# Algoritmo de Restauração Inexata aplicado à resolução de problemas de Programação Matemática em Dois Níveis

Suzana L. C. Castro Ana Friedlander

Julho 2004

Neste trabalho, estamos propondo uma nova abordagem para resolver problemas de programação matemática em dois níveis. Ela permite resolver diretamente o problema do segundo nível, sem reformulação, e sem a necessidade de recorrer a técnicas de otimização não diferenciável.

Isto é possível com a utilização do algoritmo de Restauração Inexata apresentado em ([8]). Na verdade, esta idéia surgiu ao nos depararmos com a seguinte sugestão neste artigo de Martínez & Pilotta :

*"...The philosophy of Inexact Restoration encourages case-oriented applications. Very relevant is the use of IR methods for Bilevel Programming, because we think that, in this important family of problems, the strategy of using oportunistic methods in the lower level (Feasibility Phase) could be useful...."*

# 1 Técnica de Restauração Inexata

A técnica de Restauração Inexata foi proposta por Martínez & Pilotta ([5], [7] e [6]) para resolver problemas de programação não linear. Assim como nos métodos viáveis, dentre os quais destacamos o método do gradiente reduzido generalizado (ver, por exemplo, [4]), ele caracteriza-se por tratar a viabilidade e otimalidade em diferentes fases, a cada iteração.

Porém, em oposição a estes, os métodos de restauração inexata permitem que o grau de precisão requerido na fase da restauração seja controlado a cada iteração, automaticamente.

O algoritmo que estudamos está descrito em [6] e considera o seguinte problema de programação não linear:

$$\begin{array}{ll} \text{Minimizar} & f(x) \\ \text{s.a} & \begin{cases} x \in \Omega \\ C(x) = 0 \end{cases} \end{array} \tag{1}$$

onde $\Omega$ é convexo e fechado, $f : I\!R^n \to I\!R$, $C : I\!R^n \to I\!R^m$ são tais que $\nabla f(x)$ e $C'(x)$ existem e são contínuas.

Basicamente, o algoritmo consiste num método iterativo onde cada iteração é composta por duas fases distintas. A primeira destas fases, que é denominada fase da restauração, tem como objetivo melhorar a viabilidade. A segunda fase, chamada fase da minimização, consiste em reduzir o valor de uma função de mérito, baseada no Lagrangeano, em um conjunto viável aproximado.

Os pontos da seqüência gerada pelo algoritmo são viáveis em relação a $\Omega$, mas não necessariamente em relação a $C(x) = 0$ e a distância entre duas aproximações sucessivas é controlada por uma região de confiança.

## 1.1 O Algoritmo de Restauração Inexata

Suponhamos que sejam dados os seguintes parâmetros de entrada:

$\eta > 0$, $r \in (0,1)$, $\beta > 0$, $M > 0$, $\theta_{-1} \in (0,1)$, $\delta_{min} > 0$, $x^0 \in \Omega$, $\lambda^0 \in I\!R^m$ e $w_k$ tal que $\sum_{k=0}^{\infty} w_k < \infty$

Então, os passos de uma iteração do algoritmo de Restauração ([6]) estão descritos no algoritmo a seguir:

**Algoritmo 1: Restauração Inexata**

*Passo 1.*

Definir
$\theta_k^{min} = min\{1, \theta_{k-1}, ..., \theta_{-1}\}$,
$\theta_k^{large} = min\{1, \theta_k^{min} + w_k\}$,
$\theta_{k,-1} = \theta_k^{large}$

*Passo 2:* Fase da Restauração.

Calcular $y^k \in \Omega$ tal que

$$\|C(y^k)\| \leq r\|C(x^k)\| \tag{2}$$

e

$$\|y^k - x^k\| \leq \beta\|C(x^k)\| \tag{3}$$

*Passo 3:* Cálculo da direção tangente de Cauchy.

Calcular

$$d_{tan}^k = P_k\left[y^k - \eta\nabla L(y^k,\lambda^k)\right] - y^k,$$

onde $P_k(z)$ é a projeção ortogonal de $z$ em $\Pi_k$ onde

$$\Pi_k = \left\{z \in \Omega | C'(y^k)(z-y^k) = 0\right\}$$

é uma aproximação linear do conjunto viável passando por $y^k$ (denominado Espaço Tangente Aproximado)

**Se** $y^k = x^k$ (e então $C(x^k) = C(y^k)$) e $d_{tan}^k = 0$,
**Então**, terminar a execução do algoritmo e retornar $x^k$ como a solução.

**Se** $d_{tan}^k = 0$
**Então**,
calcular $\lambda^{k+1} \in I\!R^m$ tal que $\|\lambda^{k+1}\| \leq M$,
definir
$x^{k+1} = y^k$ , $\theta_k = \theta_{k-1}$,

$$Ared_k = (1-\theta_k)\left[\|C(x^k)\| - \|C(y^k)\|\right]$$

e terminar a iteração.

**Senão**,
fazer $i = 0$, escolher $\delta_{k,0} \geq \delta_{min}$ e continuar.

*Passo 4:* Fase da Minimização no Espaço Tangente Aproximado.

Calcular $z^{k,i} \in \Pi_k$ tal que

$$L(z^{k,i},\lambda^k) \ll L(y^k,\lambda^k) \tag{4}$$

e

$$\|z^{k,i} - y^k\| \leq \delta_{k,i} \tag{5}$$

onde $L(x,\lambda) = f(x) + C(x)^T\lambda$ é a função Lagrangeana e $\delta_{k,i}$ é o raio da região de confiança.

*Passo 5:* Atualização dos multiplicadores.

Calcular $\lambda^{k,i}_{atual} \in I\!R^m$ tal que $\|\lambda^{k,i}_{atual}\| \leq M$.

*Passo 6:* Cálculo da redução Prevista.

Definir, para todo $\theta \in [0, 1]$,

$$\begin{aligned} Pred_{k,i}(\theta) = \quad & \theta \left[ L(x^k, \lambda^k) - L(z^{k,i}, \lambda^k) - \langle C(y^k), \lambda^{k,i}_{atual} - \lambda^k \rangle \right] \\ & +(1-\theta) \left[ \|C(x^k)\| - \|C(y^k)\| \right] \end{aligned} \tag{6}$$

Calcular $\theta_{k,i}$, o máximo dos elementos $\theta \in [0, \theta_{k,i-1}]$ que verifica

$$Pred_{k,i}(\theta) \geq 1/2 \left[ \|C(x^k)\| - \|C(y^k)\| \right]$$

Definir

$$Pred_{k,i} = Pred_{k,i}(\theta_{k,i})$$

*Passo 7:* Comparação da redução real e da prevista.

Calcular

$$Ared_{k,i} = \theta_{k,i} \left[ L(x^k, \lambda^k) - L(z^{k,i}, \lambda^{k,i}_{atual}) \right] + (1 - \theta_{k,i}) \left[ \|C(x^k)\| - \|C(z^{k,i})\| \right] \tag{7}$$

**Se** $Ared_{k,i} \geq 0.1 Pred_{k,i}$
**Então**,
definir $x^{k+1} = z^{k,i}$,
$\lambda^{k+1} = \lambda^{k,i}_{atual}$,
$\theta_k = \theta_{k,i}$,
$\delta_k = \delta_{k,i}$, $Ared_k = Ared_{k,i}, Pred_k = Pred_{k,i}$
e terminar a iteração $k$.

**Senão**,
escolher $\delta_{k,i+1} \in [0.1\delta_{k,i}, 0.9\delta_{k,i}]$, fazer $i = i + 1$ e voltar para o Passo 4.

Fim

As condições (2)-(3) são denominadas condições de restauração. A primeira condição garante que a viabilidade seja melhorada, ou restaurada, reduzindo o valor de $\|C(x)\|$ (já que um ponto viável é tal que $C(x) = 0$). Esta redução é proporcional ao grau de inviabilidade de $x^k$, permitindo menos esforço nas iterações iniciais.

A segunda condição, (3), exige que o ponto obtido na fase da restauração esteja "suficientemente" próximo de $x^k$, ou seja, pertença a uma bola centrada em $x^k$ e cujo raio depende do grau de inviabilidade de $x^k$. Isto evita que o algoritmo retorne a um ponto já explorado e cicle.

Observamos que se $x^k$ é viável, e então $C(x^k) = 0$, as condições de restauração são automaticamente satisfeitas e portanto $y^k = x^k$.

As condições (4)-(5) são denominadas condições de minimização. Elas determinam o decréscimo da função Lagrangeana, no espaço formado pela linearização das restrições e em uma região de confiança centrada no ponto obtido na fase da restauração. Este decréscimo exigido, denotado no algoritmo pelo símbolo $\ll$, está relacionado ao decréscimo obtido pelo ponto de Cauchy (ver [7], [6]).

As funções Ared em (7) e Pred em (6) são obtidas através de combinações convexas da viabilidade e da otimalidade, onde o peso é dado pela variável $\theta$. Elas são utilizadas como critério de aceitação de uma nova aproximação. Quando o ponto resultante da fase da minimização não é aceito por este critério, o algoritmo retorna ao passo 4, ou seja, conserva o ponto obtido na fase da restauração. Isto indica uma certa preferência do método pela viabilidade, o que para problemas práticos, em geral, é bastante interessante.

Uma das características do algoritmo de Restauração Inexata é a liberdade que temos na escolha dos métodos para realizar os passos 2 (fase da restauração) e 4 (fase da minimização). Isso permite explorar as características e propriedades particulares de cada problema, principalmente da sua região viável.

Porém, é claro, que muito da eficiência do algoritmo de Restauração Inexata também está relacionada a esta escolha. Algoritmos gerais, e que podem ser utilizados na realização destas fases, foram apresentados por Martínez ([5]) e Martínez & Pilotta ([7]). Basicamente, para a fase da restauração, eles consideram o problema de minimizar a norma da inviabilidade, enquanto que para a fase da minimização eles consideram um problema de busca linear ao longo da direção de Cauchy.

Uma proposta para realizar a fase da minimização, consiste em reescrever as condições de minimização (4)-(5) como o seguinte problema de otimização:

$$\begin{array}{ll} \underset{z}{\text{Minimizar}} & L(z,\lambda) \\ \text{s.a} & \begin{cases} z \in \Pi_k \\ \|z - y^k\| \leq \delta_k \end{cases} \end{array} \tag{8}$$

onde $\Pi_k = \left\{z \in \Omega | C'(y^k)(z - y^k) = 0\right\}$ e $L(x, \lambda) = f(x) + C(x)^T\lambda$.

Por sua vez, este problema pode ser visto como um problema de otimização com restrições lineares se tomarmos $\|.\|$ como a norma do supremo, já que $\Pi_k$ é um conjunto linear. Com isso, a fase da restauração torna-se resolver o problema (8) com uma precisão dependendo do valor da função objetivo avaliada no ponto de Cauchy.

Em Martínez & Pilotta ([7]) são apresentados resultados numéricos que mostram um bom desempenho do algoritmo de Restauração Inexata, na resolução de problemas de programação não linear.

## 1.2 Critérios de Convergência

Para garantir que o Algoritmo 1 esteja bem definido e provar a sua convergência, são necessárias as seguintes hipóteses sobre o problema (1):

(H1) $\Omega$ é convexo e compacto.

(H2) Existe $L_1 > 0$ tal que, para todo $x, y, \in \Omega$,

$$|C'(x) - C'(y)| \leq L_1|x - y|$$

(H3) Existe $L_2 > 0$ tal que, para todo $x, y \in \Omega$,

$$|\nabla f(x) - \nabla f(y)| \leq L_2|x - y|$$

O teorema seguinte garante a convergência do Algoritmo 1:

**Teorema 1.1.** *Suponhamos que o problema (1) satisfaz as hipóteses (H1)-(H3). Se a fase da restauração (Passo 2) pode ser completada com sucesso em cada iteração, então, a seqüência* $\{x^k\}$ *gerada pelo Algoritmo 1, é tal que:*

*(i)* $|C(x^k)| \to 0$

*(ii) Todo ponto limite de* $\{x^k\}$ *é viável*

*(iii) Se* $x^*$ *é um ponto limite de* $\{x^k\}$ *e* $\lim_{k\in K_1} x^k = x^*$*, então*

$$\lim_{k\in K_1} y^k = y^*$$

*onde* $y^k$ *é o ponto obtido na fase da restauração, na iteração* $k$*.*

*(iv) Existe um conjunto infinito* $K_2 \subset \{0, 1, 2, ...\}$ *tal que*

$$\lim_{k\in K_2} d^k_{tan} = 0$$

*(v) Existe um conjunto infinito* $K_3 \subset \{0, 1, 2, ...\}$ *e* $x^* \in \Omega$ *tal que*

$$\lim_{k\in K_3} x^k = \lim_{k\in K_3} y^k = x^*,$$

$$C(x^*) = 0,$$

$$\lim_{k\in K_3} d^k_{tan} = 0$$

*Além disso, se* $x^*$ *é um ponto regular, ele satisfaz as condições KKT de otimalidade do problema (1).*

A prova deste teorema está em [6]. Em linhas gerais, ele garante que, independentemente de condições sobre as restrições, existe uma subseqüência, da seqüência gerada pelo Algoritmo 1, convergente a um ponto viável e cujas direções de Cauchy $d^k_{tan}$ convergem a zero.

A otimalidade dos pontos limite está relacionada às direções $d^k_{tan}$, que são direções de descida no espaço de minimização.

Como descrito em [8], pode-se mostrar que estes pontos limite satisfazem a condição AGP de otimalidade, proposta por Martínez & Svaiter ([9]), que é uma condição satisfeita por minimizadores locais de problemas de programação não linear, independentemente de hipóteses de regularidade.

Para o problema (1), o vetor AGP $g(x)$, avaliado em $x \in I\!R^n$, é definido por:

$$g(x) = P(x - \nabla f(x)) - x$$

onde $P(z)$ é a projeção de $z$ no espaço tangente aproximado da região viável de (1).

Um ponto $x^*$ satisfaz a condição AGP se existe uma seqüência $\{x^k\}$ que converge a $x^*$ e tal que $g(x^k) \to 0$.

Para garantir a existência de uma subseqüência satisfazendo as condições de viabilidade e otimalidade do teorema 1.1, é necessário utilizar o fato de que as condições (2)-(3) e o conjunto viável $\Pi_k$ estarem definidos através do conjunto de restrições $C(x)$. Assim, as equações de determinam $C$ são bastante importantes tanto na definição das fases do algoritmo, como na prova de convergência.

# 2 Algoritmo de Restauração Inexata aplicado à resolução de problemas de Programação Matemática em Dois Níveis

Motivadas com o fato de que, por um lado, o algoritmo de Restauração Inexata permite explorar a estrutura da região viável de cada problema, e de que, por outro lado, a região viável dos problemas em dois níveis deveria ser melhor explorada, propomos a nova abordagem de utilizar a técnica de Restauração Inexata na resolução de problemas em dois níveis.

Uma outra razão, que também motiva o uso dessa técnica, é a vantagem em poder privilegiar a viabilidade. Os problemas em dois níveis, em geral, estão associados a problemas com restrições de equilíbrio, o que implica que a viabilidade pode ser de extrema importância.

Com o intuito de concretizar esta proposta, estaremos preocupadas com dois aspectos básicos: o teórico e o computacional. No aspecto teórico, além de adequar os passos do algoritmo de Restauração Inexata (Algoritmo 1) para problemas em dois níveis, já que ele foi originalmente proposto para problemas clássicos de programação não linear, pretendemos estudar condições teóricas que garantam sua convergência.

Além disso, a viabilidade de implementação computacional da técnica proposta, será sempre considerada.

## 2.1 Definição do Problema

Neste trabalho, estamos interessadas em estudar o seguinte problema de programação matemática em dois níveis:

$$\begin{array}{ll} \underset{x,y}{\text{Minimizar}} & F(x,y) \\ \text{s.a} & \left\{ \begin{array}{ll} y = \underset{y}{\text{argmin}} & f(x,y) \\ \text{s.a} & \left\{ \begin{array}{rl} h(x,y) = & 0 \\ y \geq & 0 \end{array} \right. \end{array} \right. \end{array} \tag{9}$$

onde $x \in X \subset I\!R^{n_x}$, $y \in Y \subset I\!R^{n_y}$, $F, f : I\!R^{n_x+n_y} \rightarrow I\!R$, $h : I\!R^{n_x+n_y} \rightarrow I\!R^p$ e tais que $\nabla F(x,y)$, $\nabla f(x,y)$ e $h'(x,y)$ existem e são contínuas.

Denotamos com PN2($x$) o problema do segundo nível

$$\begin{array}{ll} \underset{y}{\text{Minimizar}} & f(x,y) \\ \text{s.a} & \left\{ \begin{array}{rl} h(x,y) = & 0 \\ y \geq & 0 \end{array} \right. \end{array} \tag{10}$$

Assumimos que (9) satisfaz as seguintes hipóteses:

(H1) $X$ e $Y$ são convexos e compactos.

(H2) O problema PN2($x$) tem solução única $y(x)$ para cada valor de $x \in X$ e $y(x)$ é o único ponto estacionário de PN2($x$).

(H3) A função que a cada $x$ associa $y(x)$ a solução de PN2($x$), é contínua.

(H4) Para todo par $(x, y(x))$ com $x \in X$ e $y(x)$ a solução de PN2($x$), temos que $(x, y(x))$ é regular, no sentido de que o conjunto formado pelos gradientes das restrições ativas em $(x, y(x))$ é linearmente independente.

As condições de otimalidade de Karush-Kuhn-Tucker (KKT) do problema PN2($x$) são

$$\begin{array}{rll} \nabla_y f(x,y) + \nabla_y h(x,y)\mu - \gamma & = 0 & \\ h(x,y) & = 0 & \\ y_i \gamma_i & = 0 & (i = 1, \ldots, n_y) \\ y & \geq 0 & \\ \gamma & \geq 0 & \end{array} \tag{11}$$

onde $\mu \in I\!R^p$ e $\gamma \in I\!R^{n_y}$ são os vetores dos multiplicadores de Lagrange associados às restrições $h(x,y) = 0$ e $y \geq 0$, respectivamente, e $h'_y(x,y)$ representa a matriz jacobiana de $h(x,y)$ em relação a $y$.

Definimos a função $C : I\!R^{n_x+2n_y+p} \rightarrow I\!R^{2n_y+p}$ como:

$$C(x,y,\mu,\gamma) = \begin{pmatrix} \nabla_y f(x,y) + \nabla_y h(x,y)\mu - \gamma \\ h(x,y) \\ y_1\gamma_1 \\ \vdots \\ y_{n_y}\gamma_{n_y} \end{pmatrix} \tag{12}$$

Assumimos as seguintes hipóteses adicionais sobre (9)

(H5) Para todo $x \in X$ e $(y(x), \mu(x), \gamma(x))$ tais que $C(x, y(x), \mu(x), \gamma(x)) = 0$,

$$y^T \nabla^2_y l(x, y(x), \mu(x), \gamma(x)) y > 0$$

para todo $y \in I\!R^{n_y}$, com $y \neq 0$, e onde $l(x,y,\mu,\gamma) = f(x,y) + h(x,y)^T\mu$.

(H6) Para todo $y(x)$, solução de PN2($x$), temos

$$y_i(x) = 0 \Longrightarrow \gamma_i(x) > 0$$

e

$$\gamma_i(x) = 0 \Longrightarrow y_i(x) > 0$$

para todo $i = 1, \ldots, n_y$.

(Complementaridade estrita).

Provamos a seguir alguns resultados que usaremos mais adiante.

**Lema 2.1.** $C'_{y,\mu,\gamma}(x, y(x), \mu(x), \gamma(x))$ *é não singular para todo* $x \in X$.

*Prova.* Pela definição de $C'_{y,\mu,\gamma}(x,y,\mu,\gamma)$ temos:

$$C'_{y,\mu,\gamma}(x,y,\mu,\gamma) = \begin{pmatrix} H & A^T & -I \\ A & 0 & 0 \\ \mathrm{diag}(\gamma) & 0 & \mathrm{diag}(y) \end{pmatrix}$$

onde $H$ é a matriz hessiana do Lagrangeano em relação a $y$ ( definida em (H5)), $A = h'_y(x, y)$ é a matriz jacobiana das restrições e diag(z)$\in I\!R^{n_y \times n_y}$ é uma matriz diagonal, com componentes $z_i$.

Sejam $u \in I\!R^{n_y}$, $v \in I\!R^p$ e $w \in I\!R^{n_y}$ tais que:

$$C'_{y,\mu,\gamma}(x, y(x), \mu(x), \gamma(x)) \begin{pmatrix} u \\ v \\ w \end{pmatrix} = 0$$

ou seja,

$$Hu + A^T v - w = 0 \tag{13}$$

$$Au = 0 \tag{14}$$

$$\text{diag}(\gamma(x))u + \text{diag}(y(x))w = 0 \tag{15}$$

Utilizando a hipótese (H6) e a equação (15) temos que:

$$u^T w = 0 \tag{16}$$

Multiplicado a equação (13) por $u^T$ temos:

$$u^T H u + u^T A^T v - u^T w = 0$$

Utilizando (14) e (16) vem

$$u^T H u = 0$$

e pela hipótese (H5),

$$u = 0$$

Assim, reescrevendo o sistema temos:

$$A^T v - w = 0 \tag{17}$$

$$\text{diag}(y(x))w = 0 \tag{18}$$

Suponhamos, sem perda de generalidade que $y_i(x) = 0$ para $i = 1, ...n_c$ e, portanto $\gamma_i(x) = 0$ para $i = n_c + 1, ..., n_y$.

Por (18) temos então que

$$w_i = 0, \quad i = n_c + 1, .., n_y$$

Podemos escrever (17) como:

$$\begin{pmatrix} A^T & -I_{n_y} \end{pmatrix} \begin{pmatrix} v \\ w \end{pmatrix} = 0$$

ou ainda

$$\begin{pmatrix} A^T & -I_{n_c} \end{pmatrix} \begin{pmatrix} v \\ w_{n_c} \end{pmatrix} = 0$$

onde $I_{n_y}$ representa a matriz identidade de ordem $n_y$, $I_{n_c} \in I\!R^{n_y \times n_c}$ a matriz formada pelas $n_c$ primeiras colunas da matriz $I_{n_y}$ e $w_{n_c}$ é o vetor de dimensão $n_c$ que corresponde às variáveis não nulas de $w$.

Pela hipótese (H4) a matriz $\begin{pmatrix} A^T & -I_{n_c} \end{pmatrix}$ tem posto completo e portanto

$$v = 0 \quad w_{n_c} = 0$$

Concluímos que $(u, v, w) = (0, 0, 0)$ e portanto $C'_{y,\mu,\gamma}(x, y(x), \mu(x), \gamma(x))$ é não singular para todo $x \in X$. □

As hipótese anteriores, principalmente (H5), não permitem que o resultado do lema 2.1 seja aplicado a problemas em dois níveis do tipo linear. Porém, é possível obter um resultado semelhante, considerando, ao invés de (H5), uma hipótese específica:

(H5') $y(x)$, a única solução de PN2$(x)$, corresponde a um vértice do politopo que descreve a região viável ($Ay - b = 0$, com $A \in I\!R^{p \times n_y}$, $p < n_y$ e $b \in I\!R^p$ ) e é uma solução não degenerada. Neste caso, existem portanto $n_y$ restrições ativas na solução, ou seja:

$$\begin{array}{l} Ay(x) = b \\ y_i(x) = 0 \quad (i = p+1, .., n) \end{array}$$

**Lema 2.2.** *Suponhamos que (9) seja um problema linear satisfazendo (H1)-(H4) e (H5')-(H6). Então $C'_{y,\mu,\gamma}(x, y(x), \mu(x), \gamma(x))$ é não singular para todo $x \in X$.*

*Prova.* Pela definição de $C'_{y,\mu,\gamma}(x, y, \mu, \gamma)$ temos:

$$C'_{y,\mu,\gamma}(x, y, \mu, \gamma) = \begin{pmatrix} 0 & A^T & -I \\ A & 0 & 0 \\ \text{diag}(\gamma) & 0 & \text{diag}(y) \end{pmatrix}$$

onde diag(z)$\in \mathbb{R}^{n_y \times n_y}$ é uma matriz diagonal, com componentes $z_i$.

Suponhamos, sem perda de generalidade que $A = (B\ N)$, com $B \in \mathbb{R}^{p\times p}$ não singular.

Sejam $u \in \mathbb{R}^{n_y}$, $v \in \mathbb{R}^p$ e $w \in \mathbb{R}^{n_y}$, tais que:

$$\begin{pmatrix} 0 & A^T & -I \\ A & 0 & 0 \\ \text{diag}(\gamma(x)) & 0 & \text{diag}(y(x)) \end{pmatrix} \begin{pmatrix} u \\ v \\ w \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

ou seja,

$$A^T v - w = 0 \tag{19}$$

$$Au = 0 \tag{20}$$

$$\gamma_i(x)u_i + y_i(x)w_i = 0 \tag{21}$$

Por (H5'), (H6) e (21) temos:

$$\begin{array}{l} w_i = 0 \quad i = 1, .., p \\ \gamma_i(x) = 0 \quad i = 1, .., p \\ u_i = 0 \quad i = p+1, .., n \end{array} \tag{22}$$

Por (H4) e (19) temos

$$\begin{array}{l} v = 0 \\ w_i = 0 \quad i = p+1, .., n \end{array} \tag{23}$$

Por (H4), (20) e por (22) temos

$$u_i = 0 \quad i = 1, .., p \tag{24}$$

Então, por (22), (23) e (24), concluímos que $(u, v, w) = (0, 0, 0)$ e portanto $C(x, y(x), \mu(x), \gamma(x))$ é não singular. $\square$

**Lema 2.3.** *(Limitação local para o erro):*

*Seja $x \in X$ dado e $(y(x), \mu(x), \gamma(x))$ tal que $C(x, y(x), \mu(x), \gamma(x)) = 0$. Se $\mu(x)$ e $\gamma(x)$ são funções contínuas, então existem $\beta > 0$ e $\varepsilon > 0$ tais que*

$$\|(y, \mu, \gamma) - (y(x), \mu(x), \gamma(x))\| \leq \beta \, \|C(x, y, \mu, \gamma)\|$$

*para todo $(y, \mu, \gamma) \in V_{y,\mu,\gamma}((y(x), \mu(x), \gamma(x)), \varepsilon)$.*

*Prova.* Como $C(x,y,\mu,\gamma)$ é continuamente diferenciável e

$$C'_{y,\mu,\gamma}(x,y(x),\mu(x),\gamma(x))$$

é não singular, pelo teorema da função inversa existe $\varepsilon_1 > 0$ tal que

$$C(x,y,\mu,\gamma)$$

possui uma inversa diferenciável $G$ em $V_{y,\mu,\gamma}((x,y(x),\mu(x),\gamma(x)),\varepsilon_1)$. Então

$$G\left(C(x,y,\mu,\gamma)\right) = (x,y,\mu,\gamma) \tag{25}$$

e

$$G'(C(x,y,\mu,\gamma)) = C'_{y,\mu,\gamma}((x,y,\mu,\gamma))^{-1} \tag{26}$$

para $(x,y,\mu,\gamma) \in V_{y,\mu,\gamma}((x,y(x),\mu(x),\gamma(x)),\varepsilon_1)$.

Como $X$ é compacto, $y(x)$, $\mu(x)$ e $\gamma(x)$ são uniformemente contínuas então $\varepsilon_1$ não depende de $x$.

Expandindo $G$ em série de Taylor, em torno de $C(x,y(x),\mu(x),\gamma(x))$, obtemos:

$$\begin{gathered} G\left(C(x,y,\mu,\gamma)\right) = G\left(C(x,y(x),\mu(x),\gamma(x))\right) + \\ G'\left(C(x,y(x),\mu(x),\gamma(x))\right)\left(C(x,y,\mu,\gamma) - C(x,y(x),\mu(x),\gamma(x))\right) + \\ R\left(C(x,y,\mu,\gamma) - C(x,y(x),\mu(x),\gamma(x))\right) \end{gathered}$$

onde

$$\lim_{C(x,y,\mu,\gamma)\to C(x,y(x),\mu(x),\gamma(x))} \frac{R\left(C(x,y,\mu,\gamma) - C(x,y(x),\mu(x),\gamma(x))\right)}{\|C(x,y,\mu,\gamma) - C(x,y(x),\mu(x),\gamma(x)\|} = 0$$

Utilizando (25) e (26) e o fato de que $C(x,y(x),\mu(x),\gamma(x)) = 0$, temos que

$$\begin{aligned} (x,y,\mu,\gamma) = \ & (x,y(x),\mu(x),\gamma(x)) + C'(x,y(x),\mu(x),\gamma(x))^{-1}\, C(x,y,\mu,\gamma) + \\ & R\left(C(x,y,\mu,\gamma)\right) \end{aligned}$$

E portanto

$$(x,y,\mu,\gamma)-(x,y(x),\mu(x),\gamma(x)) = C'(x,y(x),\mu(x),\gamma(x))^{-1}C(x,y,\mu,\gamma)+ R\left(C(x,y,\mu,\gamma)\right)$$

onde

$$\lim_{C(x,y,\mu,\gamma)\to 0} \frac{R\left(C(x,y,\mu,\gamma)\right)}{\|C(x,y,\mu,\gamma)\|}=0$$

Calculando a norma temos:

$$\|(y,\mu,\gamma)-(y(x),\mu(x),\gamma(x))\| \leq |C'(x,y(x),\mu(x),\gamma(x))^{-1}|\|C(x,y,\mu,\gamma)\|+ \|R(C(x,y,\mu,\gamma))\| \tag{27}$$

Agora, dado $\alpha>0$, pela definição de $R$, existe $\varepsilon_2>0$ tal que

$$\|C(x,y,\mu,\gamma)\|\leq\varepsilon_2 \;\Rightarrow\; \left\|\frac{R\left(C(x,y,\mu,\gamma)\right)}{\|C(x,y,\mu,\gamma)\|}\right\|\leq\alpha.$$

Como $C$ é contínua e $C(x,y(x),\mu(x),\gamma(x))=0$, existe $\varepsilon_3>0$ tal que

$$\|C(x,y,\mu,\gamma)\|\leq\varepsilon_2 \tag{28}$$

para todo $(y,\mu,\gamma)\in V_{y,\mu,\gamma}((y(x),\mu(x),\gamma(x)),\varepsilon_3)$.

Assim, dado $\alpha>0$

$$\|R\left(C(x,y,\mu,\gamma)\right)\|\leq\alpha\|C(x,y,\mu,\gamma)\|$$

para todo

$$(x,y,\mu,\gamma)\in V_{y,\mu,\gamma}(y(x),\mu(x),\gamma(x)),\varepsilon_3)\cap V_{y,\mu,\gamma}((y(x),\mu(x),\gamma(x)),\varepsilon_1)$$

Utilizando (28), a equação (27) pode ser escrita como:

$$\|(y,\mu,\gamma)-(y(x),\mu(x),\gamma(x))\|\leq$$
$$|C'(y(x)(x,,\mu(x),\gamma(x))^{-1}|\;\|C(x,y,\mu,\gamma)\|+\alpha\|C(x,y,\mu,\gamma)\|$$

para todo

$$(x,y,\mu,\gamma) \in V_{y,\mu,\gamma}((y(x),\mu(x),\gamma(x)),\varepsilon_3) \cap V_{y,\mu,\gamma}((y(x),\mu(x),\gamma(x)),\varepsilon_1)$$

ou seja,

$$\|(y,\mu,\gamma) - (y(x),\mu(x),\gamma(x))\| \leq$$
$$\left(|C'(x,y(x),\mu(x),\gamma(x))^{-1}| + \alpha\right) \|C(x,y,\mu\gamma)\|$$

para todo

$$(x,y,\mu,\gamma) \in V_{y,\mu,\gamma}((y(x),\mu(x),\gamma(x)),\varepsilon_3) \cap V_{y,\mu,\gamma}((y(x),\mu(x),\gamma(x)),\varepsilon_1)$$

A continuidade uniforme das funções envolvidas garante que podemos escolher $\varepsilon_2$ e $\varepsilon_3$ independentemente de $x$. Assim, tomando $\varepsilon = min\{\varepsilon_3, \varepsilon_1\}$ e

$$\beta \geq \left(|C'_{y,\mu,\gamma}((x,y(x),\mu(x),\gamma(x))^{-1}| + \alpha\right) > 0 \tag{29}$$

temos o resultado desejado.

A existência de $\beta$ verificando (29) para qualquer $x \in X$ resulta do fato de $X$ ser compacto e das funções $y(x)$, $\mu(x)$ e $\gamma(x)$ serem contínuas. Isto implica que a imagem de $X$ é um compacto e como $C'_{y,\mu,\gamma}(x,y(x),\mu(x),\gamma(x))$ nunca é singular, $|C'_{y,\mu,\gamma}((x,y(x),\mu(x),\gamma(x))^{-1}|$ está limitado superiormente em $X$. □

No próximo lema, mostramos que existe um limitante global para o erro.

**Lema 2.4.** *(Limitação Global para o erro):*
*Sejam $x \in X$ dado e $(y(x),\mu(x),\gamma(x))$ como no lema anterior. Então existe $\beta > 0$ tal que,*

$$\|(y,\mu,\gamma) - (y(x),\mu(x),\gamma(x))\| \leq \beta \|C(x,y,\mu,\gamma)\|$$

*para todo $(x,y,\mu,\gamma) \in \Omega$*

*Prova.* Pelo lema 2.3, temos que existem $\beta_1 > 0$ e $\varepsilon > 0$ tais que,

$$\|(x, y, \mu, \gamma) - (y(x), \mu(x), \gamma(x))\| \leq \beta_1 \|C(x, y, \mu, \gamma))\|$$

para todo $(y, \mu, \gamma) \in V_{y,\mu,\gamma}((y^*, \mu^*, \gamma^*), \varepsilon)$.

Para $x \in X$ definimos os conjuntos

$$\Omega = \{(y, \mu, \gamma) \mid y \in Y,\ \mu \in \Gamma,\ \gamma \in \Theta\}$$

$$\Omega^\varepsilon = \{(y, \mu, \gamma) \in \Omega \mid \|(y, \mu, \gamma) - (y(x), \mu(x), \gamma(x))\| \geq \varepsilon\}$$

O conjunto $\Omega^\varepsilon$ possui as seguintes propriedades:

(P1) $\Omega^\varepsilon$ é compacto (fechado contido em $\Omega$)

(P2) $(y(x), \mu(x), \gamma(x)) \notin \Omega^\varepsilon$

(P3) Como a função $\|C(x, y, \mu, \gamma)\|$ é uniformemente contínua em $\Omega^\varepsilon$, ela assume valores máximo $(M)$ e mínimo $(m)$, independentes de $x$, ou seja,

$$m \leq \|C(x, y, \mu, \gamma)\| \leq M \quad , \qquad \forall (x, y, \mu, \gamma) \in \Omega^\varepsilon.$$

com $m > 0$, por (H2).

(P4) Como a função $\|(y, \mu, \gamma) - (y(x), \mu(x), \gamma(x))\|$ é uniformemente em contínua em $\Omega^\varepsilon$ ela assume valores máximo $(N)$ e mínimo $(n)$, independentes de $x$, ou seja,

$$n \leq \|(y, \mu, \gamma) - (y(x), \mu(x), \gamma(x))\| \leq N \quad , \qquad \forall (x, y, \mu, \gamma) \in \Omega^\varepsilon.$$

com $n > 0$, por (P2).

Consideremos agora um ponto $(y, \mu, \gamma) \in \Omega$. Temos, então duas possibilidades:

(i) $(y, \mu, \gamma) \in V_{y,\mu,\gamma}((y(x), \mu(x), \gamma(x)), \varepsilon)$

Neste caso, temos o resultado desejado para o $\beta$ obtido no lema anterior.

(ii) $(y, \mu, \gamma) \notin V_{y,\mu,\gamma}((y(x), \mu(x), \gamma(x)), \varepsilon)$

Utilizando as propriedades (P3) e (P4) do conjunto $\Omega^{\varepsilon}$, temos:

$$\frac{\|(x, y, \mu, \gamma) - (y(x), \mu(x), \gamma(x))\|}{\|C(x, y, \mu, \gamma)\|} \leq \frac{N}{m}$$

onde $\frac{N}{m} > 0$, pois $N$ e $m > 0$

Então para $\beta \geq max\{\beta_1, \frac{N}{m}\}$ temos o resultado desejado. □

O resultado que segue é essencial para provar a boa definição do algoritmo que vamos propor para resolver problemas de dois níveis com as hipóteses dadas.

**Teorema 2.1.** *Seja $x \in X$ e $(y(x), \mu(x), \gamma(x))$ como no lema 2.3. Seja $\{(y^k, \mu^k, \gamma^k)\}$ uma seqüência convergente ao ponto $(y(x), \mu(x), \gamma(x))$ tal que $C(x, y(x), \mu(x), \gamma(x)) = 0$. Dados $r \in (0, 1)$ e $(y, \mu, \gamma) \in \Omega$, existem $\beta > 0$ e $n \in N$ tais que $(y^n, \mu^n, \gamma^n)$ satisfaz*

$$\|C(x, y^n, \mu^n, \gamma^n)\| \leq r\|C(x, y, \mu, \gamma)\|$$

*e*

$$\|(y, \mu, \gamma) - (y^n, \mu^n, \gamma^n)\| \leq \beta\|C(x, y, \mu, \gamma)\|$$

*Prova.* Se $C(x, y, \mu, \gamma) \neq 0$, existe $N$ tal que

$$\forall n > N \Rightarrow \|C(y^n, \mu^n, \gamma^n) - C(x, y(x), \mu(x), \gamma(x))\| < r\|C(x, y, \mu, \gamma)\|$$

como $C(x, y(x), \mu(x), \gamma(x)) = 0$ temos que:

$$\forall n > N \Rightarrow \|C(x, y^n, \mu^n, \gamma^n)\| < r\|C(x, y, \mu, \gamma)\|$$

Utilizando os lemas 2.4 e 2.1, temos que para cada $x \in X$ existe $\beta > 0$ tal que,

$$\|(y, \mu, \gamma) - (y(x), \mu(x), \gamma(x))]\| \leq \beta\|C(x, y, \mu, \gamma)\|$$

Utilizando a desigualdade triangular, temos que:

$$\|(y,\mu,\gamma)-(y^n,\mu^n,\gamma^n)\| \leq$$
$$\|(y,\mu,\gamma)-(y(x),\mu(x),\gamma(x))\| + \|(y(x),\mu(x),\gamma(x)-(y^n,\mu^n,\gamma^n)\|$$

$$\leq \beta\|C(x,y,\mu,\gamma))\| + r*\beta\|C(x,y,\mu,\gamma)\|$$

Se renomeamos $(\beta + r*\beta)$, obtemos o resultado desejado. □

Estamos interessadas na resolução de (9) através do algoritmo de restauração inexata apresentado no Capítulo 2 (Algoritmo 1). Este método, no entanto, foi originalmente proposto para tratar problemas de otimização com estrutura do tipo (1), na qual não se enquadra de maneira óbvia o problema (9).

A principal diferença entre os problemas está no conjunto de restrições. Enquanto que em (1), o conjunto de restrições é descrito por equações do tipo $C(x) = 0$, em (9) ele é descrito pelas soluções de diferentes problemas de otimização, PN2($x$).

Assim, adequar o Algoritmo 1 para resolver o problema (9), significa particularizar todos os passos que dependam das restrições. Mais especificamente, é obter uma maneira de medir a viabilidade, para descrever as condições da fase da restauração, e definir o problema da fase da minimização.

Para isso, utilizamos uma função auxiliar $C$, definida pelas condições KKT de (10) que, na fase da restauração, é utilizada como medida de viabilidade e, na fase da minimização, sua derivada é utilizada para descrever o espaço viável.

Na próxima seção apresentamos uma descrição detalhada do algoritmo e a maneira adequada para realizar as diferentes fases para problemas do tipo (9).

O resultado de convergência obtido para este algoritmo está baseado no Teorema 1.1 de Martínez ([6]), que garante, considerando algumas hipóteses sobre o problema, que a seqüência gerada pelo algoritmo possui uma subseqüência convergente a pontos viáveis e que satisfazem uma condição de otimalidade, supondo que a fase da restauração possa ser completada com sucesso em todas as iterações.

Com as hipóteses dadas sobre o problema em dois níveis (9) e através de uma função $C$, associada às condições KKT de (10), provamos que a fase da restauração pode ser completada com sucesso a cada iteração e, utilizando $C'$ (jacobiana da função associada às condições KKT) para definir o conjunto viável aproximado do problema auxiliar da fase de minimização mostramos que o algoritmo proposto se enquadra na teoria do Teorema 1.1.

Neste sentido, os resultados obtidos nesta seção mostram que para todo parâmetro $r \in (0,1)$, existe $\beta(r) > 0$, tal que é possível obter um ponto satisfazendo a condição de restauração (2) e que necessariamente a condição (3) será satisfeita. Como conseqüência deste resultado, o parâmetro $\beta$ e a condição (3) podem ser excluídos do algoritmo.

## 2.2 Algoritmo

Nesta seção apresentamos o algoritmo para resolver o problema em dois níveis descrito na seção anterior e mostramos que a nossa proposta se enquadra na teoria de convergência apresentada no capítulo 2.

Dados $\Gamma \in I\!R^p$ e $\Theta \in I\!R^{n_y}$ conjuntos convexos e compactos, definimos o conjunto $\Omega$ por:

$$\Omega = \{(x, y, \mu, \gamma) \in I\!R^{n_x+2n_y+p} \mid x \in X,\ y \in Y_0,\ \mu \in \Gamma,\ \gamma \in \Theta_0\} \tag{30}$$

onde

$$Y_0 = Y \cap \{y \in Y \mid y \geq 0\}$$

e

$$\Theta_0 = \Theta \cap \{\gamma \in \Theta \mid \gamma \geq 0\}$$

Supomos, ainda, que são dados os seguintes parâmetros de entrada:

$\eta > 0$, $r \in (0,1)$, $M > 0$, $\theta_{-1} \in (0,1)$, $\delta_{min} > 0$, $x^0 \in X$, $y^0 \in Y_0$, $\lambda^0 \in I\!R^{2n_y+p}$, $w_k$ tal que $\sum_{k=0}^{\infty} w_k < \infty$ e $\mu^0 \in \Gamma$, $\gamma^0 \in \Theta_0$, aproximações dos vetores dos multiplicadores de Lagrange associados a PN2($x^0$).

Então, os passos de uma iteração do algoritmo de Restauração Inexata para resolver o problema em dois níveis do tipo (9), estão descritos no algoritmo a seguir:

**Algoritmo 3: Restauração Inexata para problemas em dois níveis**

*Passo 1*

Definir
$\theta_k^{min} = min\{1, \theta_{k-1}, ..., \theta{-1}\}$,
$\theta_k^{large} = min\{1, \theta_k^{min} + w_k\}$,
$\theta_{k,-1} = \theta_k^{large}$

*Passo 2:* Fase da Restauração.

Determinar $y^R \in Y_0$, $\mu^R \in \Gamma$ e $\gamma^R \in \Theta_0$ tais que:

$$\|C(x^k, y^R, \mu^R, \gamma^R)\| \leq r\|C(x^k, y^k, \mu^k, \gamma^k).$$

*Passo 3:* Direção Tangente de Cauchy.

Calcular

$$d_{tan}^k = P_k\left[(x^k, y^R, \mu^R, \gamma^R) - \eta\nabla L(x^k, y^R, \mu^R, \gamma^R, \lambda^k)\right] - (x^k, y^R, \mu^R, \gamma^R)$$

onde $P_k(x, y, \mu, \gamma)$ é a projeção ortogonal de $(x, y, \mu, \gamma)$ em $\Pi_k$, com

$$\Pi_k = \left\{(x, y, \mu, \gamma) \in \Omega \mid C'(x^k, y^R, \mu^R, \gamma^R)((x, y, \mu, \gamma) - (x^k, y^R, \mu^R, \gamma^R)) = 0\right\}$$

e

$$L(x, y, \mu, \gamma, \lambda) = F(x, y) + C(x, y, \mu, \gamma)^T\lambda$$

**Se** $(x^k, y^R, \mu^R, \gamma^R) = (x^k, y^k, \mu^k, \gamma^k)$ e $d_{tan}^k = 0$,
**Então** terminar a execução do algoritmo e retornar $(x^k, y^k)$ como a solução.

**Se** $d_{tan}^k = 0$
**Então**,

calcular $\lambda^{k+1} \in I\!R^{2n_y+p}$ tal que $\|\lambda^{k+1}\| \leq M$,
definir $(x^{k+1}, y^{k+1}, \mu^{k+1}, \gamma^{k+1}) = (x^k, y^R, \mu^R, \gamma^R)$ ,
$\theta_k = \theta_{k-1}$,

$$Ared_k = (1 - \theta_k) \left[ \|C(x^k, y^k, \mu^k, \gamma^k)\| - \|C(x^k, y^R, \mu^R, \gamma^R)\| \right]$$

e terminar a iteração.

**Senão**, fazer $i = 0$, escolher $\delta_{k,0} \geq \delta_{min}$ e continuar.

*Passo 4:* Fase da Minimização no Espaço Tangente Aproximado.

Calcular $(x^{k,i}, y^{k,i}, \mu^{k,i}, \gamma^{k,i}) \in \Pi_k$, tal que:

$$L(x^{k,i}, y^{k,i}, \mu^{k,i}, \gamma^{k,i}, \lambda^k) \ll L(x^k, y^R, \mu^R, \gamma^R, \lambda^k)$$

e

$$(x^{k,i}, y^{k,i}, \mu^{k,i}, \gamma^{k,i}) - (x^k, y^R, \mu^R, \gamma^R)\| \leq \delta_k$$

onde o símbolo $\ll$ representa, como no Algoritmo 1, um decréscimo suficiente da função Lagrangeana em relação ao decréscimo obtido pelo ponto de Cauchy.

*Passo 5:* Atualização dos Multiplicadores.

Calcular $\lambda_{atual}^{k,i} \in I\!R^{2n_y+p}$ tal que $\|\lambda_{atual}^{k,i}\| \leq M$.

*Passo 6:* Cálculo da Redução Prevista.

Definir, para todo $\theta \in [0, 1]$,

$$\begin{aligned} Pred_{k,i}(\theta) = \quad & \theta \left[ L(x^k, y^k, \mu^k, \gamma^k, \lambda^k) - \right. \\ & \left. L(x^{k,i}, y^{k,i}, \mu^{k,i}, \gamma^{k,i}, \lambda^k) - \langle C(x^k, y^R, \mu^R, \gamma^R), \lambda_{atual}^{k,i} - \lambda^k \rangle \right] \\ & +(1 - \theta) \left[ \|C(x^k, y^k, \mu^k, \gamma^k)\| - \|C(x^k, y^R, \mu^R, \gamma^R)\| \right] \end{aligned}$$

Calcular $\theta_{k,i}$, o máximo dos elementos $\theta \in [0, \theta_{k,i-1}]$ que verifica

$$Pred_{k,i}(\theta) \geq 1/2 \left[ \|C(x^k, y^k, \mu^k, \gamma^k)\| - \|C(x^k, y^R, \mu^R, \gamma^R)\| \right].$$

Definir

$$Pred_{k,i} = Pred_{k,i}(\theta_{k,i})$$

*Passo 7:* Comparação da Redução Real e da Prevista.

Calcular

$$\begin{aligned} Ared_{k,i} = \; & \theta_{k,i}\left[L(x^k, y^k, \mu^k, \gamma^k, \lambda^k) - L(x^{k,i}, y^{k,i}, \mu^{k,i}, \gamma^{k,i}, \lambda^{k,i}_{atual})\right] \\ & +(1-\theta_{k,i})\left[\|C(x^k, y^k, \mu^k, \gamma^k)\| - \|C(x^{k,i}, y^{k,i}, \mu^{k,i}, \gamma^{k,i})\|\right] \end{aligned}$$

**Se** $Ared_{k,i} \geq 0.1 Pred_{k,i}$
**Então**,
definir $x^{k+1} = x^{k,i}$, $y^{k+1} = y^{k,i}$, $\mu^{k+1} = \mu^{k,i}$,
$\gamma^{k+1} = \gamma^{k,i}$
$\lambda^{k+1} = \lambda^{k,i}_{atual}$,
$\theta_k = \theta_{k,i}$,
$\delta_k = \delta_{k,i}$,
$Ared_k = Ared_{k,i}, Pred_k = Pred_{k,i}$
e terminar a iteração $k$.

**Senão**,
escolher $\delta_{k,i+1} \in [0.1\delta_{k,i}, 0.9\delta_{k,i}]$, fazer $i = i + 1$ e voltar para o Passo 4.

Fim

## 2.3 Fase da Restauração

A fase da restauração consiste em obter um ponto que seja mais viável e suficientemente próximo do ponto obtido na iteração anterior. Isto é descrito no Passo 2 do Algoritmo 1 através das condições de restauração (2)-(3).

Estas condições, utilizam o valor da norma $\|C(x)\|$ como medida da viabilidade ou inviabilidade de um ponto $x$.

Para o problema (9) não temos uma maneira tão natural para medir a viabilidade, como em (1). Neste caso, então, para poder descrever as condições de restauração é preciso uma ferramenta auxiliar, que permita medir a viabilidade (ou a inviabilidade) de um ponto.

Uma maneira de medir a viabilidade de um ponto é avaliando uma função com valores reais, positiva e que se anule nos pontos viáveis do problema.

No caso do problema de programação não linear (1) esta função é dada por:

$$\varphi(x) = \|C(x)\|$$

de modo que se $\bar{x}$ é um ponto viável de (1), então $C(\bar{x}) = 0$ e portanto $\varphi(\bar{x}) = 0$.

Já para o problema (9), a definição de viabilidade está associada à proximidade com a solução do problema de otimização do segundo nível. Portanto, a função que caracterize esta viabilidade deve estar relacionada a estas soluções.

Neste trabalho, propomos a função, definida pela norma do vetor associado às condições KKT do problema do segundo nível, seja utilizada para medir a viabilidade de um ponto.

Esta escolha está motivada no fato de que, sob hipóteses de regularidade, o sistema KKT se anula em uma solução de (10), ou seja, em um ponto viável de (9).

Definimos a função $C : I\!R^{n_x+2n_y+p} \to I\!R^{2n_y+p}$ como:

$$C(x,y,\mu,\gamma) = \begin{pmatrix} \nabla_y f(x,y) + \nabla_y h(x,y)\mu - \gamma \\ h(x,y) \\ y_1\gamma_1 \\ \vdots \\ y_{n_y}\gamma_{n_y} \end{pmatrix} \tag{31}$$

Assumimos as seguintes hipóteses sobre $C(x,y,\mu,\gamma)$

(H7) Existem $L_1 > 0$ e $L_2 > 0$ tais que, para todo $(x_1,y_1,\mu_1,\gamma_1) \in \Omega$, $(x_2,y_2,\mu_2,\gamma_2) \in \Omega$, temos

$$\|\nabla F(x_1,y_1) - \nabla F(x_2,y_2)\| \le L_1\|(x_1,y_1) - (x_2,y_2)\|$$

e

$$|C'(x_1,y_1,\mu_1,\gamma_1) - C'(x_2,y_2,\mu_2,\gamma_2)| \le L_2\|(x_1,y_1,\mu_1,\gamma_1) - (x_2,y_2,\mu_2,\gamma_2)\|$$

Agora, utilizando (12), definimos a função $\varphi : I\!R^{n_x+2n_y+p} \to I\!R$, que mede a inviabilidade de um ponto no problema (9), como sendo:

$$\varphi(x, y, \mu, \gamma) = \|C(x, y, \mu, \gamma)\| \tag{32}$$

com $C(x, y, \mu, \gamma)$ dada em (12).

Utilizando a função de medida da inviabilidade (32) podemos escrever as condições de restauração do algoritmo de Restauração Inexata para o problema em dois níveis (9) da seguinte forma:

Dados $(x, y, \mu, \gamma) \in \Omega$, $r \in (0, 1)$ e $\beta > 0$, determinar $\bar{y} \in Y_0$, $\bar{\mu} \in \Gamma$ e $\bar{\gamma} \in \Theta_0$ tais que:

$$\|C(x, \bar{y}, \bar{\mu}, \bar{\gamma})\| \leq r\|C(x, y, \mu, \gamma)\| \tag{33}$$

e

$$\|(y, \mu, \gamma) - (\bar{y}, \bar{\mu}, \bar{\gamma})\| \leq \beta\|C(x, y, \mu, \gamma)\| \tag{34}$$

com $C(x, y, \mu, \gamma)$ é definida em (12).

Os resultados obtidos na seção anterior garantem a existência de $\beta > 0$ para o qual um ponto satisfazendo (33), automaticamente satisfaz a condição (34). Portanto, como conseqüência, o parâmetro $\beta$ e condição (34) podem ser excluídas do algoritmo.

## 2.4 Estimadores dos multiplicadores de Lagrange

Os vetores $\mu \in \Gamma \subset I\!R^p$ e $\gamma \in \Theta_0 \subset I\!R^{n_y}$ correspondem a um conjunto de variáveis que são introduzidas ao problema com a definição da função $C$.

Os vetores $\mu$ e $\gamma$ devem estar associados a um par $(x, y) \in X \times Y_0$, como estimativas para os multiplicadores de Lagrange associados a estes pontos.

O critério de determinação destes vetores é importante para garantir a convergência do algoritmo. Eles caracterizam o critério de parada do algoritmo, na fase da restauração e, na fase da minimização, eles definem o espaço viável.

As seguintes hipóteses sobre os estimadores são necessárias:

(E1) As funções $\tilde{\mu}(x, y(x))$ e $\tilde{\gamma}(x, y(x))$, que a cada $x$ associam as aproximações para os multiplicadores de Lagrange, são contínuas.

(E2) Seja $x^k \in X$ e $y(x^k) \in Y_0$ a solução de PN2($x^k$). Se $\{\mu^n\} \subset \Gamma$ e $\{\gamma^n\} \subset \Theta_0$ são seqüências dos vetores de multiplicadores de Lagrange aproximados, então existem $\{\mu^{n_i}\}$ e $\{\gamma^{n_i}\}$ subseqüências de $\{\mu^n\}$ e $\{\gamma^n\}$, respectivamente, tais que

$$\{\mu^{n_i}\} \to \mu^* \quad \text{e} \quad \{\gamma^{n_i}\} \to \gamma^*$$

onde $C(x^k, y(x^k), \mu(x^k), \gamma(x^k)) = 0$

Muitos algoritmos de otimização fornecem estimativas para os multiplicadores de Lagrange a cada iteração. Principalmente quando estão baseados em métodos do tipo Lagrangeano Aumentado.

Assim, quando um algoritmo com esta propriedade for utilizado na realização da fase de restauração, as estimativas podem ser aproveitadas, desde que satisfaçam (E1)-(E2).

Quando isto não for possível, uma fórmula para calcular estas estimativas deve ser incorporada na rotina de resolução do problema do segundo nível. No entanto, muitas vezes, isto pode ser complicado, já que alguns algoritmos não permitem alteração. Então, uma técnica do tipo ”parada forçada do algoritmo” a cada conjunto de iterações, para realização do teste de parada, é aconselhável.

## 2.5 Fase da Minimização

A fase da minimização consiste em diminuir o valor da função Lagrangeana, no espaço definido pela linearização das restrições, em uma região de confiança centrada no ponto obtido na fase da restauração. Isto é descrito no Passo 4 do Algoritmo 1 através das condições de minimização (4)-(5).

Estas condições utilizam as restrições em $C(x)$ para definir a função Lagrangeana e o espaço de minimização.

Neste trabalho, utilizamos a função $C$ definida em (12) para descrever as restrições do problema que resolvemos na fase da minimização. Além de visar a necessidade de funções diferenciáveis para definir as restrições, esta escolha torna-se bastante adequada se levarmos em conta a prova de convergência do Algoritmo 1. Ela explora o fato da função Lagrangeana, do espaço tangente aproximado e da medida de inviabilidade estarem todos definidos através de uma mesma função.

As condições de minimização do algoritmo de Restauração Inexata para o problema em dois níveis (9) pode ser descrita da seguinte forma:

Dados $(x^k, y^R, \mu^R, \gamma^R) \in \Omega$ satisfazendo a condição de restauração (33) e $\delta_k > 0$, determinar um ponto

$$(x^M, y^M, \mu^M, \gamma^M) \in \Pi_k(x^k, y^R, \mu^R, \gamma^R)$$

tal que:

$$L(x^M, y^M, \mu^M, \gamma^M, \lambda^k) \ll L(x^k, y^R, \mu^R, \gamma^R, \lambda^k) \tag{35}$$

e

$$\|(x^M, y^M, \mu^M, \gamma^M) - (x^k, y^R, \mu^R, \gamma^R)\| \leq \delta \tag{36}$$

onde

$$\Pi_k(x^k, y^R, \mu^R, \gamma^R) =$$
$$\big\{(x, y, \mu, \gamma) \in \Omega \mid C'(x^k, y^R, \mu^R, \gamma^R)\big((x, y, \mu, \gamma) - (x^k, y^R, \mu^R, \gamma^R)\big) = 0\big\}$$

é o conjunto viável de minimização, denominado espaço tangente aproximado, a função $C(x, y, \mu, \gamma)$ está definida em (12), e o símbolo $\ll$ significa uma redução suficiente em relação ao ponto de Cauchy (ver [7], [6]), como no Algoritmo 1.

O conjunto viável de minimização está definido através de $C'$, o jacobiano da função $C$ associada às restrições KKT do problema do segundo nível. Portanto, este conjunto depende das derivadas de segunda ordem das funções envolvidas no problema do segundo nível.

Na prática, os pontos satisfazendo as condições (35)-(36) podem ser obtidos através da resolução do seguinte problema de otimização com restrições lineares:

$$\begin{array}{ll} \text{Minimizar} & L(x, y, \mu, \gamma, \lambda^k) \\ \text{s.a} & \left\{ \begin{array}{l} (x, y, \mu, \gamma) \in \Pi_k(x^k, y^R, \mu^R, \gamma^R) \\ \|(x, y, \mu, \gamma) - (x^k, y^R, \mu^R, \gamma^R)\|_\infty \leq \delta_k \end{array} \right. \end{array} \tag{37}$$

Na prática, (37) pode ser resolvido inexatamente, já que, por (35), apenas uma condição de decréscimo suficiente, relacionada ao ponto de Cauchy, deve ser satisfeita. Em [7], está sugerido um critério de parada associado à norma do gradiente.

## 2.6 Resultado de Convergência

Diremos que um algoritmo de otimização tem a propriedade de convergência global se ele obtém, em número finito de iterações, um ponto estacionário do problema, ou gera uma seqüência na qual todo ponto de acumulação é um ponto estacionário do problema. Vamos a assumir que um algoritmo deste tipo é usado para resolver cada PN2($x$).

Então para o Algoritmo 3 temos o seguinte resultado para problemas com as hipóteses assumidas previamente.

**Teorema 2.2.** *O Algoritmo 3 está bem definido e a seqüência gerada pelo Algoritmo 3,* $\{(x^k, y^k, \mu^k, \gamma^k)\}$*, é tal que:*

*(i)* $\|C(x^k, y^k, \mu^k, \gamma^k)\| \to 0$

*(ii) Todo ponto limite* $x^*$ *de* $\{(x^k, y^k, \mu^k, \gamma^k)\}$ *é viável, no sentido de que é a solução de PN2(*$x^*$*).*

*(iii) Se* $(x^*, y^*, \mu^*, \gamma^*)$ *é um ponto limite de* $\{(x^k, y^k, \mu^k, \gamma^k)\}$ *e*

$$\lim_{k \in K_1} (x^k, y^k, \mu^k, \gamma^k) = (x^*, y^*, \mu^*, \gamma^*)$$

*então*

$$\lim_{k \in K_1} (x^k_R, y^k_R, \mu^k_R, \gamma^k_R) = (x^*, y^*, \mu^*, \gamma^*)$$

*onde* $(x^k_R, y^k_R, \mu^k_R, \gamma^k_R)$ *é o ponto obtido na fase da restauração.*

*(iv) Existe um conjunto infinito* $K_2 \subset \{0, 1, 2, ...\}$ *tal que*

$$\lim_{k \in K_2} d^k_{tan} = 0$$

*(v) Existe um conjunto infinito* $K_3 \subset \{0, 1, 2, ...\}$ *e* $(x^*, y^*, \mu^*, \gamma^*) \in \Omega$ *tal que*

$$\lim_{k \in K_3} (x^k, y^k, \mu^k, \gamma^k) = \lim_{k \in K_3} (x^k_R, y^k_R, \mu^k_R, \gamma^k_R) = (x^*, y^*, \mu^*, \gamma^*),$$

$$C(x^*, y^*, \mu^*, \gamma^*) = 0,$$

$$\lim_{k \in K_3} d^k_{tan} = 0$$

Este resultado é conseqüência dos resultados provados nas outras seções.

## 2.7 Outras classes de Problemas em Dois Níveis

Nas seções anteriores apresentamos um algoritmo e o resultado de convergência, para resolver problemas em dois níveis do tipo (9).

Para classes mais gerais, não temos um resultado para garantir que a fase da Restauração possa ser realizada com sucesso em todas as iterações. A proposta de utilizar as condições KKT do problema do segundo nível para definir a função $C$ pode gerar algumas dificuldades.

Problemas mais gerais incluem, nas condições KKT, equações de desigualdade e, portanto, não se adequam ao modelo do Algoritmo de restauração Inexata utilizado.

Nestes casos, é possível utilizar uma outra versão do algoritmo de Restauração Inexata, apresentada em [7], apropriada para restrições de desigualdade. Porém, nesta versão a própria função objetivo é utilizada como mérito, o que pode não ser muito eficiente ao considerarmos a minimização no espaço linearizado.

Algumas classes de problemas podem serem reformulados de modo a incluírem-se na classe estudada nas seções anteriores. Dentre eles, destacamos a classe de problemas em dois níveis na qual o problema do segundo nível possui restrições de desigualdade. Através da introdução de variáveis de folga, as restrições de desigualdade podem ser convertidas em restrições de igualdade e não negatividade.

Problemas do segundo nível com restrições de canalização também podem ser incluídos nas classes estudadas, através de mudança de variável. Por outro lado, os problemas, como os lineares e quadráticos em dois níveis, são casos especiais de (9).

Problemas em dois níveis envolvendo restrições no primeiro nível, não se adequam à teoria desenvolvida neste trabalho. No entanto, temos uma proposta de algoritmo, semelhante à apresentada neste trabalho, mas ainda carente de um estudo teórico mais aprofundado.

Na prática, porém, muitos problemas podem ser resolvidos com sucesso. Uma sugestão de Martínez ([6]), é incluir um critério de parada para o caso de ocorrer uma falha na fase de restauração.

### 2.7.1 Problemas em Dois Níveis sem restrição no primeiro nível e com restrições de igualdade no segundo nível

O problema de programação matemática em dois níveis sem restrição no primeiro nível e com restrições de igualdade no segundo nível tem a forma:

$$\begin{array}{ll} \underset{x,y}{\text{Minimizar}} & F(x,y) \\ \text{s.a} & \begin{cases} y = \underset{y}{\text{argmin}} & f(x,y) \\ \qquad\quad \text{s.a} & h(x,y) = 0 \end{cases} \end{array} \tag{38}$$

onde $x \in X \subset I\!R^{n_x}$, $y \in Y \subset I\!R^{n_y}$, com $X$ e $Y$ conjuntos convexos e compactos, $F, f : I\!R^{n_x+n_y} \to I\!R$, $h : I\!R^{n_x+n_y} \to I\!R^p$ e tais que $\nabla F(x,y)$, $\nabla f(x,y)$ e $h'(x,y)$ existem e são contínuas.

O problema (38) é um caso particular de (9), excluindo as restrições de não negatividade. Neste caso, a função associada às restrições KKT do segundo nível de (38), como em (12), é dada por:

$$C(x,y,\mu) = \begin{pmatrix} \nabla_y f(x,y) + \nabla_y h(x,y)\mu \\ h(x,y) \end{pmatrix} \tag{39}$$

onde $\mu \in \Gamma \subset I\!R^p$ corresponde ao vetor dos multiplicadores de Lagrange aproximados associados ao ponto $(x,y)$.

O Algoritmo 3 pode ser utilizados para resolver problemas do tipo (38), considerando $C$ definida em (39) e $\gamma = 0$.

Quanto às condições de convergência, a hipótese (H6) não é mais necessária e o lema 2.1 pode ser escrito como:

**Lema 2.5.** *$C'_{y,\mu}(x, y(x), \mu(x))$, com $C$ definida em (39) é não singular.*

*Prova.* Pela definição de $C'_{y,\mu}(x^k, y, \mu)$ temos:

$$C'_{y,\mu}(x, y(x), \mu(x)) = \begin{pmatrix} H & A^T \\ A & 0 \end{pmatrix}$$

onde $H = L'_y$, $L = \nabla_y f(x, y(x)) + \nabla h_y(x, y(x))\mu$ e $A = h'_y(x, y(x))$.

Consideremos $u \in I\!R^{n_y}$ e $v \in I\!R^p$, tais que:

$$\begin{pmatrix} H & A^T \\ A & 0 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Na notação de sistema temos:

$$\left\{\begin{array}{ll} Hu + A^T v & = 0 \\ Au & = 0 \end{array}\right. \tag{40}$$

Multiplicando a primeira equação de (40) por $u^T$ e substituindo a segunda equação, temos:

$$u^T H u + u^T A^T v = 0$$

$$\Rightarrow u^T H u = 0$$

Como, pela hipótese (H5), $H$ é definida positiva temos:

$$u = 0 \tag{41}$$

Substituindo, então, $u = 0$ na primeira equação de (40), resulta em:

$$A^T v = 0$$

Como, pela hipótese (H4) $A$ tem posto completo, temos que:

$$v = 0 \tag{42}$$

Logo, de (41) e (42), concluímos que $C(x, y(x), \mu(x))$ é não singular.

$\square$

### 2.7.2 Problema em Dois Níveis sem restrição do primeiro nível e com restrições lineares de igualdade no segundo nível

Esta classe de problemas é bastante importante na programação matemática em dois níveis, visto que ela inclui um caso bastante estudado que é a dos problemas quadráticos em dois níveis.

O problema de programação matemática em Dois Níveis sem restrição no primeiro nível e com restrições lineares de igualdade no segundo nível é da forma:

$$\begin{array}{ll} \underset{x,y}{\text{Minimizar}} & F(x,y) \\ \text{s. a} & \left\{\begin{array}{ll} y = \underset{y}{\text{argmin}} & f(x,y) \\ & \text{s.a} \left\{\begin{array}{rr} A(x)y = & b(x) \\ y \geq & 0 \end{array}\right. \end{array}\right. \end{array} \tag{43}$$

onde $x \in X \subset I\!R^{n_x}$, $y \in Y \subset I\!R^{n_y}$, com $X$ e $Y$ conjuntos convexos e compactos, $F, f : I\!R^{n_x+n_y} \to I\!R$, tais que $\nabla F(x,y)$e $\nabla f(x,y)$ existem e são contínuas e $A \in I\!R^{p\times(n_x+n_y)}$ é uma matriz com $p \geq n_y$.

Quanto aos critérios de convergência do Algoritmo 3 na resolução de problemas do tipo (43), neste caso, a hipótese (H4) pode ser substituída por:

(H4') Para todo $x \in X$, a matriz $A(x)$ tem posto completo.

e (H5) pode ser substituída por:

(H5') Para todo $x \in X$ e $(y(x), \mu(x), \gamma(x))$ tal que $C(x, y(x), \mu(x), \gamma(x)) = 0$,

$$y^T \nabla^2_y f(x, y(x)) y > 0$$

para todo $y \in I\!R^{n_y}$, $y \neq 0$.

### 2.7.3 Problemas em Dois níveis com restrições de desigualdade no segundo nível

O problema de programação matemática em dois níveis com restrições de desigualdade no segundo nível é da forma:

$$\begin{array}{ll} \underset{x,y}{\text{Minimizar}} & F(x,y) \\ \text{s.a} & \left\{ \begin{array}{rl} y = \underset{y}{\text{argmin}} & f(x,y) \\ \text{s.a} & g(x,y) \leq 0 \end{array} \right. \end{array} \tag{44}$$

onde $x \in X \subset I\!R^{n_x}$, $y \in Y \subset I\!R^{n_y}$, com $X$ e $Y$ conjuntos convexos e compactos, $F, f : I\!R^{n_x+n_y} \to I\!R$, $g : I\!R^{n_x+n_y} \to I\!R^p$ e tais que $\nabla F(x,y)$, $\nabla f(x,y)$ e $g'(x,y)$ existem e são contínuas.

Neste caso, teríamos dificuldades em utilizar uma técnica semelhante à desenvolvida para problemas do tipo (9). Isto porque, para definir a função $C$, utilizamos as condições KKT do problema do segundo nível, que neste caso são:

$$\left\{ \begin{array}{rll} \nabla_y f(x,y) + g'(x,y)^T \mu = & 0 & \\ g(x,y) \leq & 0 & \\ g_i(x,y)\mu_i = & 0 & (i = 1, \ldots, p) \\ \mu \geq & 0 & \end{array} \right.$$

As condições acima incluem restrições de desigualdade, impedindo a definição de uma função $C(x,y,\mu,\gamma)$ com a propriedade $C(x,y(x),\mu(x),\gamma(x)) = 0$, como nos casos anteriores. No entanto, o problema (44) pode ser transformado a um problema do tipo (9) adicionando variáveis de folga, da seguiinte forma:

$$\begin{array}{ll} \underset{x,y}{\text{Minimizar}} & F(x,y,z) \equiv F(x,y) \\ \text{s.a} & \left\{ \begin{array}{l} (y,z) = \underset{y,z}{\text{argmin}} \quad f(x,y,z) \equiv f(x,y) \\ \qquad\qquad \text{s.a} \left\{ \begin{array}{rl} g(x,y) + z = & 0 \\ z \geq & 0 \end{array} \right. \end{array} \right. \end{array} \tag{45}$$

### 2.7.4 Restrições no Primeiro Nível

O estudo de problemas com restrições no primeiro nível não está incluído nos objetivos deste trabalho. Porém, durante nosso estudo, surgiram alguns problemas com estas características. Sabemos que estes problemas apresentam uma dificuldade adicional pela presença de regiões viáveis desconexas.

Na prática, podemos sugerir uma adaptação do Algoritmo 3 para problemas com restrições no primeiro nível, mas cujo estudo de convergência precisa ser muito mais cuidadoso e será deixado como trabalho futuro.

Consideremos o problema (9) incluindo restrições de igualdade no primeiro nível. Teremos então:

$$\begin{array}{ll} \underset{x,y}{\text{Minimizar}} & F(x,y) \\ \text{s.a} & \left\{ \begin{array}{l} H(x,y) = 0 \\ y \in \underset{y}{\text{argmin}} \quad f(x,y) \\ \qquad\quad \text{s.a} \quad h(x,y) = 0 \end{array} \right. \end{array} \tag{46}$$

onde $x \in X \subset I\!R^{n_x}$, $y \in Y \subset R^{n_y}$, $F, f : I\!R^{n_x+n_y} \rightarrow I\!R$, $H : I\!R^{n_x+n_y} \rightarrow I\!R^p$ e $h : I\!R^{n_x+n_y} \rightarrow I\!R^q$ são tais que $\nabla F(x,y)$, $\nabla f(x,y)$, $H'(x,y)$ e $h'(x,y)$ existem e são contínuas.

A região induzida do problema (46) pode ser desconexa, como citamos no Capítulo 2. Neste caso existe $x^k \in X$ tal que não existe $y^*$ solução do problema do segundo nível parametrizado por $x^k$ e portanto a hipótese (H2) não é satisfeita.

Supondo que isto não ocorra, por exemplo se $H$ só depende da variável do primeiro nível, então poderemos descrever um algoritmo semelhante ao Algoritmo 3.

Neste caso, a função $C$ deve, além de incluir as condições KKT do problema do segundo nível, considerar as restrições do primeiro nível. Ela pode ser definida por:

$$C(x,y,\mu) = \begin{pmatrix} \nabla_y f(x,y) + \mu^T h'_y(x,y) \\ h(x,y) \\ H(x,y) \end{pmatrix} \tag{47}$$

Além disso, a fase da restauração deve ser composta de duas etapas: a primeira para restaurar em relação às restrições do primeiro nível e a segunda para resolver o problema do segundo nível.

Sejam $(x^k, y^k, \mu^k) \in \Omega$, $r \in (0,1)$ e $\beta > 0$ dados. As etapas para a fase da restauração são:

<u>Etapa 1 da Fase da Restauração</u>

Obter $(x^{R1}, y^{R1})$ tal que

$$\|H(x^{R1}, y^{R1})\| \leq r\|H(x^k, y^k, \mu^k)\| \tag{48}$$

e

$$\|(y^{R1}, \mu^{R1})\| \leq \beta\|H(x^k, y^k, \mu^k)\| \tag{49}$$

<u>Etapa 2 da Fase da Restauração</u>

Obter $y^{R2}$, resolvendo inexatamente o problema:

$$\begin{array}{ll} \underset{y}{\text{Minimizar}} & f(x^{R1}, y) \\ \text{s.a} & h(x^{R1}, y) = 0 \end{array} \tag{50}$$

usando $y^{R1}$ como aproximação inicial e as condições (48)-(49) como critério de parada.

Uma das hipóteses para o algoritmo de convergência consiste em obter garantias para que a fase da restauração possa ser completada com sucesso, em todas as iterações. Neste caso, isto significa garantir a existência de $(x^{R1}, y^{R2}, \mu^{R2})$ tais que

$$\|C(x^{R1}, y^{R2}, \mu^{R2})\| \leq r\|C(x^k, y^k, \mu^k)\|$$

e

$$\|(y^{R2}, \mu^{R2})\| \leq \beta\|C(x^k, y^k, \mu^k)\|$$

## 2.8 Observações

Para garantir que o problema em dois níveis (9) tenha solução e provar os resultados de convergência assumimos, ao longo deste capítulo, várias hipóteses sobre o problema. Algumas destas hipóteses são bastante usuais quando tratamos de algoritmos para resolver problemas de programação não linear. Outras, porém, são mais fortes e foram assumidas para tratar especificamente o problema em dois níveis. Nesta seção, estamos interessadas em pesquisar as hipóteses assumidas por outros autores na formulação de algoritmos alternativos para resolverem o problema em questão. Nos trabalhos de Dempe ([1], [2], [3]) encontramos informação bastante ampla sobre este assunto e os comentários que fazemos a seguir foram extraídos destas referências.

A formulação de condições de otimalidade para problemas em dois níveis são, em geral, deduzidas a partir de uma reformulação do problema original como um problema clássico de programação matemática. Neste enfoque, as condições necessárias de primeira ordem se baseiam no conceito de *estabilidade forte* das soluções ótimas do problema do segundo nível. Isto significa exigir a existência de vizinhanças de $x$ e $y$ onde seja possível expressar $y = y(x)$ e tal que a solução seja a única solução ótima local do problema.

Se, além disto, supõe-se que o problema do segundo nível seja convexo para cada $x$ fixo, é possível expressar o problema em dois níveis como um problema que, localmente, é de apenas um nível e obter condições de otimalidade, sempre que seja possível calcular uma derivada direcional. Para isto, assume-se que as restrições do problema verificam uma condição de Mangasarian-Fromowitz e uma condição de posto constante sobre as matrizes das restrições ativas. Estas hipóteses são comparáveis às hipóteses de regularidade pedidas no nosso trabalho. Mas, ainda permitem que a função $y(x)$ não seja contínua.

O outro enfoque, que consiste em substituir o problema do segundo nível pelas condicoes KKT, garante que as condições necessárias podem ser obtidas se o problema do segundo nível é convexo, utilizando a definição de solução ótima chamada otimista, quando o problema do segundo nível admite mais de uma solução. Neste caso, também tem-se o inconveniente da não regularidade, intrínseca destes problemas. Com o intuito de atacar esta situação, surgiram reformulações das condições KKT como sistemas não diferenciáveis.

Para obter condições necessárias de otimalidade, além da convexidade do problema do segundo nível, é preciso pedir condições de regularidade mais fortes que as de Mangasarian-Fromowitz e posto constante, mencionadas anteriormente. A questão de desenvolver algoritmos de descida para os problemas em dois níveis traz a necessidade de caracterizar os pontos estacionários. Para os problemas não diferenciáveis que surgem, utilizam-se os pontos estacionários de Clarke e de Boulligand de modo que, os algoritmos propostos convergem a pontos estacionários nestes sentidos. Essencialmente, é preciso que o problema do segundo nível tenha solução única e que esta solução verifique a condição suficiente de minimizador local, o que é similar ao nosso pedido de hessiano definido positivo.

Não entramos em muito detalhe nesta análise já que o tema é muito complexo e as técnicas utilizadas são completamente diferentes da nossa. A sugestão de fazer esta pesquisa foi dada pela Prof. Susana Scheimberg. A leitura que fizemos do material mencionado nos convenceu de que ainda há muito para ser feito na tentativa de desenvolver algoritmos para resolver problemas em dois níveis.

O enfoque via Restauração Inexata permite evitar muitas situações complicadas. Porém, para aplicarmos esta abordagem, neste trabalho, precisamos impor hipóteses fortes sobre o problema. Vemos, no entanto, que os outros autores também precisam de hipóteses bastante fortes para poderem atacar essas situações complicadas. Nossa intenção, para trabalho futuro, é estudar a possibilidade de relaxar algumas destas nossas exigências e/ou estender os resultados para o caso em que o problema do segundo nível seja um problema variacional.

# Referências

[1] S. Dempe, *Foundations of Bilevel Programming*, Kluwer Academic Publishers, Netherlands, 2002.

[2] S. Dempe, *Annottated Bibligraphy on Bilevel Programming and Mathematical Problems with equilibrium Constraints*, Optimization, 333–359, 2003.

[3] S. Dempe, *Bilevel Programming - A survey*, Technical Report TU 2003-11, Bergakademie Freiberg, 2003.

[4] L.S. Lasdon, *Reduced Gradient Methods*, Nonlinear Optimization 1981, edited by M.J.D.Powell, Academic Press, New York, 235–242, 1982.

[5] J.M. Martínez, *Two-Phase Model Algorithm with Global Convergence for Nonlinear Programming*, JOTA, 96(2), 397–436, 1998.

[6] J.M. Martínez, *Inexact-Restoratoion Method with Lagrangian Tangent Decrease and New Merit Function for Nonlinear Programming*, JOTA, 111(1), 39–58, 2001.

[7] J.M. Martínez e E.A. Pilotta, *Inexact-Restoration Algorithm for Constrained Optimization*, JOTA, 104(1), 135–163, 2000.

[8] J.M. Martínez e E.A. Pilotta, *Inexact Restoration Methods for Nonlinear Programming: Advances and Perspectives*, To appear in Optimization and Control with applications, edited by L.Q.Qi, K.L.Teo and X.Q.Yang. Kluwer Academic Publishers, 2003.

[9] J.M. Martínez e B.F. Svaiter, *A practical Optimality Condition without Constraint Qualifications for Nonlinear Programming*, JOTA, 118, 117–133, 2003.