

## **6**

# **Fundamentos de programação matemática**

### **6.1**

#### **Introdução**

A grande maioria dos problemas de engenharia pode ser solucionado de diferentes formas, uma vez que um número muito grande de soluções atende aos critérios e normas previamente estabelecidas. Desta forma seria exigido um número muito grande de repetições nos cálculos para a obtenção dos valores ideais. Porém, com a descrição, através de funções matemáticas, do comportamento da estrutura pode-se determinar os valores extremos dessas funções com a utilização de técnicas de programação matemática.

A programação matemática trata da obtenção de valores extremos, máximos ou mínimos, para funções em problemas sujeitos ou não a restrições. Este capítulo apresenta de forma sucinta os conceitos básicos de programação matemática, como a descrição de um problema de programação matemática e condições de ótimo. Descreve-se também o algoritmo de pontos interiores empregado no trabalho na etapa de otimização.

### **6.2**

#### **Conceitos básicos**

Um problema de otimização, no caso minimização, pode ser escrito de forma geral da seguinte maneira:

$$\begin{array}{lll}
\text{Minimizar} & f(x) & x \in \mathfrak{R}^n \\
\text{sujeito a} & c_i(x) = 0 & i = 1 \dots l \\
& c_i(x) \leq 0 & i = l+1 \dots m \\
& x_i^l \leq x_i \leq x_i^u & i = 1 \dots n
\end{array} \tag{6.1}$$

Onde,  $f(x)$  é a função a ser minimizada designada de função objetivo,  $x$  é um ponto do  $\mathfrak{R}^n$ ,  $c_i(x)$  são chamadas restrições do problema e representam as restrições de igualdade e de desigualdade. Supõe-se que tanto a função objetivo quanto as restrições são funções contínuas no  $\mathfrak{R}^n$ . Em geral, elas são funções não lineares e implícitas das variáveis  $x$  que definem o problema, entretanto, podem ser funções lineares e as variáveis de projeto  $x$  serem explícitas. Os valores limites que podem ser assumidos pelas variáveis de projeto são denotados pelos sobrescritos  $l$  e por  $u$  sendo estes, respectivamente, os limites inferior e superior.

Um ponto que satisfaça todas as restrições é denominado ponto viável e o conjunto de todos os pontos que satisfaçam todas as restrições é conhecido como região viável ou região factível. Uma restrição de desigualdade define uma fronteira que divide o  $\mathfrak{R}^n$  em uma região factível e outra não factível. Quando um ponto está sobre esta fronteira, a restrição é dita ativa, quando um ponto está no interior da região viável, a restrição é dita inativa e, quando um ponto está fora desta região, a restrição está violada.

### 6.3 Condições de ótimo

As condições de ótimo do problema de programação matemática enunciado em (6.1) são chamadas de condições de Kuhn-Tucker e são descritas por (6.2). A solução  $x^*$  do problema enunciado em (6.1) deve obrigatoriamente atender às condições de Kuhn-Tucker.

$$\begin{aligned}
\nabla_x L(x^*, \lambda^*) &= 0 \\
c_i(x^*) &= 0 & i = 1 \dots l \\
c_i(x^*) &\leq 0 & i = l+1 \dots m \\
\lambda_i^* &\geq 0 & i = l+1 \dots m \\
\lambda_i^* c_i(x^*) &= 0 & \forall i
\end{aligned} \tag{6.2}$$

onde  $L(x^*, \lambda^*)$  é a função lagrangiana dada pela seguinte expressão:

$$L(x^*, \lambda^*) = f(x^*) + \sum_{i=1}^m \lambda_i^* c_i(x^*) \tag{6.3}$$

$\lambda_i^*$  são os multiplicadores de Lagrange associados às restrições no ponto  $x^*$  da solução e  $\nabla_x$  é o operador gradiente em relação às variáveis  $x$ .

As condições de Kuhn-Tucker também são chamadas de condições de primeira ordem. Para problemas de programação convexa, tais como programação linear e programação quadrática as condições de Kuhn-Tucker são suficientes para a determinação de uma solução ótima global.

Entretanto, se o problema não é de programação convexa, as condições de primeira ordem não são suficientes para a determinação da solução ótima global. Deve-se então ser verificada a condição de segunda ordem. A expressão (6.4) representa a condição de segunda ordem de Kuhn-Tucker.

$$\mathbf{d}' \mathbf{W}^* \mathbf{d} \geq 0, \quad \forall \mathbf{d} \neq 0 \tag{6.4}$$

Onde:

$$\mathbf{W}_{ij}^* = \frac{\partial^2 L(x^*, \lambda^*)}{\partial x_i \partial x_j} \tag{6.5}$$

Neste caso,  $\mathbf{W}^*$  a matriz Hessiana da função Lagrangiana em  $x^*$  é positiva definida no ponto ótimo para qualquer direção viável  $\mathbf{d}$ .

## 6.4 Forma geral dos algoritmos de programação matemática

Os algoritmos de programação matemática são baseados em procedimentos iterativos que necessitam de valores iniciais para as variáveis de projeto. Após a especificação desses valores, os valores das variáveis são atualizados pela seguinte expressão:

$$\mathbf{x} = \mathbf{x}_0 + t\mathbf{d} \quad (6.6)$$

Na expressão (6.6)  $\mathbf{d}$  representa a direção de busca, que deve ser também uma direção de descida nos problemas de minimização. O parâmetro escalar  $t$  representa o tamanho do passo a ser dado ao longo da direção de busca. Observa-se que um problema de  $n$  variáveis, passa a ser tratado como um problema unidimensional, função do escalar  $t$ , após a determinação da direção de busca. O ponto obtido definirá um novo vetor  $\mathbf{x}$ . Todo o procedimento é repetido a partir do novo ponto, até que um critério de convergência, pré-estabelecido seja atendido. Dessa maneira, os algoritmos podem ser divididos em duas etapas principais: a primeira etapa é a determinação da direção de busca  $\mathbf{d}$  e a segunda é a avaliação do parâmetro escalar  $t$ . As técnicas de programação matemática variam predominantemente pela forma como são determinados  $\mathbf{d}$  e  $t$ .

De acordo com as informações necessárias para a obtenção da direção de busca podemos classificar os algoritmos de programação matemática como sendo de primeira ou de segunda ordem. Os algoritmos de primeira ordem são os que utilizam apenas os gradientes da função objetivo e das restrições. Os de segunda ordem são os que fazem uso das Hessianas dessas funções.

## 6.5 Algoritmo de pontos interiores

O algoritmo de pontos interiores, desenvolvido por HERSKOVITZ (1995), baseia-se na aplicação do método de Newton para a solução de um sistema de equações não lineares obtidas a partir da aplicação das condições de Kuhn-Tucker no problema de otimização. Dado o fato desse trabalho tratar somente de

restrições de desigualdade, as restrições de igualdade serão desconsideradas e, portanto, não estarão presentes na descrição do algoritmo. Salienta-se, porém, que o mesmo tratamento dispensado às restrições de desigualdade pode ser dado às restrições de igualdade.

O algoritmo de pontos interiores apresenta duas características bastante importantes que favorecem sua utilização. A primeira propriedade do algoritmo é que este gera pontos ou soluções intermediárias no interior da região viável. A segunda propriedade, que decorre da primeira, é o fato de que se, por alguma razão, o algoritmo for interrompido antes de alcançar o ponto de ótimo, tem-se um ponto viável e com um valor da função objetivo inferior ao anterior. Isso se deve ao fato de cada ponto intermediário gerar valores decrescentes para a função objetivo.

Com essas considerações pode-se descrever o problema de otimização da seguinte forma:

$$\begin{array}{ll} \text{minimizar} & f(\mathbf{x}) \\ \text{sujeito a} & c_i(\mathbf{x}) \leq 0 \quad i = 1 \dots m \end{array} \quad (6.7)$$

Cujas condições de Kuhn-Tucker são:

$$\begin{aligned} \mathbf{g} + \sum_{i=1}^m \lambda_i \mathbf{a}_i &= 0 \\ \lambda_i^* c_i(x^*) &= 0 \\ c_i(x^*) &\leq 0 \\ \lambda_i^* &\geq 0 \end{aligned} \quad (6.8)$$

onde  $\mathbf{g}$  é o vetor gradiente da função objetivo,  $\mathbf{a}_i$  os gradientes das restrições avaliadas no ponto corrente  $x$ .

Sendo  $\mathbf{A}$  a matriz dos gradientes das restrições e  $\mathbf{C}$  uma matriz diagonal contendo os valores das restrições, descrevem-se as duas primeiras equações de (6.8) como:

$$\begin{aligned} \mathbf{g} + \mathbf{A}'\lambda &= 0 \\ \mathbf{C}\lambda &= 0 \end{aligned} \quad (6.9)$$

Aplicando o método de Newton para resolver o problema acima, obtém-se o sistema de equações lineares:

$$\begin{bmatrix} \mathbf{W} & \mathbf{A}' \\ \Lambda \mathbf{A} & \mathbf{C} \end{bmatrix} \begin{Bmatrix} \mathbf{d}_0 \\ \lambda_0 \end{Bmatrix} = - \begin{Bmatrix} \mathbf{g} \\ 0 \end{Bmatrix} \quad (6.10)$$

Na equação (6.10),  $\Lambda$  é uma matriz diagonal para a qual  $\Lambda_{ii} = \lambda_i$ ,  $\mathbf{d}_0$  é a direção de busca,  $\lambda_0$  é a estimativa dos multiplicadores de Lagrange e  $\mathbf{W}$  é a hessiana da função Lagrangiana.  $\mathbf{d}_0$  e  $\lambda_0$  são solução do sistema.

A direção de busca encontrada com a solução de (6.10) pode apontar para uma direção não viável. Fazendo-se a expansão de uma das equações da parte inferior do sistema (6.10), chega-se a:

$$\lambda_i \mathbf{a}'_i \mathbf{d}_0 + c_i \lambda_0 = 0 \quad (6.11)$$

A equação (6.11) implica que  $\mathbf{a}'_i \mathbf{d}_0 = 0$  para todo  $i$  tal que  $c_i = 0$ . Significando geometricamente, que  $\mathbf{d}_0$  é tangente às restrições ativas, podendo apontar para fora da região viável no caso de restrição convexa.

Uma forma de se evitar este efeito é adicionar uma constante negativa ao lado direito da equação acima:

$$\lambda_i \mathbf{a}'_i \mathbf{d} + c_i \bar{\lambda}_i = -\rho \lambda_i \quad (6.12)$$

Este procedimento faz com que a direção original seja defletida, de um valor proporcional a  $\rho$ , para o interior da região viável. Sendo a deflexão proporcional a  $\rho$  e sendo  $\mathbf{d}_0$  uma direção de decréscimo de  $f$ , é possível encontrar limites em  $\rho$  para que  $\mathbf{d}$  também seja uma direção de decréscimo. Isto é possível através de:

$$\mathbf{g}' \mathbf{d} \leq k_a \mathbf{g}' \mathbf{d}_0 \quad (6.13)$$

Para  $k_a \in (0; 1)$ . Em geral, a taxa de decréscimo de  $f$  ao longo de  $\mathbf{d}$  é menor que ao longo de  $\mathbf{d}_0$ .

Considerando agora um sistema auxiliar expresso por:

$$\begin{bmatrix} \mathbf{W} & \mathbf{A}' \\ \Lambda \mathbf{A} & \mathbf{C} \end{bmatrix} \begin{Bmatrix} \mathbf{d}_1 \\ \lambda_1 \end{Bmatrix} = - \begin{Bmatrix} \mathbf{g} \\ \lambda \end{Bmatrix} \quad (6.14)$$

É possível demonstrar que:

$$\mathbf{d} = \mathbf{d}_0 + \rho \mathbf{d}_1 \quad (6.15)$$

e:

$$\bar{\lambda} = \lambda_0 + \rho \lambda_1 \quad (6.16)$$

Substituindo-se (6.15) em (6.13) pode-se escrever:

$$\rho \leq (k_a - 1) \frac{\mathbf{g}' \mathbf{d}_0}{\mathbf{g}' \mathbf{d}_1} \quad (6.17)$$

Após a definição da direção de busca faz-se uma busca linear ao longo dessa direção, de forma que seja garantido que o ponto encontrado esteja no interior da região viável.

## 6.6 Etapas do Algoritmo de Pontos Interiores

Como já mencionado o algoritmo de pontos interiores necessita de um ponto inicial. Como o algoritmo é de pontos interiores, ou seja os pontos são viáveis, o ponto inicial  $\mathbf{x}_0$  também deve ser viável. Também são necessárias estimativas iniciais dos multiplicadores de Lagrange tais que  $\lambda_i > 0$ , e da matriz  $\mathbf{B}$ , simétrica e positiva definida, para se obter a aproximação inicial de  $\mathbf{W}$ . De acordo com HERSKOVITS & SANTOS (1997), após a determinação dos valores iniciais dos vetores  $(\mathbf{d}_0, \lambda_0)$  o algoritmo segue os seguintes passos:

### 1. Cálculo da direção de busca $\mathbf{d}$ :

Faz-se a verificação do critério de convergência, utilizando a expressão (6.18).

$$\|\mathbf{d}\| \leq tol \quad (6.18)$$

Caso o critério de convergência seja atendido, parar o algoritmo. Caso contrário, determinam-se os vetores  $(\mathbf{d}_1, \lambda_1)$  através da solução do sistema linear descrito por (6.14). Calcula-se o valor de  $\rho$  por (6.19)

$$\begin{cases} \text{se } \mathbf{g}^t \mathbf{d}_1 > 0, \text{ então } \rho = \min \left[ k_f \|\mathbf{d}_0\|^2, (k_a - 1) \mathbf{g}^t \mathbf{d}_0 / \mathbf{g}^t \mathbf{d}_1 \right] \\ \text{se } \mathbf{g}^t \mathbf{d}_1 \leq 0, \text{ então } \rho = k_f \|\mathbf{d}_0\|^2 \end{cases} \quad (6.19)$$

Sendo  $k_f$  uma constante escalar pré definida e estritamente positiva.

Calcula-se a direção de busca  $\mathbf{d}$ :

$$\mathbf{d} = \mathbf{d}_0 + \rho \mathbf{d}_1 \quad (6.20)$$

E

$$\bar{\lambda} = \lambda_0 + \rho \lambda_1 \quad (6.21)$$

## 2. Busca linear sobre $\mathbf{d}$ :

Determina-se o tamanho do passo  $t$ , tal que

$$\begin{cases} c_i(\mathbf{x} + t\mathbf{d}) \leq 0, & \text{se } \bar{\lambda}_i \geq 0 \\ c_i(\mathbf{x} + t\mathbf{d}) \leq c_i(x), & \text{se } \bar{\lambda}_i < 0 \end{cases} \quad (6.22)$$

e o novo ponto  $\mathbf{x}$

$$\mathbf{x} = \mathbf{x}_0 + t\mathbf{d} \quad (6.23)$$

## 3. Atualização da matriz $\mathbf{B}$ :

É realizada a atualização da matriz  $\mathbf{B}$  (aproximação da matriz Hessiana da função Lagrangeana  $\mathbf{W}$ ), através do método BFGS (Broyden - Fletcher - Goldfarb - Shanno).

A aproximação  $\mathbf{B}_0$  da matriz Hessiana  $\mathbf{W}_0$  da função Lagrangeana  $\mathbf{B}_0$  é dada pela seguinte função:

$$\mathbf{B}_0 = b_o \mathbf{I} \quad (6.24)$$

Onde  $b_o$  é um parâmetro definido pelo usuário do algoritmo. Para maior estabilidade do processo numérico o esquema BFGS de aproximação da matriz Hessiana deve ser reinicializado após um número  $n_r$  de iterações,

definido pelo usuário. Segundo PARENTE (2000) o reinício de  $\mathbf{B}$  serve para descartar a influência de pontos muito distantes do ponto corrente.

4. Nova estimativa para os multiplicadores de Lagrange através da expressão (6.25):

$$\lambda_i = \max \left[ \lambda_0, k_e \|\mathbf{d}_0\|^2 \right] \quad (6.25)$$

Sendo  $k_e$  uma constante escalar pré definida e estritamente positiva.

5. Fazer  $\mathbf{x}$  igual a  $\mathbf{x}_0$  e retornar ao passo 1.