

5

Programação Matemática – Princípios Básicos

5.1

Considerações Gerais

Este capítulo tem por objetivo apresentar os conceitos básicos de Programação Matemática (PM), necessários à compreensão do processo de otimização de dimensões, bem como descrever os algoritmos de otimização utilizados.

Em problemas típicos de engenharia, é possível identificar uma infinidade de soluções capazes de atender à todas as especificações de projeto. Em um problema de otimização deseja-se obter um *projeto ótimo*, maximizando ou minimizando uma função, adotada como medida quantitativa do desempenho do sistema analisado, a qual denomina-se *função objetivo*.

A maioria dos problemas de otimização são formulados em termos de parâmetros e restrições. Os parâmetros escolhidos para descrever o projeto de uma estrutura são conhecidos como *variáveis de projeto* enquanto que as restrições, impostas para que projeto seja *admissível*, são conhecidas como *condições de restrição*.

A Programação Matemática é a disciplina que estuda a minimização ou maximização de funções em problemas com ou sem restrições. Matematicamente, um problema de otimização, no caso minimização, podem ser enunciados como:

$$\begin{array}{lll} \text{Minimizar} & f(x) & x \in \mathcal{R}^n \\ \text{sujeito a} & c_i(x) = 0 & i = 1 \dots l \\ & c_i(x) \leq 0 & i = l + 1 \dots m \\ & x_i^l \leq x_i \leq x_i^u & i = 1 \dots n \end{array} \quad (5.1)$$

onde x é o vetor de variáveis, sobre o qual são impostos os limites mínimos e máximos (restrições laterais), $f(x)$ é a função a ser minimizada e as funções $c_i(x)$ representam as restrições de igualdade e desigualdade impostas ao sistema. Assume-se que tanto a função objetivo quanto as restrições são funções contínuas

no $\mathfrak{R}^n$. Em geral, $f(x)$ e $c_i(x)$ são funções não-lineares e implícitas das variáveis (x) que definem o problema.

- **Espaço de Busca** $\rightarrow$ É o conjunto, espaço ou região que compreende as possíveis ou viáveis soluções do problema a ser otimizado, sendo caracterizado pelas funções de restrição.

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

5.2 Condições de Ótimo

Para que a solução x^* seja um mínimo local do problema de otimização enunciado em (5.1) é necessário que esta atenda as condições de 1ª ordem, também chamadas de condições de Kuhn-Tucker, enunciadas por:

$$\begin{aligned} \nabla_x L(x^*, \lambda^*) &= 0 \\ c_i(x^*) &= 0 & i = 1 \dots l \\ c_i(x^*) &\leq 0 & i = l+1 \dots m \\ \lambda_i^* &\geq 0 & i = l+1 \dots m \\ \lambda_i^* c_i(x^*) &= 0 & \forall i \end{aligned} \quad (5.2)$$

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

$$L(x^*, \lambda^*) = f(x^*) + \sum_{i=1}^l \lambda_i^* c_i(x^*) \quad (5.3)$$

onde λ_i^* são os multiplicadores de Lagrange associados às restrições no ponto x^* solução.

Para determinadas classes de problemas de programação matemática as condições de Kuhn-Tucker são suficientes para a determinação de uma solução ótima local. Inclui-se nessas classes os problemas de programação convexa, tais como os de programação linear e quadrática. O problema de programação convexa é caracterizado por função objetivo e restrições convexas.

Porém, nos casos mais comuns, onde os problemas não são de programação convexa, as condições de primeira ordem não são suficientes para a determinação da solução ótima local. Assim, além das condições expressas em (5.2), deve ser verificada a condição de segunda ordem expressa na equação (5.4) a seguir:

$$\mathbf{d}^t \mathbf{W}^* \mathbf{d} \geq 0, \quad \forall \mathbf{d} \neq 0 \text{ tal que } \mathbf{d}^t \mathbf{a}_i^* = 0 \quad (5.4)$$

onde $\mathbf{a}_i^* = \nabla c_i(x^*)$ para todas as restrições ativas e $\mathbf{W}^* = \nabla^2 L(x^*)$ é a Hessiana da função Lagrangiana. O que significa que $\mathbf{W}^*$ em x^* é positiva definida no ponto ótimo para qualquer direção estacionária $\mathbf{d}$.

5.3

Forma Geral dos Algoritmos de Otimização

Os problemas de programação matemática são definidos de acordo com as características da função-objetivo e das restrições. Desta forma, os problemas de otimização podem se dividir em diferentes formas, como mostra a Tabela 2.1.

TABELA 5.1
Divisão dos problemas de Programação Matemática

Tipos de Otimização	$f(x)$	$c_i(x)$
Programação Linear	linear	linear
Programação Quadrática	quadrática	linear
Programação Não-Linear	não-linear	não-linear
	não-linear	linear
	linear	não-linear

Algoritmos de otimização para problema de programação linear e programação quadrática têm solução em um número finito de passos. Já os algoritmos de programação não-linear podem não ter solução em um número finito de passos, mas espera-se que a sequência gerada convirja para um mínimo local. Portanto, um problema adicional no processo de otimização ocorre quando a função objetivo e as restrições são funções não-lineares do vetor de variáveis de projeto, $x \in \mathcal{R}^n$.

Os algoritmos de programação não-linear, restrita e irrestrita, são procedimentos iterativos. Nestes, a partir de uma solução inicial x_0 , denominada solução básica, e com base no cálculo de derivadas para indicar a direção na qual se deve caminhar, novas soluções são geradas através da expressão (5.5).

$$x = x_0 + t\mathbf{d} \quad (5.5)$$

Assim, 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 , que representa o tamanho do passo a ser dado ao longo da direção de busca. A partir da expressão (5.5) diversos algoritmos podem ser construídos utilizando diferentes técnicas para a determinação da direção de busca e do tamanho do passo.

Por trabalharem com o cálculo de derivadas, os algoritmos de PM são denominados de algoritmos de ordem n , onde n é a maior derivada utilizada. Um algoritmo é dito de primeira ordem se utilizar apenas os gradientes da função objetivo e das restrições para calcular a direção de busca. Por outro lado, se o algoritmo utiliza informações sobre as Hessianas destas funções, ele é dito de segunda ordem.

5.4

Método de Newton para Problemas de Otimização sem Restrição

O princípio do método de Newton é minimizar a função $f(x)$ através de uma aproximação local por uma função quadrática. Para esse propósito a função $f(x)$ é expandida por série de Taylor até o termo de segunda ordem, ou seja:

$$f(x) = f(x_0) + \nabla f(x_0)(x - x_0) + \frac{1}{2}(x - x_0)^t \nabla^2 f(x_0)(x - x_0) \quad (5.6)$$

se

$$\mathbf{d} = \Delta x = (x - x_0) \rightarrow x = \mathbf{d} + x_0 \quad (5.7)$$

e

$$\mathbf{g} = \nabla f(x_0) \quad \text{e} \quad \mathbf{H} = \nabla^2 f(x_0) \quad (5.8)$$

Substituindo-se (5.7) e (5.8) em (5.6) tem-se

$$f(\mathbf{d} + x_0) = f(x_0) + \mathbf{d}^t \mathbf{g} + \frac{1}{2} \mathbf{d}^t \mathbf{H} \mathbf{d} \quad (5.9)$$

onde $\mathbf{d}$ é o incremento de x_0 , $\mathbf{g}$ é vetor gradiente de f e $\mathbf{H}$, uma matriz simétrica positiva definida, é a hessiana da função f no ponto x_0 . A equação (5.9) é uma equação quadrática cuja variável é $\mathbf{d}$. Portanto, o algoritmo de otimização procura determinar uma direção de busca $\mathbf{d}$ tal que $f(\mathbf{d} + x_0) < f(x_0)$ em cada passo, ou seja, uma direção de decréscimo em f . Assim:

$$\min f(\mathbf{d} + x_0) = \min(\mathbf{d}^t \mathbf{g} + \frac{1}{2} \mathbf{d}^t \mathbf{H} \mathbf{d}) \quad (5.10)$$

Escrevendo a condição de otimalidade de (5.10) ($\nabla_{\mathbf{d}} f(\mathbf{d} + x_0) = 0$), obtém-se $\mathbf{d}$:

$$\mathbf{d} = -\mathbf{H}^{-1} \mathbf{g} \quad (5.11)$$

Assim, (5.11) fornece um mínimo global único para a função aproximadora de f . Além disso, se f é um função quadrática o método de Newton determina o ponto de mínimo em um único passo. Este método, entretanto, tem como desvantagem o elevado esforço computacional envolvido nos cálculos para a montagem da matriz $\mathbf{H}$, sobretudo em problemas com grande número de variáveis.

Os métodos Quase-Newton surgiram para resolver esse problema sem perder as boas propriedades de convergência do método de Newton. Nesses métodos, uma aproximação da Hessiana (ou de sua inversa) é construída a partir dos valores dos gradientes ao longo das iterações. Esses métodos, dos quais o BFGS (Broyden - Fletcher - Goldfarb - Shanno) é o mais popular, possuem convergência superlinear e são amplamente utilizados em problemas de otimização.

5.5 Busca Linear

Após a determinação da direção de busca $\mathbf{d}$ é necessário calcular o tamanho do passo a ser dado nessa direção, a fim de se obter o novo vetor das variáveis de projeto em (5.5). O tamanho do passo é calculado fazendo-se uma minimização da função unidimensional p definida através da expressão:

$$p(t) = f(x_0 + t\mathbf{d}) \quad (5.12)$$

A partir desta definição, pode-se verificar que:

$$p(0) = f(x_0) \quad (5.13)$$

e

$$p'(0) = \frac{\partial f(x)^t}{\partial x} \frac{dx}{dt} \Big|_{t=0} \quad (5.14)$$

onde p' indica a derivada em relação à t .

A busca linear pode ser exata ou aproximada, dependendo do método utilizado para a minimização. A busca aproximada é uma forma mais moderna, na qual o objetivo é determinar t de forma que f apresente um certo nível de decréscimo, segundo um critério preestabelecido, como:

$$p(t) \leq f(x_0) + t\gamma \mathbf{d}^t \mathbf{g}, \quad \gamma \in (0,1) \quad (5.15)$$

De acordo com esta equação, o parâmetro γ controla o tamanho do passo. Assim, um γ pequeno permite a utilização de passos maiores e a utilização de um γ grande força a utilização de passos pequenos.

Uma forma bastante popular de busca linear é fazer uma aproximação quadrática de p e calcular t como o mínimo desta aproximação, verificando se a equação (5.15) é satisfeita. Uma forma ainda mais simples é o método de Armijo (Herskovits, 1995), no qual t é igual ao primeiro número da seqüência $\{1, \alpha, \alpha^2, \alpha^3, \dots\}$, $\alpha \in (0; 1)$, para o qual $p(t)$ satisfaz a condição (5.15).

5.6 Programação Quadrática

A Programação Quadrática (PQ) tem como objetivo determinar o vetor solução x^* do problema colocado na seguinte forma::

$$\begin{array}{ll} \text{minimizar} & \mathbf{q}^t \mathbf{x} + \frac{1}{2} \mathbf{x}^t \mathbf{Q} \mathbf{x} \\ \text{sujeito a} & \mathbf{a}_i^t \mathbf{x} = b_i \quad i = 1 \dots l \\ & \mathbf{a}_i^t \mathbf{x} \leq b_i \quad i = l + 1 \dots m \end{array} \quad (5.16)$$

onde $\mathbf{a}$ é uma matriz que contém os coeficientes dos gradientes das restrições, $\mathbf{b}$ é o vetor dos termos independentes das restrições.

Sendo $\mathbf{Q}$ uma matriz positiva definida, o problema quadrático é convexo e pode-se garantir a existência de um único mínimo local.

A solução deste problema pode ser obtida em três etapas bem definidas (Eboli, 1989 e Parente, 2000):

1. As l restrições de igualdade são eliminadas do problema diminuindo-se o número das variáveis independentes para $n - l$, obtendo-se um problema de programação quadrática (reduzida), chamado problema padrão de PQ, só com as restrições de desigualdade.
2. O problema quadrático reduzido é transformado em um Problema Linear Complementar (PLC), que pode ser resolvido através de métodos de pivoteamento como o de Lemke.
3. Recupera-se a solução para o espaço original com o cálculo das variáveis eliminadas na primeira etapa, obtendo-se os valores de $\mathbf{x}$ e $\boldsymbol{\lambda}$.

5.7

Algoritmo de Han-Powell - Programação Quadrática Seqüencial

O algoritmo de otimização de Han-Powell proposto por Han em 1976 e 1977 e por Powell em 1978 (Eboli, 1989), foi implementado e aplicado a problemas de Engenharia Estrutural no DEC/PUC-Rio por Eboli (1989), Parente (2000) e Farfán (2000). Este algoritmo utiliza a técnica de Programação Quadrática Seqüencial (PQS) através da resolução de um subproblema quadrático (PQ).

O método de PQS pode ser considerado como o resultado da aplicação do método de Newton à minimização da função Lagrangiana do problema. Este método fornece a cada iteração os vetores $\mathbf{d}$ (correção de $\mathbf{x}$) e $\Delta\boldsymbol{\lambda}$ (correção dos multiplicadores de Lagrange $\boldsymbol{\lambda}$).

Este fato pode ser demonstrado considerando o seguinte problema:

$$\begin{array}{ll} \text{minimizar} & f(\mathbf{x}) \\ \text{sujeito a} & c_i(\mathbf{x}) = 0 \end{array} \quad (5.17)$$

cuja função Lagrangiana é dada por:

$$L(x, \lambda) = f(x) + \sum_i \lambda_i c_i(x) \quad (5.18)$$

Desenvolvendo $\nabla L(x, \lambda)$ em séries de Taylor em torno de (x^k, λ^k) até a primeira ordem tem-se:

$$\nabla L(x^k + \mathbf{d}^{k+1}, \lambda^k + \Delta \lambda^{k+1}) = \nabla L(x^k, \lambda^k) + [\nabla^2 L(x^k, \lambda^k)] \begin{Bmatrix} \mathbf{d}^{k+1} \\ \Delta \lambda^{k+1} \end{Bmatrix} \quad (5.19)$$

Considerando $\mathbf{d}^{k+1} = x^{k+1} - x^k$ e $\Delta \lambda^{k+1} = \lambda^{k+1} - \lambda^k$ e aplicando a condição de estacionariedade a (5.19) no ponto $(x^k + \mathbf{d}^{k+1}, \lambda^k + \Delta \lambda^{k+1})$ resulta em:

$$[\nabla^2 L(x^k, \lambda^k)] \begin{Bmatrix} \mathbf{d}^{k+1} \\ \Delta \lambda^{k+1} \end{Bmatrix} = -\nabla L(x^k, \lambda^k) \quad (5.20)$$

ou expresso na forma matricial como:

$$\begin{bmatrix} \mathbf{W}^k & \mathbf{A}^{k^t} \\ \mathbf{A}^k & 0 \end{bmatrix} \begin{Bmatrix} \mathbf{d}^{k+1} \\ \Delta \lambda^{k+1} \end{Bmatrix} = - \begin{Bmatrix} \mathbf{g}^k + \mathbf{A}^k \lambda^k \\ \mathbf{c}^k \end{Bmatrix} \quad (5.21)$$

Finalmente, substituindo $\lambda^{k+1} = \lambda^k + \Delta \lambda^{k+1}$ tem-se:

$$\begin{bmatrix} \mathbf{W}^k & \mathbf{A}^{k^t} \\ \mathbf{A}^k & 0 \end{bmatrix} \begin{Bmatrix} \mathbf{d}^{k+1} \\ \lambda^{k+1} \end{Bmatrix} = - \begin{Bmatrix} \mathbf{g}^k \\ \mathbf{c}^k \end{Bmatrix} \quad (5.22)$$

onde, $\mathbf{A}^k$ é a matriz dos gradientes das restrições, $\mathbf{W}^k$ é a Hessiana da Lagrangiana, e $\mathbf{g}^k$ é o gradiente de $f(x)$ sendo todos avaliados no ponto x^k . A solução de (5.22) equivale à solução do subproblema de PQ (Eboli, 1989):

$$\begin{array}{ll} \text{minimizar} & \mathbf{g}^{k^t} \mathbf{d} + \frac{1}{2} \mathbf{d}^t \mathbf{W}^k \mathbf{d} \\ \text{sujeito a} & \mathbf{c}^k + \mathbf{A}^{k^t} \mathbf{d} = 0 \end{array} \quad (5.23)$$

Ou seja, cada iteração k da solução do problema original é idêntica à solução do PQ obtido pela linearização das restrições e pela expansão quadrática de f em torno de x_0 .

Em problemas em que todas as restrições são de igualdade, a direção de busca e os multiplicadores de Lagrange podem ser obtidos pela solução do sistema de equações lineares gerado pelo método de Newton, como mostrado em (5.22).

Para considerar o caso de restrições de desigualdade, Wilson propôs resolver o problema geral de PM:

$$\begin{array}{lll} \text{minimizar} & f(x) & \\ \text{sujeito a} & c_i(x) = 0 & i = 1 \dots l \\ & c_i(x) \leq 0 & i = l + 1 \dots m \end{array} \quad (5.24)$$

definindo uma direção de busca $\mathbf{d}$ e uma nova estimativa dos multiplicadores de Lagrange λ através da solução do PQ:

$$\begin{array}{lll} \text{minimizar} & \mathbf{g}^{k^t} \mathbf{d} + \frac{1}{2} \mathbf{d}^t \mathbf{W}^k \mathbf{d} & \\ \text{sujeito a} & c_i^k + \mathbf{a}_i^{k^t} \mathbf{d} = 0 & i = 1 \dots l \\ & c_i^k + \mathbf{a}_i^{k^t} \mathbf{d} \leq 0 & i = l + 1 \dots m \end{array} \quad (5.25)$$

cujo método de solução foi visto na seção anterior.

5.7.1

Etapas do Algoritmo Não-Linear Han-Powell (PQS)

As etapas que formam o algoritmo Han-Powell são (Parente, 2000):

1. Dado um ponto inicial x_0 e uma aproximação da Hessiana da função Lagrangiana $\mathbf{B}_0$ fazer $k=0$. $\mathbf{B}_0$ é dada pela seguinte função:

$$\mathbf{B}_0 = b_o \mathbf{I} \quad (5.26)$$

onde b_o é um parâmetro definido pelo usuário do algoritmo. O número de reinícios da matriz $\mathbf{B}$ é controlado pelo parâmetro n_r definido pelo

usuário. Segundo Parente (2000) o reinício de **B** serve para descartar a influencia de pontos muito distantes do ponto corrente.

2. Para $k=k+1$, montar e resolver o problema de programação quadrática definido pela equação (5.25) determinando os vetores $\mathbf{d}^k$ e λ^k :

$$\begin{array}{ll} \text{Minimizar} & \mathbf{g}^{k-1^t} \mathbf{d} + \frac{1}{2} \mathbf{d}^t \mathbf{B}^{k-1} \mathbf{d} \quad \mathbf{d} \in \mathfrak{R}^n \\ \text{sujeito a} & c_i^{k-1} + \mathbf{a}_i^{k-1^t} \mathbf{d} = 0 \quad i = 1 \dots l \\ & c_i^{k-1} + \mathbf{a}_i^{k-1^t} \mathbf{d} \leq 0 \quad i = l+1 \dots m \end{array} \quad (5.27)$$

onde c_i^{k-1} é o vetor com as restrições, $\mathbf{a}_i^{k-1^t}$ é uma matriz com o gradiente das restrições e $\mathbf{B}^{k-1}$ é uma aproximação da Hessiana no ponto $\mathbf{x}^{k-1}$.

3. Verificar os critérios de parada do algoritmo:

$$\begin{cases} \left| \mathbf{g}^{k-1^t} \mathbf{d}^k \right| \leq tol_1 \\ \max(c_i^k) \leq tol_2 \end{cases} \quad (5.28)$$

onde o primeiro critério representa a variação da função objetivo na direção $\mathbf{d}^k$ e o segundo critério verifica explicitamente o valor da restrição mais violada.

4. Faz-se então uma busca linear unidimensional para determinar o tamanho do passo t^k , na direção $\mathbf{d}^k$ de forma que o novo estimador da solução $\mathbf{x}^k = \mathbf{x}^{k-1} + t^k \mathbf{d}^k$ seja um ponto que contribua para o decréscimo da função objetivo. A busca é feita sobre a função de penalidade (p), construída no intuito de impor um alto custo à violação das restrições. Esta função é definida pela expressão:

$$p(t) = p(\mathbf{x} + t\mathbf{d}) = f(\mathbf{x}) + \sum_{i=1}^l r_i |c_i(\mathbf{x})| + \sum_{i=l+1}^m r_i \max[c_i(\mathbf{x}), 0] \quad (5.29)$$

onde os r_i são os fatores de penalidades. A busca é aproximada, isto é a solução t^* não é o mínimo de $p(t)$, mas atende a um certo decréscimo

pré-estipulado em $p(t)$ considerado satisfatório. O coeficiente de decréscimo da função é pelo parâmetro γ definido pelo usuário.

5. Atualização da matriz $\mathbf{B}^k$ do subproblema quadrático através do método BFGS.
6. Retorno à etapa 2.

5.8

Método dos Pontos Interiores

O algoritmo utilizado neste trabalho baseia-se na aplicação do método de Newton para a solução do sistema de equações não-lineares obtidas a partir da aplicação das condições de Kuhn-Tucker do problema de otimização (Herskovitz, 1995). Apenas o algoritmo para restrições de desigualdade é discutido, uma vez que os problemas de projeto ótimo a serem resolvidos não possuem restrições de igualdade. No entanto, as mesmas idéias aqui apresentadas também são válidas para os problemas que possuem simultaneamente restrições de igualdade e de desigualdade e podem ser vistas com mais detalhes em (Herskovitz, 1995; Herskovitz & Santos, 1997). O algoritmo de Pontos Interiores foi implementado e aplicado a problemas de Engenharia Estrutural no DEC/PUC-Rio por Parente (2000).

O método de Pontos Interiores tem como característica gerar uma sequência de pontos no interior da região viável que converge para a solução do problema. Outra propriedade importante destes algoritmos é que cada um dos pontos intermediários possui valores decrescentes da função objetivo.

Considere o problema de otimização:

$$\begin{array}{ll} \text{minimizar} & f(x) \\ \text{sujeito a} & c_i(x) \leq 0 \quad i = 1 \dots m \end{array} \quad (5.30)$$

cujas condições de Kuhn-Tucker são:

$$\begin{aligned} \mathbf{g} + \sum_{i=1}^m \lambda_i \mathbf{a}_i &= 0 \\ \lambda_i^* c_i(x^*) &= 0 \\ c_i(x^*) &\leq 0 \\ \lambda_i^* &\geq 0 \end{aligned} \quad (5.31)$$

Sendo $\mathbf{A}$ a matriz dos gradientes das restrições e $\mathbf{C}$ uma matriz diagonal contendo os valores das restrições, as duas primeiras equações podem ser escritas como:

$$\begin{aligned}\mathbf{g} + \mathbf{A}^t \boldsymbol{\lambda} &= 0 \\ \mathbf{C} \boldsymbol{\lambda} &= 0\end{aligned}\tag{5.32}$$

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

$$\begin{bmatrix} \mathbf{W} & \mathbf{A}^t \\ \boldsymbol{\Lambda} \mathbf{A} & \mathbf{C} \end{bmatrix} \begin{Bmatrix} \mathbf{d}_0 \\ \boldsymbol{\lambda}_0 \end{Bmatrix} = - \begin{Bmatrix} \mathbf{g} \\ 0 \end{Bmatrix}\tag{5.33}$$

Na equação acima, $\boldsymbol{\Lambda}$ é uma matriz diagonal para a qual $\Lambda_{ii} = \lambda_i$, $\mathbf{d}_0$ é a direção de busca e $\boldsymbol{\lambda}_0$ é a estimativa dos multiplicadores de Lagrange. Pode-se demonstrar que $\mathbf{d}_0$ é uma direção de decréscimo de f e que $\mathbf{d}_0 = 0$ se $\mathbf{x}$ for um ponto estacionário (Parente, 2000).

O problema é que a direção de busca fornecida por (5.33) nem sempre é uma direção viável. Expandindo-se uma equação da parte inferior do sistema (5.33), chega-se a:

$$\lambda_i \mathbf{a}_i^t \mathbf{d}_0 + c_i \lambda_{0_i} = 0\tag{5.34}$$

Esta equação implica que $\mathbf{a}_i^t \mathbf{d}_0 = 0$ para todo i tal que $c_i = 0$. Geometricamente, isto significa que $\mathbf{d}_0$ é tangente às restrições ativas, indicando que a direção aponta para fora da região viável.

Uma solução para evitar este efeito é adicionar uma constante negativa do lado direito da equação acima:

$$\lambda_i \mathbf{a}_i^t \mathbf{d} + c_i \bar{\lambda}_i = -\rho \lambda_i\tag{5.35}$$

onde $\bar{\lambda}_i$ é a nova estimativa de λ_i .

Este procedimento faz com que a direção original seja defletida, de um valor proporcional a ρ , para o interior da região viável. Como a deflexão é proporcional

a ρ e $\mathbf{d}_0$ é uma direção de decréscimo de f , é possível encontrar limites em ρ para que $\mathbf{d}$ também seja uma direção de decréscimo. Este objetivo pode ser atingido impondo-se que:

$$\mathbf{g}^t \mathbf{d} \leq k_a \mathbf{g}^t \mathbf{d}_0 \quad (5.36)$$

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 o sistema auxiliar:

$$\begin{bmatrix} \mathbf{W} & \mathbf{A}^t \\ \Lambda \mathbf{A} & \mathbf{C} \end{bmatrix} \begin{Bmatrix} \mathbf{d}_1 \\ \lambda_1 \end{Bmatrix} = - \begin{Bmatrix} \mathbf{g} \\ \lambda \end{Bmatrix} \quad (5.37)$$

é fácil mostrar que:

$$\mathbf{d} = \mathbf{d}_0 + \rho \mathbf{d}_1 \quad (5.38)$$

e

$$\bar{\lambda} = \lambda_0 + \rho \lambda_1 \quad (5.39)$$

Substituindo (5.38) em (5.36) chega-se a:

$$\rho \leq (k_a - 1) \frac{\mathbf{g}^t \mathbf{d}_0}{\mathbf{g}^t \mathbf{d}_1} \quad (5.40)$$

Definida a direção de busca, é necessário realizar uma busca linear restrita ao longo dessa direção, de forma a garantir que o ponto gerado esteja no interior da região viável. Além disso, é necessário atualizar os valores dos multiplicadores de Lagrange de maneira a assegurar a convergência para a solução correta.

5.8.1

Etapas do Algoritmo de Pontos Interiores

O algoritmo de Pontos Interiores para problemas de restrições de desigualdade necessita de um ponto inicial viável x_0 , uma estimativa para os multiplicadores de Lagrange de forma que $\lambda_i > 0$ e uma matriz $\mathbf{B}$ simétrica e

positiva definida, que é uma aproximação de $\mathbf{W}$. O algoritmo pode ser dividido nos seguintes passos (Herskovits & Santos, 1997):

1. Obter a direção de busca $\mathbf{d}$:

- a) Determinar os vetores $(\mathbf{d}_0, \lambda_0)$ através da solução do sistema linear definido em (5.33).
- b) Verificar o critério de convergência:

$$\|\mathbf{d}\| \leq tol \quad (5.41)$$

- c) Determinar os vetores $(\mathbf{d}_1, \lambda_1)$ através da solução do sistema linear definido em (5.37).
- d) Calcular o valor de ρ :

$$\begin{cases} \text{se } \mathbf{g}^t \mathbf{d}_1 > 0, \text{ então } \rho = \min \left[k_f \|\mathbf{d}_0\|^2, (k_a - 1) \mathbf{g}^t \mathbf{d}_0 / \mathbf{g}^t \mathbf{d}_1 \right] \\ \text{se } \mathbf{g}^t \mathbf{d}_1 \leq 0, \text{ então } \rho = k_f \|\mathbf{d}_0\|^2 \end{cases} \quad (5.42)$$

sendo $k_f > 0$.

- e) Calcular a direção de busca $\mathbf{d}$:

$$\mathbf{d} = \mathbf{d}_0 + \rho \mathbf{d}_1 \quad (5.43)$$

e

$$\bar{\lambda} = \lambda_0 + \rho \lambda_1 \quad (5.44)$$

- 2. Fazer busca linear sobre $\mathbf{d}$, determinando o tamanho do passo t que satisfaça um critério sobre o decréscimo da função objetivo e para o qual:

$$\begin{cases} c_i(x + t\mathbf{d}) \leq 0, & \text{se } \bar{\lambda}_i \geq 0 \\ c_i(x + t\mathbf{d}) \leq c_i(x), & \text{se } \bar{\lambda}_i < 0 \end{cases} \quad (5.45)$$

e o novo ponto x :

$$x = x_0 + t\mathbf{d} \quad (5.46)$$

3. Atualizar a matriz $\mathbf{B}$, que é uma aproximação da Hessiana da função Lagrangiana, através do método BFGS.
4. Definir uma nova estimativa para os multiplicadores de Lagrange:

$$\lambda_i = \max \left[\lambda_{0_i}, k_e \|\mathbf{d}_0\|^2 \right] \quad (5.47)$$

sendo $k_e > 0$.

5. Fazer x igual a x_0 e retornar ao passo 1.

A aproximação inicial e o reinício da Hessiana da função Lagrangiana são controlados pelos mesmos parâmetros utilizados pelo algoritmo de Programação Quadrática Sequencial.