1, \dots, n_g \end{array}$$

a função Lagrange aumentada fica:

$$L(\mathbf{x}, \lambda, r) = f(\mathbf{x}) + r \sum_{j=1}^{n_g} \left\langle \frac{\lambda_j}{2r} - g_j \right\rangle^2 \quad \text{onde } \langle a \rangle = \max(a, 0)$$

A condição de estacionaridade de L vale:

$$\frac{\partial f}{\partial x_i} - 2r \sum_{j=1}^{n_g} \left\langle \frac{\lambda_j}{2r} - g_j \right\rangle \frac{\partial g_j}{\partial x_i} = 0$$

e a condição exata vale:

$$\frac{\partial f}{\partial x_i} - \sum_{j=1}^{n_g} \lambda_j^* \frac{\partial g_j}{\partial x_i} = 0 \quad \text{além de } \lambda_j^* g_j = 0$$

Assim, a estimativa do multiplicador pode ser dada por:

$$\lambda_j^* = \max(\lambda_j - 2rg_j, 0)$$

A rotina do SCILAB que resolve problemas de otimização não-lineares com restrições é a SOLNP (veja como utilizá-la na apostila em anexo). Essa rotina é baseada no método de Lagrange Aumentado descrito acima.

4.4 Método de Programação Linear Sequencial (PLS ou "SLP")

Os dois métodos apresentados a seguir fazem parte de um grupo de métodos de solução chamado aproximação seqüencial cuja descrição será continuada na seção 7.1. Esses métodos permitem a integração prática de um software comercial de análise como elementos finitos com um algoritmo de otimização para solução de problemas genéricos em otimização. Os softwares de otimização disponíveis com os algoritmos descritos anteriormente, muitas vezes, exigem que hajam subrotinas que avaliem a função objetivo e restrições. Quando a função objetivo e restrições são resultados de uma análise de elementos finitos por exemplo, é muito difícil transformar o pacote comercial de MEF numa subrotina a ser chamada pelo programa de otimização se não se tem acesso ao código fonte do programa. Assim nesses casos a integração com o software de otimização é feita através de métodos de aproximação seqüencial. O PLS (ou "SLP" em inglês) será o primeiro a ser estudado.

Considere um problema não-linear de otimização:

$$\begin{array}{ll} \text{Min} & f(\mathbf{x}) \\ \mathbf{x} & g_j(\mathbf{x}) \geq 0, \\ \text{tal que} & j = 1, \dots, n_g \\ & \mathbf{x}_{\min} \leq \mathbf{x} \leq \mathbf{x}_{\max} \end{array}$$

A idéia básica do PLS para solucionar o problema acima é aproximá-lo por sucessivos subproblemas lineares de otimização que podem ser resolvidos usando métodos de programação linear (Simplex). Essa aproximação é obtida aproximando-se as funções objetivo e restrição por séries de Taylor e limitando a variação do valor de cada variável de projeto em cada subproblema linear através dos chamados limites móveis. Assim, obtemos a seguinte formulação para cada subproblema linear:

$$\begin{array}{ll} \text{Min} & f(\mathbf{x}_0) + \sum_{i=1}^n (x_i - x_{0i}) \left(\frac{\partial f}{\partial x_i} \right) \Big|_{\mathbf{x}_0} \\ \mathbf{x} & \\ \text{tal que} & g_j(\mathbf{x}_0) + \sum_{i=1}^n (x_i - x_{0i}) \left(\frac{\partial g_j}{\partial x_i} \right) \Big|_{\mathbf{x}_0} \geq 0 \\ & j = 1, \dots, n_g \\ & a_{li} \leq x_i - x_{0i} \leq a_{ui} \end{array}$$

onde $\left(\frac{\partial f}{\partial x_i} \right) \Big|_{\mathbf{x}_0}$ e $\left(\frac{\partial g_j}{\partial x_i} \right) \Big|_{\mathbf{x}_0}$ são denominados sensibilidade (ou gradientes) das funções $f(\mathbf{x})$ e $g_j(\mathbf{x})$, respectivamente.

Para entendermos melhor o conceito do método, vamos considerar o problema de encontrar o máximo de uma função unidimensional $f(x)$. A fig.4.4.1 ilustra a função.

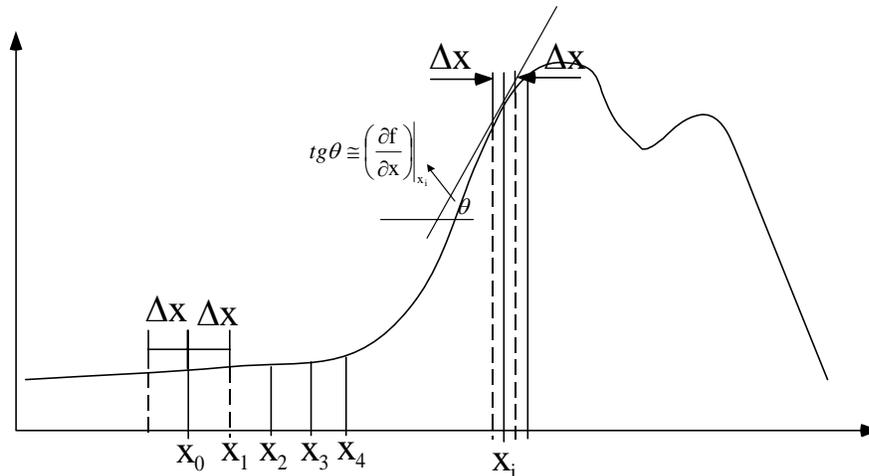


Fig.4.4.1: Procedimento sequencial para encontrar o máximo da função.

Consideremos um ponto x_0 como ponto de partida e a aproximação da função $f(x)$ em torno de x_0 usando uma série de Taylor:

$$f(x) \cong f(x_0) + (x - x_0) \left(\frac{\partial f}{\partial x} \right) \Big|_{x_0} + \varepsilon(\Delta x^2)$$

A equação acima consiste na equação de uma reta. Agora consideremos o problema de encontrar o máximo, não da função $f(x)$, mas de sua aproximação (reta) (ver Fig.4.4.1). Logicamente a aproximação de reta somente é válida próxima ao ponto x_0 , distanciando-se muito do valor real da função a medida que nos afastamos de x_0 . Assim, nesse problema de otimização devemos definir os valores máximos e mínimos em que a aproximação é válida (Δx), que são os chamados *limites móveis* de x_0 . Obviamente a solução desse problema de otimização estará num dos extremos dos limites móveis, por exemplo $x_1 = x_0 + \Delta x$. Na próxima iteração tomamos como novo ponto de partida a solução do problema anterior, ou seja, x_1 , e assim por diante. As iterações procedem com x_i sendo substituído por x_{i+1} , até que ocorre a convergência da solução, definido, por exemplo, pela variação do valor da função objetivo.

Um problema fácil de se identificar é a definição dos limites móveis. Se forem muito grandes, o erro da aproximação será grande, podendo causar inclusive a perda do ponto ótimo. Se forem muito pequenos o custo computacional para obter a solução será muito grande. A Fig.4.4.1 nos sugere uma forma de tentar minimizar esse problema. Note que na região em que a função apresenta um valor baixo de derivadas (região esquerda da figura), os limites móveis podem ser grandes (pois a função se comporta quase como uma reta horizontal), já na região em que a função apresenta grandes derivadas os limites móveis devem ser pequenos. Além disso, a medida que nos aproximamos da solução ótima (derivada zero) devemos reduzir os limites móveis, caso contrário podemos "passar" pelo ponto ótimo gerando oscilação no valor da função objetivo. Em geral, os valores de limites móveis utilizados estão em torno de 15% do valor de x_i .

Melhores aproximações (quadrática, etc.) também podem ser adotadas considerando-se mais termos na série de Taylor como é o caso do PQS descrito adiante. No entanto, a informação sobre gradientes mais elevados se faz necessária. Os gradientes (ou derivadas) também são chamados de *sensibilidade* do problema de otimização. O seu cálculo preciso define a obtenção ou não da solução ótima. O cálculo da sensibilidade é descrito em detalhe no capítulo 5.

Vejam agora o conceito dos limites móveis, considerando um exemplo com uma função bidimensional:

$$\begin{aligned} \text{Min} \quad & f(\mathbf{x}) = -2x_1 - x_2 \\ & x_1, x_2 \\ \text{tal que} \quad & g_1 = 25 - x_1^2 - x_2^2 \geq 0 \\ & g_2 = 7 - x_1^2 + x_2^2 \geq 0 \\ & x_1, x_2 \geq 0 \end{aligned}$$

O ponto de partida usado é $\mathbf{x}_0^T = (1,1)$ e o valor dos limites móveis usados será igual à 1.

Solução:

Note que a função objetivo já é linear, assim somente é necessário aproximar as restrições. Calculando a série de Taylor das restrições temos:

$$g_1(\mathbf{x}_0) = 25 - 1 - 1 = 23; \quad g_2(\mathbf{x}_0) = 7 - 1 + 1 = 7$$

$$\nabla g_1 = \begin{Bmatrix} -2x_1 \\ -2x_1 \end{Bmatrix}; \quad (\nabla g_1)_{\mathbf{x}_0} = \begin{Bmatrix} -2 \\ -2 \end{Bmatrix}; \quad \text{e} \quad \nabla g_2 = \begin{Bmatrix} -2x_1 \\ 2x_2 \end{Bmatrix}; \quad (\nabla g_2)_{\mathbf{x}_0} = \begin{Bmatrix} -2 \\ 2 \end{Bmatrix}$$

Assim, as séries de Taylor de g_1 e g_2 ficam:

$$g_{1L}(\mathbf{x}) = 23 + \begin{bmatrix} -2 & -2 \end{bmatrix} \begin{Bmatrix} x_1 - 1 \\ x_2 - 1 \end{Bmatrix} = 27 - 2x_1 - 2x_2 \geq 0$$

$$g_{2L}(\mathbf{x}) = 7 + \begin{bmatrix} -2 & 2 \end{bmatrix} \begin{Bmatrix} x_1 - 1 \\ x_2 - 1 \end{Bmatrix} = 7 - 2x_1 + 2x_2 \geq 0$$

Assim o problema de programação linear fica:

$$\begin{aligned} \text{Min} \quad & f(\mathbf{x}) = -2x_1 - x_2 \\ & x_1, x_2 \\ \text{tal que} \quad & g_{1L} = 25 - 2x_1 - 2x_2 \geq 0 \\ & g_{2L} = 7 - 2x_1 + 2x_2 \geq 0 \\ & 0 \leq x_1, x_2 \leq 2 \end{aligned}$$

Cuja solução vale: $\mathbf{x}_1^T = (2,2) \Rightarrow f = -6$. Se não houvessem limites móveis a solução seria: $\mathbf{x}_1^T = (8,5;5,0) \Rightarrow f = -22$. A Fig.4.4.2 ilustra a situação descrita acima.

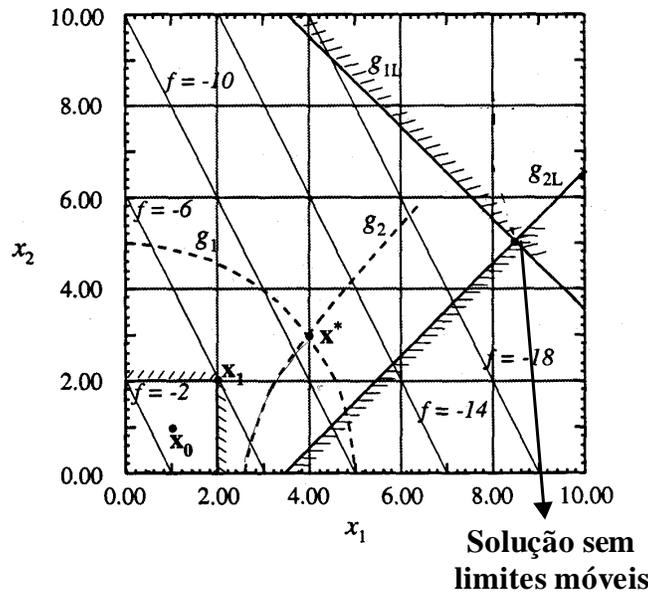


Fig.4.4.2: Linearização das restrições e limites móveis.

Dessa forma, embora o valor da função objetivo seja menor, as restrições exatas (g_1 e g_2) são substancialmente violadas.

Vejamos algumas considerações sobre os limites móveis:

- Limites móveis devem ser pequenos o suficiente de maneira que x_L seja próximo do valor ótimo;
- Os limites móveis devem ser reduzidos a medida que a convergência se aproxima;
- Valores iniciais mais comuns dos limites móveis: $0,1x$ à $0,3x$;
- Redução nos valores de limites móveis de 10 à 50% dos valores anteriores quando $\Delta f \leq \epsilon$;
- $a_L \geq x_{\min}$ e $a_U \leq x_{\max}$, onde a_L e a_U são os limites móveis e x_{\min} e x_{\max} os limites absolutos das variáveis (ver formulação acima);
- Se $x_i \cong 0$, a_{Li} e a_{Ui} não devem ser definidos em função de x_i ;
- A oscilação da solução entre 2 pontos indica que uma estratégia apropriada de redução dos limites móveis deve ser adotada;

Consideremos agora a solução do problema de otimização abaixo usando o método PLS.

$$\begin{aligned} \text{Min}_{x_1, x_2} \quad & f(\mathbf{x}) = \frac{3}{x_1} + \frac{\sqrt{3}}{x_2} \\ \text{tal que} \quad & 18x_1 + 6\sqrt{3}x_2 \leq 3 \\ & 0,05 \leq x_1 \leq 0,1546 \\ & 0,05 \leq x_2 \leq 0,1395 \end{aligned}$$

O ponto de partida é igual à: $\mathbf{x}_0^T = (0,1;0,1) \Rightarrow f = 47,32$ e os limites móveis são iguais a $a_{Ui} = a_{Li} = 0,01$.

Solução:

Somente a função objetivo precisa ser aproximada, já que a restrição é uma função linear. Assim, sendo:

$$\nabla f(\mathbf{x}_0) = \begin{Bmatrix} -300 \\ -173,2 \end{Bmatrix}$$

podemos definir o subproblema de programação linear:

$$\begin{array}{l} \text{Min} \quad f_L(\mathbf{x}) = 47,32 - 300(x_1 - 0,1) + \\ x_1, x_2 \quad -173,2(x_2 - 0,1) \\ \text{tal que} \quad 18x_1 + 6\sqrt{3}x_2 \leq 3 \\ \quad \quad \quad 0,09 \leq x_1 \leq 0,11 \\ \quad \quad \quad 0,09 \leq x_2 \leq 0,11 \end{array}$$

Cuja solução é igual à: $\mathbf{x}_1^T = (0,10316; 0,11) \Rightarrow f_L = 44,641 \Rightarrow f = 44,8274$.

Linearizando a função em torno de x_1 e formulando o novo subproblema de programação linear (PL):

$$\begin{array}{l} \text{Min} \quad f_L(\mathbf{x}) = 44,8274 - 281,9(x_1 - 0,10316) + \\ x_1, x_2 \quad -143,1(x_2 - 0,11) \\ \text{tal que} \quad 18x_1 + 6\sqrt{3}x_2 \leq 3 \\ \quad \quad \quad 0,09316 \leq x_1 \leq 0,11316 \\ \quad \quad \quad 0,1 \leq x_2 \leq 0,12 \end{array}$$

Cuja solução vale: $\mathbf{x}_2^T = (0,10893; 0,1) \Rightarrow f_L = 44,6313 \Rightarrow f = 44,8607$.

Assim, embora o valor da função aproximada diminuiu, o valor da função exata aumentou dessa forma os valores dos limites móveis devem ser reduzidos. Adotando agora $a_{Ui} = a_{Li} = 0.005$ e $\mathbf{x}_0^T = (0,10316; 0,11)$, obtemos após a solução do novo subproblema linear:

$$\mathbf{x}_1^T = (0,10604; 0,105) \Rightarrow f_L = 44,7294 \Rightarrow f = 44,785$$

Na segunda iteração obtemos: $\mathbf{x}_2^T = (0,10316; 0,11) = \mathbf{x}_1^T$ que consiste numa oscilação. Isso indica que devemos reduzir novamente os limites móveis para continuar a melhorar f . No entanto nas duas últimas iterações f variou de 44,8274 para 44,7856 o que é menos do que 0,1%, assim é razoável parar.

No entanto, em cada um dos subproblemas de PL a restrição linear estava ativa, assim podemos obter a solução ótima de forma analítica :

$$18x_1 + 6\sqrt{3}x_2 = 3 \Rightarrow x_2 = \frac{3 - 18x_1}{6\sqrt{3}} \Rightarrow f = \frac{3}{x_1} + \frac{6}{3 - 18x_1} \Rightarrow x_1 = x_2 = 0,10566 \Rightarrow f = 44,7846$$

A Fig.4.4.3 ilustra a situação do problema acima.

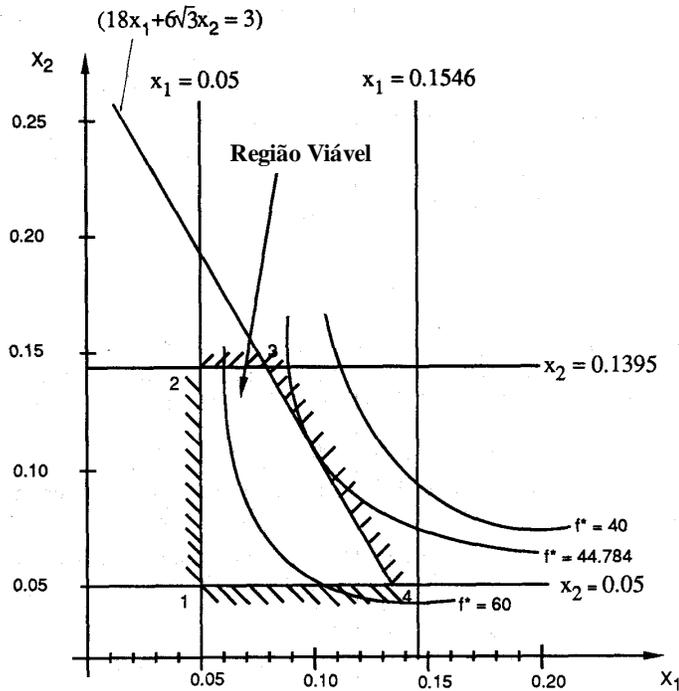


Fig.4.4.3: Domínio viável do problema e curvas de nível de f.

O problema anterior levanta uma questão importante. Como decidir em cada iteração se o resultado obtido no PL é uma melhora ou não? Podemos ter duas situações:

- Há uma melhora de f mas uma violação da restrição;
- Há uma melhora na satisfação das restrições mas uma piora em f;

Se f e as restrições melhoram, então os resultados devem ser aceitos. Se f melhora, mas as restrições deterioram ou vice-versa, demonstra-se a regra descrita a seguir. Considere, a função Lagrangeana (L) definida como:

$$L = f(\mathbf{x}_{iL}^*) - \sum_{j=1}^r \lambda_j g_j(\mathbf{x}_{iL}^*)$$

onde λ_j e \mathbf{x}_{iL}^* é foram obtidos no PL. Se L do resultado corrente for menor do que L do resultado anterior, então aceite o resultado, caso contrário altere os limites móveis.

Vejamos um exemplo para ilustrar esse conceito. Considere a solução do problema abaixo:

$$\begin{aligned} \text{Min} \quad & f(\mathbf{x}) = 3x_1 + \sqrt{3}x_2 \\ & x_1, x_2 \\ \text{tal que} \quad & g(\mathbf{x}) = 3 - \frac{18}{x_1} - \frac{6\sqrt{3}}{x_2} \geq 0 \\ & 8 \leq x_1 \leq 20 \\ & 8 \leq x_2 \leq 20 \end{aligned}$$

O ponto de partida vale $\mathbf{x}_0^T = (12;8) \Rightarrow f = 49,856$ e os limites móveis valem $a_{U_i} = a_{L_i} = 0,3x_i$.

Solução:

Como nenhuma das restrições está ativa, o Lagrangeano inicial vale: $L_0 = f = 49,856$. Linearizando a restrição e formulando o subproblema de PL temos:

$$\begin{aligned} & \text{Min}_{x_1, x_2} \quad 3x_1 + \sqrt{3}x_2 \\ & \text{tal que} \quad 0,125x_1 + 0,1624x_2 \geq 2,598 \\ & \quad \quad \quad 8,4 \leq x_1 \leq 15,6 \\ & \quad \quad \quad 8,0 \leq x_2 \leq 10,4 \end{aligned}$$

Cuja solução é igual à:

$$\mathbf{x}_1^T = (8,4; 9,534) \Rightarrow f = 41,713 \Rightarrow g = -0,2329; \lambda_1 = 10,667; \lambda_2 = 1,667$$

e o Lagrangeano vale: $L_1 = 41,713 - 10,667(-0,2329) = 44,197 < L_0$. Portanto, aceitamos a solução. Note que embora a restrição g_2 esteja ativa no subproblema linear, está inativa no problema não-linear.

Na próxima iteração o novo subproblema de PL é:

$$\begin{aligned} & \text{Min}_{x_1, x_2} \quad 3x_1 + \sqrt{3}x_2 \\ & \text{tal que} \quad 0,2551x_1 + 0,1143x_2 \geq 3,4658 \\ & \quad \quad \quad 8,0 \leq x_1 \leq 10,92 \\ & \quad \quad \quad 8,0 \leq x_2 \leq 12,3938 \end{aligned}$$

Cuja solução é igual à:

$$\mathbf{x}_2^T = (10; 8) \Rightarrow f = 43,858 \Rightarrow g = -0,09896; \lambda_1 = 11,76; \lambda_2 = 0,38$$

e o Lagrangeano vale: $L_2 = 43,858 - 11,76(-0,09896) = 45,022 > L_1$. Portanto, devemos rejeitar a solução e resolver novamente o PL, alterando os limites móveis.

Implementação do PLS

A Fig.4.4.4 mostra o fluxograma da implementação do PLS utilizando-se um software de elementos finitos comercial. Os passos seguidos são similares aos seguidos nos exemplos acima, apenas que o valor da função objetivo resulta de uma análise de MEF.

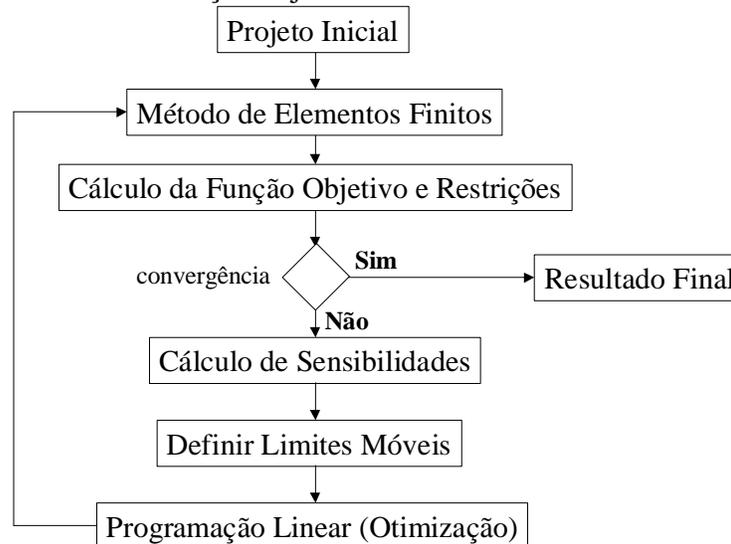


Fig.4.4.4: Fluxograma da implementação do PLS.

Entre as vantagens do PLS temos:

- Adequado para um grande nº de variáveis de projeto;
- São necessárias apenas as primeiras derivadas;
- Fácil implementação;
- Rotinas de programação linear estão largamente disponibilizadas gratuitamente na internet;

Entre as desvantagens:

- Convergência depende da escolha apropriada dos limites móveis;
- Ponto de partida \mathbf{x}_0 deve estar dentro do domínio viável;
- Precisão deve aumentar a medida que se aproxima do valor ótimo (redução dos limites móveis);

No caso em que o ponto de partida não esteja no domínio viável, deve-se relaxar as restrições nos primeiros ciclos através da seguinte formulação do subproblema de PL:

$$\begin{array}{l}
 \text{Min}_{\mathbf{x}, \beta} \quad f(\mathbf{x}_0) + \sum_{i=1}^n (x_i - x_{0i}) \left(\frac{\partial f}{\partial x_i} \right)_{\mathbf{x}_0} + k\beta \\
 \text{tal que} \\
 g_j(\mathbf{x}_0) + \sum_{i=1}^n (x_i - x_{0i}) \left(\frac{\partial g_j}{\partial x_i} \right)_{\mathbf{x}_0} + \beta \geq 0 \\
 \qquad \qquad \qquad j = 1, \dots, n_g \\
 a_{li} \leq x_i - x_{0i} \leq a_{ui} \\
 \beta \geq 0
 \end{array}$$

onde β é a margem de violação da restrição e k deve ser alto o suficiente para priorizar a minimização de β .

4.5 Método de Programação Quadrática Sequencial (PQS ou "SQP")

A idéia básica é similar a do PLS, no entanto o problema não-linear é aproximado por uma sucessão de subproblemas de programação quadrática (ver seção 4.3.2.6) com restrições lineares. Existem vários métodos de SQP e o apresentado aqui é denominado método do Lagrangeano projetado.

Consideremos novamente o problema não-linear de otimização:

$$\begin{array}{l}
 \text{Min}_{\mathbf{x}} \quad f(\mathbf{x}) \\
 \text{tal que} \quad g_j(\mathbf{x}) \geq 0, \quad j = 1, \dots, n_g
 \end{array}$$

A cada iteração a direção de movimento é obtida resolvendo-se o seguinte subproblema de programação quadrática (PQ):

$$\begin{aligned} \text{Min}_{\mathbf{s}} \quad & \phi(\mathbf{s}) = f(\mathbf{x}^{(i)}) + \mathbf{s}^T \nabla f(\mathbf{x}^{(i)}) + (1/2)\mathbf{s}^T \mathbf{A}(\mathbf{x}^{(i)}, \lambda^{(i)})\mathbf{s} \\ \text{tal que} \quad & g_j(\mathbf{x}^{(i)}) + \mathbf{s}^T \nabla g_j(\mathbf{x}^{(i)}) \geq 0, \quad j = 1, \dots, n_g \end{aligned}$$

onde \mathbf{A} é uma aproximação positiva-definida da matriz Hessiana da função Lagrangeana. Como já visto na seção 4.3.2.6, a solução desse problema fornece \mathbf{s} e λ_{i+1} . Como $\mathbf{x}^{(i+1)} = \mathbf{x}^{(i)} + \alpha\mathbf{s}$, α é obtido minimizando-se a função:

$$\psi(\alpha) = f(\mathbf{x}) + \sum_{j=1}^{n_g} \mu_j |\min(0, g_j(\mathbf{x}))| \quad \text{onde: } \mu_j^{(i)} = \max \left[\left| \lambda_j^{(i)} \right|, \frac{1}{2} \left(\mu_j^{(i-1)} + \left| \lambda_j^{(i-1)} \right| \right) \right]$$

onde o sobrescrito (i) denota o número da iteração. Para a primeira iteração: $\mu_j^{(1)} = |\lambda_1|$.

A matriz \mathbf{A} é inicializada com a matriz identidade (ou alguma matriz positiva-definida) e atualizada usando uma equação do tipo BFGS:

$$\mathbf{A}_{\text{novo}} = \mathbf{A} - \frac{\mathbf{A}\Delta\mathbf{x}\Delta\mathbf{x}^T\mathbf{A}}{\Delta\mathbf{x}^T\mathbf{A}\Delta\mathbf{x}} + \frac{\Delta l \Delta l^T}{\Delta\mathbf{x}^T\Delta\mathbf{x}}$$

onde: $\Delta\mathbf{x} = \mathbf{x}_{i+1} - \mathbf{x}_i$ e $\Delta l = \nabla_x L(\mathbf{x}_{i+1}, \lambda_i) - \nabla_x L(\mathbf{x}_i, \lambda_i)$

onde L é a função Lagrangeana e ∇_x denota o gradiente de L em relação a \mathbf{x} . Para garantir que \mathbf{A} seja positiva-definida Δl é modificado se $\Delta\mathbf{x}^T\Delta l \leq 0,2\Delta\mathbf{x}^T\mathbf{A}\Delta\mathbf{x}$ e substituído por:

$$\Delta l' = \theta\Delta l + (1-\theta)\mathbf{A}\Delta\mathbf{x} \quad \text{onde: } \theta = \frac{0,8\Delta\mathbf{x}^T\mathbf{A}\Delta\mathbf{x}}{\Delta\mathbf{x}^T\mathbf{A}\Delta\mathbf{x} - \Delta\mathbf{x}^T\Delta l}$$

No caso do PLS em cada iteração se resolve um subproblema de PL onde se caminha pelos extremos do domínio viável local (ver seção 4.2).

Dessa forma, no SQP o problema não-linear é aproximado por uma sucessão de subproblemas de PQ definido acima.

Como exemplo, considere a solução do problema de otimização abaixo:

$$\begin{aligned} \text{Min}_{x_1, x_2} \quad & f(\mathbf{x}) = 3x_1 + \sqrt{3}x_2 \\ \text{tal que} \quad & g_1 = 3 - \frac{18}{x_1} - \frac{6\sqrt{3}}{x_2} \geq 0 \\ & g_2 = x_1 - 5,73 \geq 0 \\ & g_3 = x_2 - 7,17 \geq 0 \end{aligned}$$

Como ponto de partida é adotado: $x_1 = 11,61$; $x_2 = 7,17$

Solução:

Temos duas restrições ativas no ponto de partida: $g_1 = 0$; $g_3 = 0$; $f = 47,25$

Calculando os gradientes da função objetivo e restrições ativas:

$$\nabla f = \begin{Bmatrix} 3 \\ \sqrt{3} \end{Bmatrix}; \nabla g_1 = \begin{Bmatrix} 0,1335 \\ 0,2021 \end{Bmatrix}; \nabla g_3 = \begin{Bmatrix} 0 \\ 1 \end{Bmatrix}; \nabla g_2 = \begin{Bmatrix} 1 \\ 0 \end{Bmatrix}$$

e assumindo $\mathbf{A} = \mathbf{I} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, resulta no seguinte subproblema de PQ:

$$\begin{aligned} \text{Min } \phi(s) &= 47,25 + 3s_1 + \sqrt{3}s_2 + 0,5s_1^2 + 0,5s_2^2 \\ s_1, s_2 \\ \text{tal que } g_1 &= 0,1335s_1 + 0,2021s_2 \geq 0 \\ g_2 &= 5,88 + s_1 \geq 0 \\ g_3 &= s_2 \geq 0 \end{aligned}$$

cuja solução obtida usando o comando QUAPRO do SCILAB é:

$$\lambda_2 = \lambda_3 = 0; \lambda_1 = 12,8; s_1 = -1,29; s_2 = 0,855 \Rightarrow \mathbf{x}_1 = \begin{Bmatrix} 11,61 \\ 7,17 \end{Bmatrix} + \alpha \begin{Bmatrix} -1,29 \\ 0,855 \end{Bmatrix};$$

onde α é obtido minimizando-se $\psi(\alpha)$. Como na primeira iteração $\mu_1 = |\lambda_1|$, assim:

$$\begin{aligned} \psi &= 3(11,61 - 1,29\alpha) + \sqrt{3}(7,17 + 0,855\alpha) + 12,8 \left| 3 - \frac{18}{11,61 - 1,29\alpha} - \frac{6\sqrt{3}}{7,17 + 0,855\alpha} \right| \Rightarrow \\ \Rightarrow \alpha &= 2,2; \mathbf{x}^{(1)} = (8,77; 9,05)^T; f(\mathbf{x}^{(1)}) = 41,98; g_1(\mathbf{x}^{(1)}) = -0,201 \end{aligned}$$

Atualizando \mathbf{A} :

$$\begin{aligned} L &= 3x_1 + \sqrt{3}x_2 - 12,8(3 - 18/x_1 + 6\sqrt{3}/x_2) \Rightarrow \nabla_x L = (3 - 230,4/x_1^2, \sqrt{3} - 133/x_2^2)^T \Rightarrow \\ \Rightarrow \Delta \mathbf{x} &= \mathbf{x}^{(1)} - \mathbf{x}^{(0)} = \begin{Bmatrix} -2,84 \\ 1,88 \end{Bmatrix}; \end{aligned}$$

$$\Delta l = \nabla_x L(\mathbf{x}^{(1)}) - \nabla_x L(\mathbf{x}^{(0)}) = \begin{Bmatrix} -1,31 \\ 0,963 \end{Bmatrix} \Rightarrow \text{Como } \mathbf{A} = \mathbf{I} \Rightarrow \Delta \mathbf{x}^T \mathbf{A} \Delta \mathbf{x} = 11,6;$$

$$\Delta \mathbf{x}^T \Delta l = 5,53 > 0, 2\Delta \mathbf{x}^T \mathbf{A} \Delta \mathbf{x} \Rightarrow \mathbf{A}_{\text{nov}} = \mathbf{I} - \frac{\Delta \mathbf{x} \Delta \mathbf{x}^T}{\Delta \mathbf{x}^T \Delta \mathbf{x}} + \frac{\Delta l \Delta l^T}{\Delta \mathbf{x}^T \Delta \mathbf{x}} = \begin{bmatrix} 0,453 & 0,352 \\ 0,352 & 0,775 \end{bmatrix}$$

Para a segunda iteração, o subproblema PQ fica:

$$\begin{aligned} \text{Min } \phi(s) &= 41,98 + 3s_1 + \sqrt{3}s_2 + 0,5(0,453s_1^2 + \\ s_1, s_2 &+ 0,775s_2^2 + 0,704s_1s_2) \\ \text{tal que } g_1 &= -0,201 + 0,234s_1 + 0,127s_2 \geq 0 \\ g_2 &= 3,04 + s_1 \geq 0 \\ g_3 &= 1,88 + s_2 \geq 0 \end{aligned}$$

cuja solução obtida usando o comando QUAPRO do SCILAB é:

$$\lambda_2 = \lambda_3 = 0; \lambda_1 = 14,31; s_1 = 1,059; s_2 = -0,376$$

onde α é obtido minimizando-se $\psi(\alpha) = f(\alpha) + \mu_1 g_1(\alpha)$ e:

$$\mu_1 = \max\left(\lambda_1, \frac{1}{2}(\lambda_1 + \mu_1^{\text{velho}})\right) = 14,31 \Rightarrow \alpha = 0,5 \Rightarrow \mathbf{x}^{(2)} = (9,30; 8,86)^T;$$

$$f(\mathbf{x}^{(2)}) = 43,25; \quad g_1(\mathbf{x}^{(2)}) = -0,108$$

portanto foi feito um bom progresso em direção à solução ótima que vale $\mathbf{x}^* = (9,46; 9,46)^T$.