



**ESTRATÉGIAS PARA RESOLUÇÃO
DO PROBLEMA MPEC**

Flavio Sakakisbara Yano

Dissertação de Mestrado
Pós-Graduação em Matemática Aplicada
MAP - 073

Estratégias para Resolução do Problema MPEC

Flavio Sakakisbara Yano

Dissertação apresentada ao Instituto de Biociências, Letras e Ciências Exatas da Universidade Estadual Paulista "Júlio de Mesquita Filho", Câmpus de São José do Rio Preto, São Paulo, para a obtenção o título de Mestre em Matemática Aplicada.

Orientador: Prof. Dr. Roberto Andreani

São José do Rio Preto,

21 de fevereiro de 2003.

Aos meus pais, Aluísio e Rosa.

À minha irmã, Fernanda.

À minha namorada, Luciani
e às minhas pequenas Rafa e Gabi.

Dedico

Agradecimentos

Primeiramente, aos meus pais, por terem batalhado tanto para me oferecer as melhores oportunidades e as melhores condições para crescer na vida. Hoje, com esta dissertação de mestrado, espero estar lhes dando um pouco alegria, retribuindo o esforço por eles realizado. Aos meus pais, muito obrigado!

Um agradecimento especial ao meu orientador Prof. Dr. Roberto Andreani, pelo companheirismo, pela paciência e atenção dada desde a iniciação científica, mesmo diante das dificuldades que a vida lhe impôs.

Gostaria de agradecer também ao Prof. Dr. Geraldo Nunes Silva, a Profa. Dra. Maria do Socorro Rangel e ao Prof. Dr. Alagacone Sri Ranga, pela ajuda e atenção sem medida.

À Fapesp, pelo auxílio financeiro durante esse período.

À minha irmã, pelo incentivo dado a cada dia.

À minha namorada, pelo entendimento nos dias que me dediquei inteiramente a esse trabalho.

Às minha pequenas Rafa e Gabi, minhas fontes de inspiração.

Aos meus amigos pós-graduandos, pela ajuda.

Enfim, agradeço a todas as pessoas que contribuíram para a realização deste trabalho.

Resumo

Problemas de programação matemática com restrições de equilíbrio (MPEC) são problemas de programação não-linear onde as restrições tem uma estrutura análoga às condições necessárias de primeira ordem de um problema de otimização com restrições. Em formulações usuais do MPEC todos os pontos factíveis são não-regulares no sentido que não satisfazem a *constraint qualification* de Mangassarian-Fromovitz. Portanto, todos os pontos factíveis satisfazem a clássica condição necessária de Fritz-John. Em princípio, isto poderia causar sérias dificuldades ao aplicarmos algoritmos de programação não-linear ao MPEC. Entretanto, muitos pontos factíveis do MPEC não satisfazem uma condição de otimalidade mais forte que Fritz-John, denominada condição AGP. Esta é a razão na qual, em geral, os algoritmos de programação não-linear são satisfatórios quando aplicados ao MPEC. Nosso objetivo neste trabalho é discutir a aplicabilidade dos algoritmos de programação não-linear ao MPEC.

Palavras-chave: programação matemática com restrições de equilíbrio, condições de otimalidade, algoritmos de minimização.

Abstract

Mathematical programming problems with equilibrium constraints (MPEC) are nonlinear programming problems where the constraints have a form that is analogous to first-order optimality conditions of constrained optimization. In usual formulations of MPEC all the feasible points are non-regular in the sense that they do not satisfy the Mangassarian-Fromovitz constraint qualification of nonlinear programming. Therefore, all the feasible points satisfy the classical Fritz-John necessary optimality conditions. In principle, this can cause serious difficulties for nonlinear programming algorithms applied to MPEC. However, the most feasible points of MPEC do not satisfy a stronger optimality condition, called AGP condition. This is the reason why, in general, nonlinear programming algorithms are successful when applied to MPEC. Our objective in this work is to discuss the application of nonlinear programming algorithms to MPEC.

Keywords: mathematical programming with equilibrium constraints, optimality conditions, minimization algorithms.

Sumário

Introdução	8
1 Teoria dos Multiplicadores de Lagrange	11
1.1 Restrições em Formato Geral	11
1.2 Condições Necessárias para Restrições de Igualdade	13
1.3 Função Lagrangeana	16
1.4 Restrições de Desigualdade	18
1.4.1 Condições de Karush-Kuhn-Tucker (KKT)	19
1.4.2 Condições de Fritz-John	21
2 Considerações Teóricas do Problema MPEC	28
2.1 Factibilidade	28
2.2 Otimalidade	31
2.3 Condição AGP	35
2.3.1 Motivação	35
2.3.2 A Condição Necessária	36
2.4 Aplicabilidade da Condição AGP ao MPEC	39
3 Resolução Numérica do Problema MPEC	41
3.1 Algoritmo de Restauração Inexata	41
3.1.1 Idéias Principais do Algoritmo	43
3.1.2 Descrição do Algoritmo	44

3.2 Estratégias	48
3.3 Resultados Numéricos	52
Conclusão	61
Referências Bibliográficas	63

Introdução

Seja $f : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}$ uma função continuamente diferenciável. O problema MPEC (mathematical programming with equilibrium constraints) consiste em minimizar f sujeito a

$$h(x, y) = 0 \quad (1)$$

$$g(x, y) \leq 0 \quad (2)$$

$$F(x, y) + \nabla_y t(x, y)\lambda + \nabla_y s(x, y)z = 0 \quad (3)$$

$$t(x, y) = 0 \quad (4)$$

$$z \geq 0, \quad s(x, y) \leq 0, \quad z_i[s(x, y)]_i = 0, \quad i = 1, \dots, q \quad (5)$$

onde $[s(x, y)]_i$ é a i -ésima componente de $s(x, y)$, $h(x, y) \in \mathbb{R}^{m_1}$, $g(x, y) \in \mathbb{R}^l$, $F(x, y) \in \mathbb{R}^m$, $t(x, y) \in \mathbb{R}^{m_2}$ e $s(x, y) \in \mathbb{R}^q$. Assumiremos que h, g, F são C^1 e que s, t são C^2 .

As restrições (1) e (2) são conhecidas como restrições ordinárias e as restrições (3), (4) e (5) são conhecidas como restrições de equilíbrio.

Se $F(x, y)$ é o gradiente (com respeito a y) de uma função potencial $P(x, y)$ então as restrições de equilíbrio são as condições de KKT do seguinte problema

$$\min_y P(x, y) \text{ s.a. } t(x, y) = 0, \quad s(x, y) \leq 0. \quad (6)$$

Um caso particular do problema MPEC é o problema bilevel, que se caracteriza por ter (6) como restrição ao invés de (3)-(4)-(5).

Se $F(x, y)$ não é um gradiente, as restrições de equilíbrio representam a solução do seguinte problema de inequações variacionais: encontrar $y \in \Omega$ tal que

$$\langle F(x, y), y - z \rangle \geq 0, \quad \forall z \in \Omega$$

onde $\Omega = \{z \in \mathbb{R}^m \mid t(x, z) = 0, s(x, z) \leq 0\}$.

Algumas aplicações do MPEC envolvem, por exemplo, redes de transporte, modelos econômicos e desenho mecânico [9, 10, 14, 22].

O MPEC é um problema de programação não-linear, então, poderíamos tentar resolvê-lo aplicando algoritmos gerais de programação não-linear (PNL). A principal dificuldade associada a esta estratégia se deve às restrições

$$z_i \geq 0, s_i(x, y) \leq 0, s_i(x, y)z_i = 0$$

pois para todo ponto factível devemos ter $s_i(x, y) = 0$ ou $z_i = 0$. Agora

- (a) se $s_i(x, y) = 0$ então o gradiente associado com esta restrição ativa é LD com o gradiente associado à restrição $s_i(x, y)z_i = 0$,
- (b) se $z_i = 0$ então o gradiente associado com esta restrição ativa é LD com o gradiente associado à restrição $s_i(x, y)z_i = 0$.

Logo, o conjunto dos vetores gradientes das restrições ativas são LD para todo ponto factível, ou seja, nenhum ponto factível para o MPEC é regular. Nem mesmo a *constraint qualification* de Mangassarian-Fromovitz é satisfeita.

A convergência de muitos algoritmos gerais de PNL dependem fortemente da hipótese de regularidade, então sua performance poderia ser ruim, e talvez até, inaplicável para o MPEC. Mesmo assim, nosso objetivo é discutir a aplicabilidade dos algoritmos gerais de PNL para estes tipos de problemas.

No Capítulo 1, desta dissertação, introduzimos uma teoria necessária para um bom entendimento do assunto em questão. Esta teoria se refere, basicamente, às condições de otimalidade que envolvem multiplicadores de Lagrange tais como KKT, Fritz-John e a *constraint qualification* de Mangassarian Fromovitz.

No Capítulo 2, exploraremos algumas características do problema MPEC tais como factibilidade e otimalidade. Além disso, este capítulo se destina a mostrar a base teórica que nos dá suporte para acreditar que podemos obter bons resultados ao aplicarmos algoritmos de PNL ao problema MPEC. Esta teoria está baseada

em uma condição prática de otimalidade, denominada AGP, que é o assunto principal de algumas seções deste capítulo.

No Capítulo 3, apresentaremos as estratégias e os algoritmos (métodos) de PNL utilizados dando uma atenção especial ao algoritmo de restauração inexata, já que, a implementação deste algoritmo também fez parte do nosso trabalho. Assim, apresentaremos os resultados numéricos obtidos ao tentarmos resolver alguns problemas MPEC da literatura para, a partir daí, concluirmos e finalizarmos este trabalho.

Capítulo 1

Teoria dos Multiplicadores de Lagrange

O conjunto factível de um problema de otimização, geralmente, é especificado em termos de restrições de igualdade e desigualdade. Levando-se em conta esta estrutura obtemos uma sofisticada coleção de condições de otimalidade envolvendo algumas variáveis auxiliares, denominadas multiplicadores de Lagrange. Estas, facilitam a caracterização da solução ótima. Neste capítulo, iremos apresentar as condições de otimalidade que envolvem multiplicadores de Lagrange. Um estudo detalhado deste assunto pode ser encontrado em [6, 12, 20].

1.1 Restrições em Formato Geral

Nesta seção, consideraremos o seguinte problema:

$$\begin{array}{ll} \min & f(x) \\ \text{s.a.} & x \in \Omega. \end{array} \tag{1.1}$$

As curvas no conjunto Ω desempenham um papel importante na obtenção de condições de otimalidade, assim, as condições de otimalidade que vamos apresentar estão baseadas apenas no comportamento da função objetivo em cima de

curvas factíveis que passam pelo ponto considerado. Um estudo detalhado deste assunto pode ser encontrado em [20].

Definição 1.1.1. *Dado $x^* \in \Omega$ chamamos de curva em Ω de classe C^k partindo de x^* a função contínua $\gamma : [0, \varepsilon] \rightarrow \Omega$ tal que $\varepsilon > 0$, $\gamma(0) = x^*$ e $\gamma \in C^k[0, \varepsilon]$.*

Teorema 1.1.2 (Condição Necessária de Primeira Ordem baseada em Curvas). *Seja x^* um minimizador local de (1.1) e γ uma curva em Ω de classe C^1 partindo de x^* . Então $\nabla f(x^*)^T \gamma'(0) \geq 0$.*

Prova: Defina $\varphi : [0, \varepsilon] \rightarrow \mathbb{R}$ por $\varphi(t) = f(\gamma(t))$.

Como x^* é minimizador local, existe $\varepsilon_1 \in (0, \varepsilon)$ tal que $\varphi(t) \geq \varphi(0) = f(x^*)$ para todo $t \in (0, \varepsilon_1)$. Assim,

$$\frac{\varphi(t) - \varphi(0)}{t} \geq 0, \quad \forall t \in (0, \varepsilon_1).$$

Desta forma,

$$\varphi'(0) = \lim_{t \rightarrow 0} \frac{\varphi(t) - \varphi(0)}{t} \geq 0.$$

Mas, pela regra da cadeia,

$$\varphi'(t) = f'(\gamma(t))\gamma'(t) = \nabla f(\gamma(t))^T \gamma'(t).$$

Em particular, para $t = 0$ temos $\varphi'(0) = \nabla f(\gamma(0))^T \gamma'(0) = \nabla f(x^*)^T \gamma'(0) \geq 0$. ■

Definição 1.1.3. *Dado $x \in \Omega$, dizemos que γ é uma curva em Ω de classe C^k passando por x se $\gamma : [-\varepsilon, \varepsilon] \rightarrow \Omega$, $\varepsilon > 0$, $\gamma(0) = x$ e $\gamma \in C^k$.*

Teorema 1.1.4. *Se $x^* \in \Omega$ é um minimizador local de (1.1) e γ é uma curva em Ω de classe C^1 passando por x^* então $\nabla f(x^*)^T \gamma'(0) = 0$*

Prova: Defina $\gamma_1 : [0, \varepsilon] \rightarrow \Omega$ por $\gamma_1(t) = \gamma(t)$ e $\gamma_2 : [0, \varepsilon] \rightarrow \Omega$ por $\gamma_2(t) = \gamma(-t)$.

Pelo Teorema 1.1.2, temos que

$$\nabla f(x^*)^T \gamma'_1(0) \geq 0 \quad \text{e} \quad \nabla f(x^*)^T \gamma'_2(0) \geq 0.$$

Mas, $\gamma'_1(0) = \gamma'(0)$ e $\gamma'_2(0) = -\gamma'(0)$ e, portanto, $\nabla f(x^*)^T \gamma'(0) = 0$. ■

1.2 Condições Necessárias para Restrições de Igualdade

Nesta seção, consideraremos problemas com restrições de igualdade da forma:

$$\begin{aligned} \min \quad & f(x) \\ \text{s.a.} \quad & h(x) = 0 \end{aligned} \quad (1.2)$$

onde $f : \mathbb{R}^n \rightarrow \mathbb{R}$ é continuamente diferenciável e $h : \mathbb{R}^n \rightarrow \mathbb{R}^m$ é definida por $h = (h_1, h_2, \dots, h_m)^T$ onde $h_i : \mathbb{R}^n \rightarrow \mathbb{R}$, $i = 1, \dots, m$, também são continuamente diferenciáveis. Neste caso, temos $\Omega = \{x \in \mathbb{R}^n \mid h(x) = 0\}$.

Definição 1.2.1. Se $x \in \Omega$, chamamos de conjunto tangente a Ω por x , denotado por $M(x)$, o conjunto dos vetores tangentes a curvas em Ω passando por x , ou seja,

$$M(x) = \{v \in \mathbb{R}^n \mid v = \gamma'(0) \text{ para alguma curva } \gamma \text{ passando por } x\}.$$

Lema 1.2.2. Se x^* é um minimizador local de (1.2) então $\nabla f(x^*) \perp M(x^*)$, ou seja, para todo elemento $v \in M(x^*)$, temos que $\nabla f(x^*)^T v = 0$.

Prova: Seja $v \in M(x^*)$. Então existe uma curva γ passando por x^* tal que $\gamma'(0) = v$.

Do Teorema 1.1.4 segue que $\nabla f(x^*)^T v = 0$.

Como v é arbitrário concluímos que $\nabla f(x^*) \perp M(x^*)$. ■

O teorema dos Multiplicadores de Lagrange diz, basicamente, que para um dado minimizador local x^* existem escalares $\lambda_1, \dots, \lambda_m$, denominados multiplicadores de Lagrange, tais que

$$\nabla f(x^*) + \sum_{i=1}^m \lambda_i \nabla h_i(x^*) = 0. \quad (1.3)$$

Utilizando a notação

$$h'(x) = \begin{pmatrix} \frac{\partial h_1}{\partial x_1} & \cdots & \frac{\partial h_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial h_m}{\partial x_1} & \cdots & \frac{\partial h_m}{\partial x_n} \end{pmatrix} = \begin{pmatrix} h'_1(x) \\ \vdots \\ h'_m(x) \end{pmatrix} = \begin{pmatrix} \nabla h_1(x)^T \\ \vdots \\ \nabla h_m(x)^T \end{pmatrix}$$

podemos interpretar esta condição da seguinte forma:

(a) O vetor gradiente $\nabla f(x^*)$ pertence ao espaço gerado pelos vetores gradientes das restrições em x^* , isto é, $\nabla f(x^*) \in \text{Im}(h'(x^*)^T)$.

(b) O vetor gradiente $\nabla f(x^*)$ é ortogonal ao $\text{Nu}(h'(x^*))$ [este resultado é consequência direta do item (a), já que, $\text{Nu}(h'(x^*)) \perp \text{Im}(h'(x^*)^T)$] onde

$$\text{Nu}(h'(x^*)) = \{v \in \mathbb{R}^n \mid \nabla h_i(x^*)^T v = 0, i = 1, \dots, m\}.$$

Observando o resultado enunciado no Lema 1.2.2 e a interpretação para a equação dos multiplicadores de Lagrange, dada no item (b), é natural que indagemos sobre qual é a relação entre $\text{Nu}(h'(x^*))$ e $M(x^*)$.

Lema 1.2.3. *Para todo $x \in \Omega$ temos que $M(x) \subset \text{Nu}(h'(x))$.*

Prova: Seja $v \in M(x)$ e $\gamma : [-\varepsilon, \varepsilon] \rightarrow \Omega$ tal que $\gamma'(0) = v$, $\gamma(0) = x$.

Defina $\phi(t) = h(\gamma(t))$, para todo $t \in [-\varepsilon, \varepsilon]$.

Desta forma, temos que $\phi(t) = 0$ para todo $t \in [-\varepsilon, \varepsilon]$.

Logo, $\phi'(t) = 0$ para todo $t \in (-\varepsilon, \varepsilon)$. Mas, pela regra da cadeia, temos que

$$\phi'(t) = h'(\gamma(t))\gamma'(t) = 0, \forall t \in (-\varepsilon, \varepsilon).$$

Em particular, para $t = 0$ temos que $0 = h'(x)\gamma'(0) = h'(x)v$, ou seja, $v \in \text{Nu}(h'(x))$. ■

Será que também vale a recíproca inclusão, ou seja, $\text{Nu}(h'(x)) \subset M(x)$? Em geral, esta relação não é verdadeira, conforme ilustra o seguinte exemplo: consideremos $h(x_1, x_2) = x_1 x_2$ e $x = (0, 0)^T$, então $M(x) = \{v \in \mathbb{R}^2 \mid v_1 v_2 = 0\}$, mas $h'(x) = (0, 0)$ e claramente $\text{Nu}(h'(x)) = \mathbb{R}^2$.

Definição 1.2.4. Dizemos que $x \in \Omega = \{x \in \mathbb{R}^n \mid h(x) = 0\}$ é um ponto regular se $\{\nabla h_1(x), \dots, \nabla h_m(x)\}$ formam um conjunto linearmente independente, ou equivalentemente, se o posto de $h'(x)$ é igual a m .

Agora, com a definição de regularidade, podemos enunciar o seguinte teorema cuja demonstração pode ser encontrada em [12, 20].

Teorema 1.2.5. Seja $\Omega = \{x \in \mathbb{R}^n \mid h(x) = 0\}$, $h \in C^k$, $x \in \Omega$ um ponto regular. Então, para todo $v \in Nu(h'(x))$ existe uma curva γ de classe C^k passando por x tal que $\gamma'(0) = v$, ou seja, $Nu(h'(x)) \subset M(x)$. Portanto, $M(x) = Nu(h'(x))$.

Desta forma, se x^* é minimizador regular de (1.2) então a equação (1.3) dos multiplicadores de Lagrange é consequência direta do Lema 1.2.2 e do Teorema 1.2.5. Agora, enunciaremos formalmente o teorema dos Multiplicadores de Lagrange. Várias formas de demonstrar este teorema podem ser encontradas em [6, 12, 20].

Teorema 1.2.6 (Multiplicadores de Lagrange - Condições Necessárias).

Seja x^* um minimizador local regular de f sujeito a $h(x) = 0$. Então existe um único vetor $\lambda^* = (\lambda_1^*, \dots, \lambda_m^*)^T$, chamado vetor dos multiplicadores de Lagrange, tal que

$$\nabla f(x^*) + \sum_{i=1}^m \lambda_i \nabla h_i(x^*) = 0 \quad (1.4)$$

Se, além disso, $f, h_i \in C^2$, temos que

$$y^T \left(\nabla^2 f(x^*) + \sum_{i=1}^m \lambda_i^* \nabla^2 h_i(x^*) \right) y \geq 0, \quad \forall y \in Nu(h'(x^*)) \quad (1.5)$$

A seguir, vamos ilustrar que pode ocorrer o fato de não existirem os multiplicadores de Lagrange para um minimizador local que não seja regular.

Exemplo 1.2.1. Considere o problema de minimizar

$$f(x) = x_1 + x_2$$

sujeito a estas duas restrições

$$h_1(x) = (x_1 - 1)^2 + x_2^2 - 1 = 0, \quad h_2(x) = (x_1 - 2)^2 + x_2^2 - 4 = 0.$$

A geometria deste problema está dada na figura a seguir. Daí, podemos notar que no ponto de mínimo local $x^* = (0, 0)$ que é o único ponto factível do problema, o vetor gradiente $\nabla f(x^*) = (1, 1)$ não pode ser escrito como combinação linear dos gradientes das restrições $\nabla h_1(x^*) = (-2, 0)$ e $\nabla h_2(x^*) = (-4, 0)$.

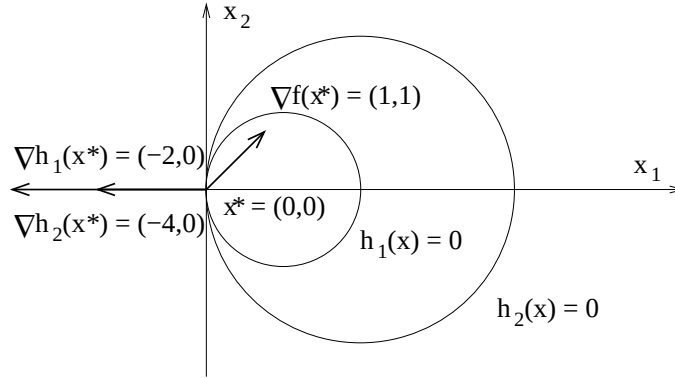


Figura 1.1: Geometria do problema

Logo, a condição dos multiplicadores de Lagrange

$$\nabla f(x^*) + \lambda_1^* \nabla h_1(x^*) + \lambda_2^* \nabla h_2(x^*) = 0$$

não pode ser satisfeita quaisquer que sejam λ_1^* e λ_2^* . A dificuldade deste problema é que o núcleo do jacobiano $Nu(h'(x^*)) = \{y \mid y_1 = 0, y_2 \in \mathbb{R}\}$ tem dimensão maior que o plano tangente $M(x^*) = \{y \mid y = 0\}$. A otimalidade de x^* implica que $\nabla f(x^*) \perp M(x^*)$, mas para que os multiplicadores existam é necessário que $\nabla f(x^*) \perp Nu(h'(x^*))$.

1.3 Função Lagrangeana

Muitas vezes, é mais conveniente escrever as condições necessárias em termos da função Lagrangeana $L : \mathbb{R}^{n+m} \rightarrow \mathbb{R}$, definida por:

$$L(x, \lambda) = f(x) + \sum_{i=1}^m \lambda_i h_i(x).$$

Desta forma, derivando a função Lagrangeana, obtemos as seguintes relações:

$$\begin{aligned}\nabla_x L(x^*, \lambda^*) &= \nabla f(x^*) + \sum_{i=1}^m \lambda_i^* \nabla h_i(x^*), \\ \nabla_\lambda L(x^*, \lambda^*) &= h(x^*), \\ \nabla_{xx}^2 L(x^*, \lambda^*) &= \nabla^2 f(x^*) + \sum_{i=1}^m \lambda_i^* \nabla^2 h_i(x^*).\end{aligned}$$

Então, se x^* é um minimizador local regular, as condições dos multiplicadores de Lagrange do Teorema 1.2.6 juntamente com a restrição $h(x^*) = 0$ são escritas, compactamente, da seguinte forma:

$$\nabla_x L(x^*, \lambda^*) = 0 \quad \nabla_\lambda L(x^*, \lambda^*) = 0, \quad (1.6)$$

$$y^T \nabla_{xx}^2 L(x^*, \lambda^*) y \geq 0, \quad \forall y \in Nu(h'(x^*)). \quad (1.7)$$

As condições necessárias de primeira ordem (1.6) representam um sistema não-linear com $n + m$ equações e $n + m$ incógnitas - as coordenadas de x^* e λ^* . Todo minimizador local regular x^* , juntamente com o vetor dos multiplicadores de Lagrange associado, representarão uma solução deste sistema. Entretanto, uma solução do sistema não corresponde necessariamente a um minimizador local; esta solução pode ser, por exemplo, um maximizador local.

Assim, os pontos regulares que sejam soluções do sistema não linear

$$\begin{cases} \nabla_x L(x, \lambda) = 0 \\ \nabla_\lambda L(x, \lambda) = 0 \end{cases}$$

são apenas candidatos a minimizador local. Esses pontos serão chamados estacionários ou críticos. Naturalmente, os pontos não-regulares de Ω também serão candidatos a minimizador local.

Exemplo 1.3.1. *Considere o problema*

$$\min \frac{1}{2}(x_1^2 + x_2^2 + x_3^2)$$

$$\text{s. a. } x_1 + x_2 + x_3 = 3.$$

As condições necessárias de primeira ordem (1.6) fornecem

$$\begin{aligned}x_1^* + \lambda^* &= 0, \\x_2^* + \lambda^* &= 0, \\x_3^* + \lambda^* &= 0, \\x_1^* + x_2^* + x_3^* &= 3.\end{aligned}$$

Este representa um sistema com quatro equações e quatro incógnitas que tem uma única solução

$$x_1^* = x_2^* = x_3^* = 1, \quad \lambda^* = -1.$$

O gradiente da restrição é $(1, 1, 1)$, então todos os pontos factíveis são regulares. Portanto, $x^* = (1, 1, 1)$ é o único candidato a minimizador local. Além disso, como para este problema $\nabla_{xx}^2 L(x^*, \lambda^*)$ é a matriz identidade, a condição necessária de segunda ordem (1.7) também é satisfeita. Se quisermos maximizar o problema anterior, isto é,

$$\begin{aligned}\min \quad & -\frac{1}{2}(x_1^2 + x_2^2 + x_3^2) \\ \text{s.a.} \quad & x_1 + x_2 + x_3 = 3.\end{aligned}$$

então as condições necessárias de primeira ordem (1.6) fornecem $x^* = (1, 1, 1)$ e $\lambda^* = 1$. Entretanto, a condição necessária de segunda ordem (1.7) não é satisfeita e como todos os pontos factíveis são regulares podemos concluir que este problema não tem solução.

1.4 Restrições de Desigualdade

Agora, consideraremos problemas envolvendo restrições de igualdade e desigualdade definidos da seguinte forma:

$$\begin{aligned}\min \quad & f(x) \\ \text{s.a.} \quad & h(x) = 0, \quad g(x) \leq 0\end{aligned}\tag{1.8}$$

onde $f : \mathbb{R}^n \rightarrow \mathbb{R}$ é continuamente diferenciável, $h : \mathbb{R}^n \rightarrow \mathbb{R}^m$ é definida por $h = (h_1, h_2, \dots, h_m)^T$ e $g : \mathbb{R}^n \rightarrow \mathbb{R}^r$ é definida por $g = (g_1, g_2, \dots, g_r)^T$ onde $h_i, g_j : \mathbb{R}^n \rightarrow \mathbb{R}$ também são continuamente diferenciáveis. Neste caso, temos $\Omega = \{x \in \mathbb{R}^n \mid h(x) = 0 \text{ e } g(x) \leq 0\}$.

Nossa primeira linha de pensamento está baseada em uma versão estendida da noção de regularidade que utilizamos para os problemas com restrições de igualdade. Isto nos conduzirá diretamente ao desenvolvimento da teoria que é, entretanto, limitada pelas hipóteses de regularidade. Tais hipóteses são muito comuns em problemas com restrições não-lineares de igualdade mas, são um tanto restritivas em problemas com restrições de desigualdade. A razão é que em muitos tipos de problema pode haver muitas restrições de desigualdade que são satisfeitas por igualdade no ponto de mínimo local cujos vetores gradientes associados são dependentes. Com esta motivação, desenvolveremos mais tarde, uma teoria dos multiplicadores de Lagrange que não está baseada na hipótese de regularidade.

1.4.1 Condições de Karush-Kuhn-Tucker (KKT)

Para todo ponto factível $x \in \Omega$, o conjunto dos índices das restrições (de desigualdade) ativas é definido e será denotado por:

$$I(x) = \{j \in \{1, \dots, r\} \mid g_j(x) = 0\}.$$

Se $j \notin I(x)$ diremos que a j -ésima restrição de desigualdade é inativa em x .

Definição 1.4.1. Dizemos que $x \in \Omega$ é regular se os vetores gradientes das restrições de igualdade $\nabla h_i(x)$, $i = 1, \dots, m$ e os vetores gradientes das restrições ativas $\nabla g_j(x)$, $j \in I(x)$ são linearmente independentes.

Observação: Dizemos também, excepcionalmente, que x é regular quando não temos restrições de igualdade e todas as restrições de desigualdade são inativas.

O teorema a seguir é uma generalização do Teorema dos Multiplicadores de Lagrange e será dada em termos da função lagrangeana definida por:

$$L(x, \lambda, \mu) = f(x) + \sum_{i=1}^m \lambda_i h_i(x) + \sum_{j=1}^r \mu_j g_j(x).$$

Teorema 1.4.2 (Condições Necessárias de KKT). *Seja x^* um minimizador local regular do problema (1.8). Então existem e são únicos os vetores dos multiplicadores de Lagrange $\lambda^* = (\lambda_1, \dots, \lambda_m)^T$ e $\mu^* = (\mu_1, \dots, \mu_r)^T$ tais que*

$$\nabla_x L(x^*, \lambda^*, \mu^*) = \nabla f(x^*) + \sum_{i=1}^m \lambda_i^* \nabla h_i(x^*) + \sum_{j=1}^r \mu_j^* \nabla g_j(x^*) = 0$$

$$\mu_j \geq 0, \quad j = 1, \dots, r$$

$$\mu_j = 0, \quad j \notin I(x^*)$$

Se além disso, $f, h_i, g_j \in C^2$ então temos que

$$y^T \nabla_{xx}^2 L(x^*, \lambda^*, \mu^*) y \geq 0,$$

para todo $y \in \mathbb{R}^n$ tal que $\nabla^T h_i(x^*) y = 0$, $i = 1, \dots, m$ e $\nabla^T g_j(x^*) y = 0$, $\forall j \in I(x^*)$.

Observações:

1. No ponto de mínimo local, as restrições inativas não causam problemas, elas podem ser ignoradas ao expressarmos as condições de otimalidade. Por outro lado, as restrições ativas podem ser tratadas como uma extensão das restrições de igualdade.

2. A condição $\mu_j^* = 0$ para todo $j \notin I(x^*)$ pode ser escrita como

$$\mu_j^* g_j(x^*) = 0, \quad j = 1, \dots, r$$

e é chamada de condição de folga complementar. Este nome se deve ao fato de que para cada j , sempre que a restrição $g_j(x) \leq 0$ seja inativa em x^* , ou seja, $g_j(x^*) < 0$ então a restrição $\mu_j^* \geq 0$ deve ser satisfeita por igualdade ($\mu_j^* = 0$) e vice-versa. Uma versão mais forte da condição de folga complementar será dada nas condições de Fritz-John.

3. Pela observação feita no item anterior, se x é regular então podemos escrever as condições de KKT da seguinte forma:

$$\nabla f(x) + \sum_{i=1}^m \lambda_i \nabla h_i(x) + \sum_{j=1}^r \mu_j \nabla g_j(x) = 0 \quad (1.9)$$

$$h(x) = 0 \quad (1.10)$$

$$\mu_j g_j(x) = 0, \quad j = 1, \dots, r \quad (1.11)$$

$$\mu_j \geq 0, \quad j = 1, \dots, r \quad (1.12)$$

$$g_j(x) \leq 0, \quad j = 1, \dots, r \quad (1.13)$$

As $n + m + r$ equações (1.9)-(1.11) formam um sistema não-linear nas incógnitas $x \in \mathbb{R}^n$, $\lambda \in \mathbb{R}^m$, $\mu \in \mathbb{R}^r$. As soluções deste sistema que satisfazem (1.12) e (1.13) são os pontos estacionários do problema. Naturalmente, os pontos não-regulares factíveis também são candidatos a minimizador local.

4. De (1.12) e (1.13) segue que o produto $\mu_j g_j(x)$ é não-positivo para todo j . Desta forma, as r equações definidas em (1.11) podem ser escritas compactamente como $\mu^T g(x) = 0$.

1.4.2 Condições de Fritz-John

Desenvolveremos, agora, uma generalização das condições necessárias de KKT que não necessita da hipótese de regularidade. Com a ajuda desta generalização iremos deduzir as condições de otimalidade já obtidas e mais algumas novas condições necessárias. Para ter uma idéia do que representam as condições de Fritz-John, considere o problema com restrições de igualdade

$$\begin{aligned} \min \quad & f(x) \\ \text{s.a.} \quad & h(x) = 0 \end{aligned}$$

e, note que, um resultado equivalente ao dado no teorema dos Multiplicadores de Lagrange é que para um minimizador local x^* (regular ou não) existem escalares

$\mu_0^*, \lambda_1^*, \dots, \lambda_m^*$ nem todos nulos tais que $\mu_0^* \geq 0$ e

$$\mu_0^* \nabla f(x^*) + \sum_{i=1}^m \lambda_i^* \nabla h_i(x^*) = 0. \quad (1.14)$$

Para ver a equivalência, note que há duas possibilidades:

(a) x^* é regular e, neste caso, a equação (1.14) é satisfeita com $\mu_0^* = 1$.

(b) x^* não é regular e, neste caso, os vetores gradientes $\nabla h_i(x^*)$ são linearmente dependentes, ou seja, existem $\lambda_1^*, \dots, \lambda_m^*$ nem todos nulos tais que

$$\sum_{i=1}^m \lambda_i^* \nabla h_i(x^*) = 0,$$

e portanto, a equação (1.14) é satisfeita com $\mu_0^* = 0$.

Condições necessárias que envolvem um multiplicador não-negativo para o vetor gradiente da função objetivo, tal como a representada na equação (1.14), são conhecidas como condições necessárias de Fritz-John. No teorema a seguir, enunciaremos formalmente as condições de otimalidade de Fritz-John que também vão fornecer uma característica mais precisa das restrições correspondentes aos multiplicadores não-nulos.

Teorema 1.4.3 (Condições Necessárias de Fritz-John). *Seja x^* um minimizador local do problema*

$$\begin{aligned} \min \quad & f(x) \\ \text{s.a.} \quad & h(x) = 0, \quad g(x) \leq 0 \end{aligned}$$

onde f, h_i, g_j são funções continuamente diferenciáveis de $\mathbb{R}^n$ em $\mathbb{R}$. Então existe um escalar μ_0^* e multiplicadores $\lambda_1^*, \dots, \lambda_m^*, \mu_1^*, \dots, \mu_r^*$ satisfazendo as seguintes condições:

(i)

$$\mu_0^* \nabla f(x^*) + \sum_{i=1}^m \lambda_i^* \nabla h_i(x^*) + \sum_{j=1}^r \mu_j^* \nabla g_j(x^*) = 0.$$

(ii)

$$\mu_j^* \geq 0, \quad j = 0, 1, \dots, r.$$

(iii) $\mu_0^*, \lambda_1^*, \dots, \lambda_m^*, \mu_1^*, \dots, \mu_r^*$ não são todos nulos.

(iv) Em qualquer vizinhança N de x^* existe um $x \in N$ tal que $\lambda_i^* h_i(x) > 0$ para todo i com $\lambda_i^* \neq 0$ e $\mu_j^* g_j(x) > 0$ para todo j com $\mu_j^* \neq 0$.

Exemplo 1.4.1. Considere o problema

$$\begin{aligned} \min \quad & f(x) = x_1 + x_2 \\ \text{s.a.} \quad & g_1(x) = (x_1 - 1)^2 + x_2^2 - 1 \leq 0, \\ & g_2(x) = (x_1 - 2)^2 + x_2^2 - 4 \leq 0. \end{aligned}$$

Esta é uma versão com restrições de desigualdade do Exemplo 1.2.1, cuja geometria está ilustrada na Figura 1.1. Podemos observar desta figura que o único ponto factível é $x^* = (0, 0)$, que é portanto o minimizador global. A condição

$$\mu_0^* \nabla f(x^*) + \mu_1^* \nabla g_1(x^*) + \mu_2^* \nabla g_2(x^*) = 0$$

pode ser escrita como

$$\mu_0^* - 2\mu_1^* + 4\mu_2^* = 0, \quad \mu_0^* = 0.$$

Logo, coerentemente com o Exemplo 1.2.1, onde concluímos que não existem os multiplicadores de Lagrange, temos que $\mu_0^* = 0$. Desta forma, podemos observar que todos positivos μ_1^* e μ_2^* tais que $\mu_1^* = 2\mu_2^*$ satisfazem as condições necessárias de Fritz-John (i)-(iii), e pela geometria do problema ilustrada na Figura 1.1, podemos verificar que a condição (iv) também é satisfeita.

A condição (iv) implica na condição de folga complementar ($\mu_j^* g_j(x^*) = 0, \forall j$) pois se $\mu_j^* > 0$ a j -ésima restrição de desigualdade deve ser violada ($g_j(x) > 0$) arbitrariamente próximo de x^* implicando que $g_j(x^*) = 0$ e, se $g_j(x^*) < 0$ devemos ter $\mu_j^* = 0$ pois, caso contrário, existiria uma vizinhança N de x^* que $\mu_j^* g_j(x) < 0$ para todo $x \in N$.

Esta condição substitui a condição de folga complementar (mais fraca)

$$\mu_j^* = 0, \forall j \notin I(x^*)$$

do Teorema 1.4.2 e restringe o conjunto dos candidatos a multiplicadores que satisfazem as condições de Fritz-John. Além disso, a condição (iv) é algo um tanto incomum na literatura e pode ser escrita equivalentemente da seguinte forma: "em toda vizinhança N de x^* existe $x \in N$ que viola simultaneamente todas as restrições correspondentes aos multiplicadores de Lagrange não-nulos."

De Fritz-John para as Condições de KKT

Para ilustrar a utilidade das condições de Fritz-John e da condição (iv), em particular, suponha que convertamos um problema com uma única restrição de igualdade $\min_{h(x)=0} f(x)$ em um problema com restrições de desigualdade

$$\begin{aligned} \min \quad & f(x) \\ \text{s.a.} \quad & h(x) \leq 0, \quad -h(x) \leq 0. \end{aligned}$$

As condições de Fritz-John asseguram a existência de escalares não-negativos $\mu_0^*, \lambda^+, \lambda^-$ nem todos iguais a zeros tais que

$$\mu_0^* \nabla f(x^*) + \lambda^+ \nabla h(x^*) - \lambda^- \nabla h(x^*) = 0. \quad (1.15)$$

Dentre os candidatos a multiplicadores que satisfazem a equação (1.15), além da condição de folga complementar

$$\lambda^+ h(x^*) = \lambda^- h(x^*) = 0,$$

estão incluídos aqueles da forma $\mu_0^* = 0, \lambda^+ = \lambda^- > 0$. Entretanto, estes multiplicadores não satisfazem a condição (iv) do Teorema 1.4.3, mostrando que se $\mu_0^* = 0$ então devemos ter $\lambda^+ \neq 0$ e $\lambda^- = 0$ ou $\lambda^+ = 0$ e $\lambda^- \neq 0$. Para um x^* regular, estes multiplicadores violam a equação (1.15) e daí segue que $\mu_0^* > 0$.

Desta forma, dividindo a equação (1.15) por μ_0^* obtemos, com a hipótese de regularidade de x^* , a já conhecida condição necessária de primeira ordem

$\nabla f(x^*) + \lambda^* \nabla h(x^*) = 0$ com $\lambda^* = (\lambda^+ - \lambda^-)/\mu_0^*$. Note que não seria possível fazer esta dedução sem a condição (iv).

Enunciaremos agora alguns teoremas cuja demonstração segue, como na discussão acima, o seguinte procedimento: verificar que não é possível satisfazer as condições de Fritz-John (i)-(iv) com $\mu_0^* = 0$ para daí concluir que $\mu_0^* > 0$, assim, dividindo todos λ_i^* e μ_j^* por μ_0^* vamos obter as condições de KKT com multiplicadores de Lagrange iguais a λ_i^*/μ_0^* e μ_j^*/μ_0^* . Na verdade, como os multiplicadores podem ser normalizados como uma multiplicação por escalar positivo, é suficiente considerar apenas duas possibilidades para μ_0^* : $\mu_0^* = 0$ ou $\mu_0^* = 1$.

O teorema que enunciaremos a seguir representa uma versão um pouco mais forte das condições de KKT, substituindo a condição de folga complementar $\mu_j^* = 0$ se $j \notin I(x^*)$ pela condição (iv) do Teorema 1.4.3. A demonstração deste teorema pode ser encontrada em [6].

Teorema 1.4.4. *Seja x^* um minimizador local regular do problema*

$$\begin{aligned} \min \quad & f(x) \\ \text{s.a.} \quad & h(x) = 0, \quad g(x) \leq 0 \end{aligned}$$

onde f, h_i, g_j são funções continuamente diferenciáveis de $\mathbb{R}^n$ em $\mathbb{R}$. Então existem multiplicadores de Lagrange $\lambda_1^*, \dots, \lambda_m^*$ e $\mu_1^*, \dots, \mu_r^*$ satisfazendo as seguintes condições:

(i)

$$\nabla f(x^*) + \sum_{i=1}^m \lambda_i^* \nabla h_i(x^*) + \sum_{j=1}^r \mu_j^* \nabla g_j(x^*) = 0.$$

(ii)

$$\mu_j^* \geq 0, \quad j = 1, \dots, r.$$

(iii) *Em qualquer vizinhança N de x^* existe um $x \in N$ tal que $\lambda_i^* h_i(x) > 0$ para todo i com $\lambda_i \neq 0$ e $\mu_j^* g_j(x) > 0$ para todo j com $\mu_j^* \neq 0$.*

Agora, apresentaremos algumas condições que garantem a existência dos multiplicadores de Lagrange. Tais condições são conhecidas como *constraint qualifications*. O teorema a seguir foi demonstrado por Arrow, Hurwitz e Uzawa para o caso onde temos somente restrições de desigualdade e foi generalizado, para o caso onde também temos restrições de igualdade, por Mangassarian e Fromovitz em [13]. A demonstração deste teorema também pode ser encontrada em [6].

Teorema 1.4.5 (Constraint Qualification de Mangassarian-Fromovitz).

Seja x^ um minimizador local do problema*

$$\begin{aligned} \min \quad & f(x) \\ \text{s.a.} \quad & h(x) = 0, \quad g(x) \leq 0 \end{aligned}$$

onde f, h_i, g_j são funções continuamente diferenciáveis de $\mathbb{R}^n$ em $\mathbb{R}$. Suponha que os vetores gradientes $\nabla h_i(x^)$, $i = 1, \dots, m$ são linearmente independentes e que existe um vetor d tal que*

$$\nabla h_i(x^*)^T d = 0, \quad \forall i = 1, \dots, m, \quad \nabla g_j(x^*)^T d < 0, \quad \forall j \in I(x^*).$$

Então x^ satisfaz as condições necessárias do Teorema 1.4.4.*

Note que a existência do vetor d satisfazendo as condições de teorema anterior é equivalente a hipótese (mais fácil de ser verificada) que diz que para cada $j \in I(x^*)$ existe um vetor d^j tal que

$$\nabla h_i(x^*)^T d^j = 0, \quad i = 1, \dots, m,$$

$$\nabla g_j(x^*)^T d^j < 0, \quad \nabla g_{\bar{j}}(x^*)^T d^j \leq 0, \quad \forall \bar{j} \in I(x^*), \quad \bar{j} \neq j.$$

Desta forma, o vetor $d = \sum_{j \in I(x^*)} d^j$ satisfaz as condições do teorema. O próximo teorema, devido a Slater, é um caso especial do teorema anterior. A demonstração deste teorema também pode ser encontrada em [6].

Teorema 1.4.6 (Constraint Qualification de Slater para Desigualdades Convexas). *Seja x^* um minimizador local do problema*

$$\begin{aligned} \min \quad & f(x) \\ \text{s.a.} \quad & h(x) = 0, \quad g(x) \leq 0 \end{aligned}$$

onde f e g_j são funções continuamente diferenciáveis de $\mathbb{R}^n$ em $\mathbb{R}$ e as funções h_i são lineares. Suponha que as funções g_j são convexas e que existe um vetor factível $\bar{x}$ satisfazendo

$$g_j(\bar{x}) < 0, \quad \forall j \in I(x^*).$$

Então x^ satisfaz as condições necessárias do Teorema 1.4.4.*

Capítulo 2

Considerações Teóricas do Problema MPEC

O objetivo principal deste capítulo é solidificar a base teórica das estratégias que vamos propor para resolver o problema MPEC. Para isto, vamos explorar alguns aspectos que envolvem o problema tais como: factibilidade e otimalidade.

2.1 Factibilidade

Muitos algoritmos de PNL tem a propriedade de que seus pontos de acumulação são pontos (KKT) estacionários da norma das restrições ao quadrado. Obviamente, seria muito interessante garantir a factibilidade dos pontos de acumulação, mas isto é impossível quando não temos a hipótese de convexidade. Num caso extremo, o conjunto factível poderia ser vazio. Mesmo neste caso, seria interessante entender como os algoritmos de PNL se comportam. Sem perda de generalidade, suponha que as restrições dos problemas de PNL são dadas por:

$$H(u) = 0, \quad u \geq 0. \quad (2.1)$$

(Restrições arbitrárias podem ser escritas desta forma através da introdução de variáveis de folga apropriadas). Então um ponto estacionário da norma das res-

trições ao quadrado é um ponto que satisfaz as condições necessárias de primeira ordem para o problema

$$\min \|H(u)\|_2^2 \quad \text{s.a.} \quad u \geq 0. \quad (2.2)$$

Em síntese, para a maioria dos algoritmos de PNL quando minimizamos $f(u)$ sujeito as restrições dadas em (2.1), os pontos limites das sequências geradas pelos algoritmos pertencem a um dos seguintes conjuntos:

- (a) Pontos (KKT) estacionários de (2.2) que não satisfazem (2.1);
- (b) Pontos factíveis que não satisfazem nenhuma *constraint qualification*;
- (c) Pontos KKT do problema original.

Geralmente, inicia-se a teoria de convergência dos algoritmos de PNL provando que pontos limites são pontos KKT de (2.2). Para alguns problemas (por exemplo, se a região factível é vazia) nada mais pode ser provado. Portanto, é importante caracterizar as situações em que pontos KKT de (2.2) são pontos factíveis. Esta é a motivação do Teorema 2.1.1.

Introduzindo as variáveis de folga $v \in \mathbb{R}^{m_1}$, $w \in \mathbb{R}^q$ em (2) e (5), podemos reformular as restrições da seguinte forma:

$$\begin{aligned} h(x, y) &= 0, \quad g(x, y) + v = 0, \quad v \geq 0, \\ F(x, y) + \nabla_y t(x, y)\lambda + \nabla_y s(x, y)z &= 0, \\ t(x, y) &= 0 \end{aligned} \quad (2.3)$$

$$s(x, y) + w = 0, \quad w \geq 0, \quad z \geq 0, \quad w_i z_i = 0, \quad i = 1, \dots, q.$$

Portanto, dados $r, p > 0$ resolver (2.3) corresponde a

$$\begin{aligned} \min \quad & \|h(x, y)\|^2 + \|g(x, y) + v\|^2 + \\ & \|F(x, y) + \nabla_y t(x, y)\lambda + \nabla_y s(x, y)z\|^2 + \\ & \|t(x, y)\|^2 + \|s(x, y) + w\|^2 + \left(\sum_{i=1}^q (w_i z_i)^r\right)^p \\ \text{s.a.} \quad & v \geq 0, \quad w \geq 0, \quad z \geq 0. \end{aligned} \quad (2.4)$$

A formulação (2.4) corresponde a reformulação do problema de complementaridade não-linear dada em [1]. Geralmente, temos interesse no caso em que $r = 2$, $p = 1$, que corresponde a soma das normas das restrições ao quadrado, como na formulação (2.2).

O próximo teorema, cuja demonstração pode ser encontrada em [5], apresenta algumas condições para que pontos estacionários de (2.4) seja pontos factíveis do MPEC.

Teorema 2.1.1. *Suponha que as seguintes condições são satisfeitas:*

- (a) $(x^*, y^*, v^*, w^*, \lambda^*, z^*)$ é um ponto KKT de (2.4);
- (b) A matriz $\nabla_y F(x^*, y^*) + \sum_{i=1}^q z_i^* \nabla_{yy} [s(x^*, y^*)]_i$ é definida positiva no núcleo de $\nabla_y^T t(x^*, y^*)$;
- (c) $h(x^*, y^*) = 0$, $g(x^*, y^*) \leq 0$;
- (d) $t(x^*, y)$ é uma função afim (em y) e as funções $[s(x^*, y)]_i$ são convexas;
- (e) Existe $\tilde{y}$ tal que $t(x^*, \tilde{y}) = 0$ e $s(x^*, \tilde{y}) \leq 0$;
- (f) $p \geq 1$, $r \geq 1$ e $p + r > 2$.

Então, (x^*, y^*) é um ponto factível do MPEC.

O seguinte teorema, cuja demonstração também pode ser encontrada em [5], mostra que a tese do teorema anterior permanece verdadeira supondo que a matriz “Hessiana” reduzida é semidefinida positiva se supormos que F e s são funções afins com respeito a y .

Teorema 2.1.2. *Suponha que as seguintes condições são satisfeitas:*

- (a) $(x^*, y^*, v^*, w^*, \lambda^*, z^*)$ é um estacionário de (2.4);
- (b) As funções $F(x^*, y)$, $t(x^*, y)$ e $s(x^*, y)$ são funções afins (em y);

(c) A matriz $\nabla_y F(x^*, y^*)$ é semidefinida positiva no núcleo de $\nabla_y^T t(x^*, y^*)$;

(d) $h(x^*, y^*) = 0$, $g(x^*, y^*) \leq 0$;

(e) Existem $\tilde{y}$, $\tilde{\lambda}$, $\tilde{z} \geq 0$ tais que

$$F(x^*, \tilde{y}) + \nabla_y t(x^*, \tilde{y})\tilde{\lambda} + \nabla_y s(x^*, \tilde{y})\tilde{z} = 0,$$

$$t(x^*, \tilde{y}) = 0 \quad e \quad s(x^*, \tilde{y}) \leq 0;$$

(f) $p \geq 1$, $r \geq 1$ e $p + r > 2$.

Então, (x^*, y^*) é um ponto factível do MPEC.

Os Teoremas 2.1.1 e 2.1.2 estabelecem condições suficientes para que algoritmos de PNL encontrem pontos factíveis do MPEC. A grosso modo, as condições sobre o problema dizem que as restrições de equilíbrio representam as condições de otimalidade de um problema convexo. As hipóteses $h(x^*, y^*) = 0$, $g(x^*, y^*) \leq 0$ tem consequências algorítmicas. Elas nos dizem que é importante preservar a factibilidade das restrições ordinárias, já que, não podemos garantir a factibilidade destas restrições sabendo apenas que temos um ponto estacionário da norma das restrições ao quadrado. Portanto, seria interessante talvez, dar preferência a algoritmos que mantêm a factibilidade das restrições ordinárias em todas as iterações. Quando as restrições ordinárias tem uma estrutura simples (por exemplo, quando elas definem uma caixa) manter a factibilidade é uma tarefa fácil, mas a situação é mais complicada quando as restrições são fortemente não-lineares.

2.2 Otimalidade

Suponha que temos um método que converge para pontos factíveis do MPEC. Essencialmente, condições suficientes para se verificar esta propriedade foram

dadas na seção anterior mas, evidentemente, muitos algoritmos geram sequências que convergem para pontos factíveis, mesmo quando as hipóteses do Teorema 2.1.1 não são satisfeitas. Para muitos algoritmos de PNL se tem provado a seguinte propriedade:

Propriedade 1: Se um ponto limite é factível e regular então este ponto satisfaz as condições de KKT para o problema dado.

Considerando a formulação (2.3) podemos notar a presença das restrições

$$w_i z_i = 0, \quad w_i \geq 0, \quad z_i \geq 0, \quad i = 1, \dots, q. \quad (2.5)$$

Para cada i fixo, sendo $h(w, z) = w_i z_i$, $g_1(w, z) = w_i$ e $g_2(w, z) = z_i$, temos que os vetores gradiente das restrições são os seguintes:

$$\underbrace{\begin{pmatrix} z_i e_i \\ w_i e_i \end{pmatrix}}_{\nabla h(w, z)} \quad \underbrace{\begin{pmatrix} e_i \\ 0 \end{pmatrix}}_{\nabla g_1(w, z)} \quad \underbrace{\begin{pmatrix} 0 \\ e_i \end{pmatrix}}_{\nabla g_2(w, z)}$$

onde e_i é o i -ésimo vetor canônico do $\mathbb{R}^q$ e 0 é o vetor nulo do $\mathbb{R}^q$. Para todo par $(w, z) \in \mathbb{R}^{2q}$ satisfazendo (2.5) devemos ter $w_j = 0$ ou $z_j = 0$ para todo $j \in \{1, \dots, i, \dots, q\}$. Suponhamos que $w_i = 0$ e $z_i > 0$. Desta forma, os vetores gradiente das restrições ativas são:

$$\underbrace{z_i \begin{pmatrix} e_i \\ 0 \end{pmatrix}}_{\nabla h(w, z)} \quad \underbrace{\begin{pmatrix} e_i \\ 0 \end{pmatrix}}_{\nabla g_1(w, z)}$$

e daí, podemos notar a dependência linear dos vetores, já que, $\nabla h(w, z) = z_i \nabla g_1(w, z)$. Analogamente, se $z_i = 0$ e $w_i > 0$ os vetores gradiente das restrições ativas são:

$$\underbrace{w_i \begin{pmatrix} 0 \\ e_i \end{pmatrix}}_{\nabla h(w, z)} \quad \underbrace{\begin{pmatrix} 0 \\ e_i \end{pmatrix}}_{\nabla g_2(w, z)}$$

e, novamente, temos a dependência linear dos vetores ($\nabla h(w, z) = w_i \nabla g_2(w, z)$). Se $w_i = z_i = 0$ então $\nabla h(w, z) = 0$ e, portanto, a dependência linear dos vetores gradiente das restrições ativas é um fato evidente. Logo, podemos concluir que para todo ponto factível, os vetores gradiente das restrições ativas são sempre LD. Portanto, devido a presença das restrições (2.5), nenhum ponto factível do MPEC é regular.

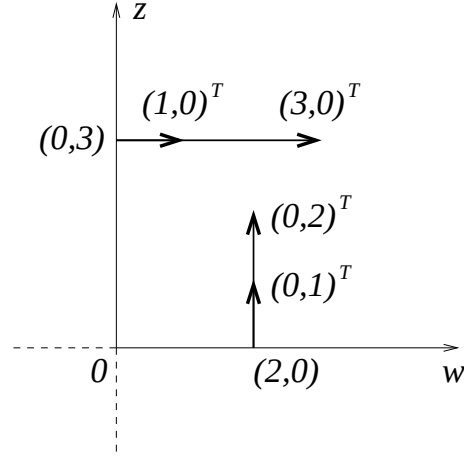
Exemplo 2.2.1. *Consideremos $w, z \in R$.*

Restrições: $h(w, z) = wz = 0$, $g_1(w, z) = w \geq 0$, $g_2(w, z) = z \geq 0$.

Vetores Gradientes:

$$\underbrace{\begin{pmatrix} z \\ w \end{pmatrix}}_{\nabla h(w, z)} \quad \underbrace{\begin{pmatrix} 1 \\ 0 \end{pmatrix}}_{\nabla g_1(w, z)} \quad \underbrace{\begin{pmatrix} 0 \\ 1 \end{pmatrix}}_{\nabla g_2(w, z)}$$

Os pontos factíveis estão sobre os eixos positivos (linha contínua na figura).



Exemplos:

$$w = 0, z = 0 \Rightarrow \nabla h(0, 0) = (0, 0)^T \Rightarrow LD,$$

$$w = 0, z = 3 \Rightarrow \nabla h(0, 3) = (3, 0)^T, \nabla g_2(0, 3) = (1, 0)^T \Rightarrow \nabla h(0, 3) - 3\nabla g_1(0, 3) = 0 \Rightarrow LD,$$

$$w = 2, z = 0 \Rightarrow \nabla h(2, 0) = (0, 2)^T, \nabla g_2(2, 0) = (0, 1)^T \Rightarrow \nabla h(2, 0) - 2\nabla g_2(2, 0) = 0 \Rightarrow LD.$$

A figura ilustra estes exemplos mostrando que, de fato, nenhum ponto factível é regular.

Como nenhum ponto factível do MPEC é regular, segue que um algoritmo que satisfaz a Propriedade 1 pode “convergir” para qualquer ponto factível. Em outras palavras, a Propriedade 1, não acrescenta em nada à teoria de convergência dos algoritmos para o MPEC.

Pior que a falta da regularidade é o fato de que nem mesmo a *constraint qualification* de Mangassarian-Fromovitz é satisfeita qualquer que seja o ponto

factível do MPEC. Este fato se deve, novamente, a presença das restrições (2.5) que não permitem que as hipóteses do Teorema 1.4.5 sejam satisfeitas. Para cada i fixo, vimos que

$$\begin{aligned}\nabla h(w, z) &= 0 && \text{se } w_i = 0, z_i = 0, \\ \nabla h(w, z) &= z_i \nabla g_1(w, z) && \text{se } w_i = 0, z_i > 0, \\ \nabla h(w, z) &= w_i \nabla g_2(w, z) && \text{se } w_i > 0, z_i = 0.\end{aligned}$$

Se $w_i = z_i = 0$ já não temos a independência linear dos vetores gradiente das restrições de igualdade pois $\nabla h(w, z) = 0$. Agora, seja $d \in \mathbb{R}^{2q}$ um vetor que satisfaz $\nabla h(w, z)^T d = 0$. Se $w_i = 0, z_i > 0$ temos $z_i \nabla g_1(w, z)^T d = 0$ mas, como $z_i > 0$ segue que $\nabla g_1(w, z)^T d = 0$. Analogamente, se $w_i > 0, z_i = 0$ temos $\nabla g_2(w, z)^T d = 0$. Logo, não existe vetor $d \in \mathbb{R}^{2q}$ que satisfaça as hipóteses do Teorema 1.4.5. No caso em que $w, z \in \mathbb{R}$ ($q = 1$), este fato pode ser facilmente visualizado na figura do Exemplo 2.2.1.

Portanto, nem os algoritmos que satisfazem a Propriedade 2 são satisfatórios, a menos que algo melhor possa ser provado.

Propriedade 2: Se um ponto limite é factível então este ponto satisfaz as condições de Fritz-John para o problema dado.

Pontos que não satisfazem a *constraint qualification* de MF necessariamente satisfazem as condições de Fritz-John. Portanto, se a Propriedade 2 é satisfeita, temos novamente, um situação onde todos os pontos factíveis são candidatos a serem um ponto limite do algoritmo.

Estes fatos motivaram muitos autores a inventar algoritmos específicos para resolver o MPEC. Nosso objetivo não é analisar algoritmos específicos para o MPEC e sim, discutir a aplicabilidade dos algoritmos gerais de PNL. Portanto, nossa questão é: “Será que o comportamento dos algoritmos gerais de PNL é tão ruim, podendo estes, convergir para qualquer ponto factível do MPEC ?”

Felizmente, a resposta é não. O conjunto dos pontos para os quais os algoritmos podem convergir é, geralmente, muito menor que o conjunto dos pontos que

satisfazem as condições de Fritz-John. Em [21], *Martínez e Svaiter* introduziram uma nova condição prática de otimalidade chamada AGP e verificaram que, em geral, o conjunto dos pontos que satisfazem esta condição é estritamente menor que o conjunto dos pontos que satisfazem as condições de Fritz-John.

2.3 Condição AGP

2.3.1 Motivação

Na teoria de convergência global da maioria dos algoritmos de PNL procura-se provar que alguns pontos limites (talvez todos os pontos limites) das sequências geradas pelos algoritmos satisfazem uma condição necessária. Em geral, esta condição necessária diz que a condição de KKT é satisfeita se o ponto limite satisfaz alguma *constraint qualification*. Logo, uma sequência gerada por um algoritmo poderia convergir para pontos onde a *constraint qualification* não é satisfeita. Isto significa que os algoritmos de PNL poderiam apresentar alguma deficiência na presença de pontos atratores que sejam factíveis e não-regulares mas que não sejam minimizadores locais.

Entretanto, convergência para pontos factíveis não-regulares que não são minimizadores locais não é frequente. Isto não é nenhuma surpresa, já que, os algoritmos são desenvolvidos levando em conta, a cada iteração, a necessidade de diminuir o valor da função objetivo. Então, convergência para pontos que não possuem uma propriedade especial é incomum, pelo menos para os métodos bem desenvolvidos.

Estas observações motivaram *Martínez e Svaiter* a buscar uma condição de otimalidade que ajustasse melhor a conduta dos métodos. Uma outra motivação para encontrar uma condição de otimalidade mais forte que a condição de Fritz-John surgiu do próprio problema MPEC, já que, nenhum ponto factível deste problema satisfaz a *constraint qualification* de Mangassarian-Fromovitz e, desta

forma, todos os pontos factíveis satisfazem a condição de Fritz-John. Então, era necessário estudar uma condição de otimalidade que caracterizasse melhor os minimizadores do problema do que a condição de Fritz-John.

Desta forma, estes autores introduziram em [21] uma nova condição prática de otimalidade que, a grosso modo, diz que a “aproximação do gradiente projetado tende a zero”. Por esta razão, esta nova condição prática de otimalidade é denominada AGP. Esta, é satisfeita por minimizadores locais de problemas com restrições independentemente de qualquer *constraint qualification* e usando apenas diferenciabilidade de primeira ordem. Mais ainda, o conjunto dos pontos que satisfazem esta condição está contido, e pode estar estritamente contido, no conjunto dos pontos que satisfazem a condição de Fritz-John. Assim como nas condições de KKT, a condição AGP também é suficiente se o problema é convexo.

2.3.2 A Condição Necessária

Nesta seção consideraremos problemas da forma:

$$\min f(x) \text{ s.a. } x \in \Omega, \quad (2.6)$$

onde

$$\Omega = \{x \in \mathbb{R}^n \mid h(x) = 0, g(x) \leq 0\},$$

$f : \mathbb{R}^n \rightarrow \mathbb{R}$, $h : \mathbb{R}^n \rightarrow \mathbb{R}^m$, $g : \mathbb{R}^n \rightarrow \mathbb{R}^r$ e todas as funções possuem primeira derivada contínua.

Seja $\gamma \in [-\infty, 0]$. Para todo $x \in \mathbb{R}^n$ definimos o conjunto $\Omega(x, \gamma)$ como sendo o conjunto dos pontos $z \in \mathbb{R}^n$ que satisfazem

$$g_i(x) + g'_i(x)(z - x) \leq 0 \text{ se } \gamma < g_i(x) < 0,$$

$$g'_i(x)(z - x) \leq 0 \text{ se } g_i(x) \geq 0$$

e

$$h'(x)(z - x) = 0.$$

O conjunto $\Omega(x, \gamma)$ (um poliedro fechado e convexo) pode ser interpretado como uma aproximação linear do conjunto dos pontos $z \in \mathbb{R}^n$ que satisfaz

$$h(z) = h(x), \quad g_i(z) \leq g_i(x) \text{ se } g_i(x) \geq 0, \quad g_i(z) \leq 0 \text{ se } g_i(x) \in (\gamma, 0).$$

Para todo $x \in \mathbb{R}^n$, definimos

$$d(x, \gamma) = P_{\Omega(x, \gamma)}(x - \nabla f(x)) - x,$$

onde $P_C(y)$ denota a projeção ortogonal de y em C para todo $y \in \mathbb{R}^n$, $C \subset \mathbb{R}^n$ fechado e convexo. O vetor $d(x, \gamma)$ será chamado de Aproximação do Gradiente Projetado. Enunciaremos, agora, o principal resultado desta seção cuja demonstração pode ser encontrada em [21].

Teorema 2.3.1. *Suponha que x^* é um minimizador local de (2.6) e sejam $\gamma \in [-\infty, 0]$, $\varepsilon > 0$, $\delta > 0$ dados. Então, existe $x \in \mathbb{R}^n$ tal que $\|x - x^*\| \leq \delta$ e $\|d(x, \gamma)\| \leq \varepsilon$.*

Este teorema diz, basicamente, que se x^* é um minimizador local de (2.6), então podemos encontrar pontos arbitrariamente próximos de x^* com o vetor da aproximação do gradiente projetado suficientemente pequeno.

O seguinte corolário especifica a condição AGP de uma maneira que se aproxima da estrutura algorítmica: dado um minimizador local, uma sequência que satisfaz $x^k \rightarrow x^*$ e $d(x^k, \gamma) \rightarrow 0$ necessariamente existe. Esta propriedade é satisfeita por muitos algoritmos de PNL sendo $\{x^k\}$, nestes casos, a sequência gerada pelo algoritmo.

Corolário 2.3.2. *Se x^* é um minimizador local de (2.6) e $\gamma \in [-\infty, 0]$, existe uma sequência $\{x^k\} \subset \mathbb{R}^n$ tal que $\lim x^k = x^*$ e $\lim d(x^k, \gamma) = 0$.*

Finalizaremos esta seção, enunciando o teorema que garante que a condição AGP implica na condição de Fritz-John cuja demonstração também pode ser encontrada em [21].

Teorema 2.3.3. *Seja $\gamma \in [-\infty, 0]$. Suponha que x^* é um ponto factível e que existe uma sequência $x^k \rightarrow x^*$ tal que $d(x^k, \gamma) \rightarrow 0$. Então x^* é um ponto Fritz-John de (2.6).*

O seguinte corolário é uma consequência trivial do teorema anterior e diz que, sob a *constraint qualification* de Mangassarian-Fromovitz (MF), a condição AGP implica nas condições de KKT.

Corolário 2.3.4. *Se as hipóteses do Teorema 2.3.3 são válidas e x^* satisfaz MF, então as condições de KKT são satisfeitas em x^* .*

Observação: O conjunto dos pontos de Fritz-John pode ser estritamente maior que o conjunto dos pontos que satisfazem a condição AGP. Considere o problema

$$\min x \quad \text{s.a. } x^3 \leq 0.$$

Tome $x^* = 0$. Claramente, x^* satisfaz a condição de Fritz-John mas, para qualquer sequência $\{x^k\}$ que converge para x^* , $d(x^k, \gamma)$ não tende a zero. Obviamente, x^* não satisfaz a *constraint qualification* de MF.

Resumindo, a condição AGP é uma condição de otimalidade que

- é satisfeita por minimizadores locais de problemas com restrições independentemente de qualquer *constraint qualification*;
- implica na condição de Fritz-John, ou seja, o conjunto dos pontos AGP está contido no conjunto dos pontos que satisfazem a condição de Fritz-John podendo esta inclusão, ser estrita, como mostra a observação acima;
- implica nas condições de KKT quando a *constraint qualification* de MF é satisfeita.

Logo, o MPEC é um bom exemplo de problema onde a condição AGP é útil, já que, o conjunto dos pontos AGP se aproxima mais do conjunto dos minimizadores locais do que o conjunto dos pontos que satisfazem a condição de Fritz-John. Então, algoritmos em que temos a garantia de convergência para pontos AGP são candidatos a resolver o MPEC.

2.4 Aplicabilidade da Condição AGP ao MPEC

A aplicabilidade da condição AGP ao problema MPEC vem Teorema 2.4.1. Este, mostra que o conjunto dos pontos que satisfazem a condição AGP é consideravelmente menor que o conjunto dos pontos que satisfazem a condição de Fritz-John.

Considere as restrições do problema MPEC dadas na forma (2.3). Assim, o problema MPEC é um caso particular do problema

$$\begin{aligned}
 \min \quad & \Phi(x, v, w, z) \\
 \text{s.a.} \quad & H(x, v, w, z) = 0, \\
 & w_i z_i = 0, \quad i = 1, \dots, q, \\
 & v \geq 0, \quad w \geq 0, \quad z \geq 0,
 \end{aligned} \tag{2.7}$$

onde $x \in \mathbb{R}^n$, $v \in \mathbb{R}^m$, $w \in \mathbb{R}^q$, $z \in \mathbb{R}^q$, $H(x, v, w, z) \in \mathbb{R}^l$. Usaremos alguns abusos de notação para evitar um número excessivo de variáveis e diferentes dimensões. Lembre-se que a falta de regularidade dos pontos factíveis deste problema se deve a presença simultânea das restrições $w_i z_i = 0$ e $w \geq 0$, $z \geq 0$.

Suponha que (x^*, v^*, w^*, z^*) é um ponto factível do problema (2.7) e que $w_i^* + z_i^* > 0$ para todo $i \in I_* \subset \{1, \dots, q\}$. Definiremos um problema associado a (2.7) que consiste, simplesmente, em retirar as restrições $w_i \geq 0$, $z_i \geq 0$ para $i \in I_*$. Então o problema associado é

$$\begin{aligned}
 \min \quad & \Phi(x, v, w, z) \\
 \text{s.a.} \quad & H(x, v, w, z) = 0, \\
 & w_i z_i = 0, \quad i = 1, \dots, q, \\
 & v \geq 0, \quad w_i \geq 0, \quad z_i \geq 0 \quad \forall i \notin I_*
 \end{aligned} \tag{2.8}$$

O teorema a seguir mostra que se (x^*, v^*, w^*, z^*) é um ponto AGP de (2.7) então este ponto é, necessariamente, um ponto AGP de (2.8). Isto significa que o conjunto dos pontos de acumulação (dos algoritmos) que satisfazem a condição

AGP para o problema (2.7) são pontos AGP de um problema associado onde as restrições que fazem com que todos os pontos factíveis sejam não-regulares, provavelmente, não estão mais presentes. A consequência é que a maioria dos pontos (Fritz-John) factíveis do MPEC são descartados como possíveis pontos limite do algoritmo. A demonstração do teorema pode ser encontrada em [5].

Teorema 2.4.1. *Suponha que (x^*, v^*, w^*, z^*) é um ponto AGP factível de (2.7) tal que $w_i^* + z_i^* > 0 \forall i \in I_*$. Então, (x^*, v^*, w^*, z^*) é um ponto AGP de (2.8).*

O teorema diz que os pontos AGP do problema (2.7) são pontos AGP de (2.8). Portanto, o conjunto dos candidatos a ponto limite dos algoritmos que convergem para pontos AGP é restringido pelos pontos AGP de (2.8). Quando, em um dado ponto (excluindo aqueles tais que $w_i = z_i = 0$), a *constraint qualification* de Mangassarian-Fromovitz para o problema (2.8) é satisfeita (em particular, quando os vetores gradientes das restrições ativas de (2.8) forem linearmente independentes), a condição AGP implica nas condições de KKT. Como as condições de KKT para (2.8) evidentemente implicam nas condições de KKT para (2.7), podemos estabelecer o seguinte teorema.

Teorema 2.4.2. *Suponha que (x^*, v^*, w^*, z^*) é um ponto AGP factível de (2.7) tal que $w_i^* + z_i^* > 0 \forall i = 1, \dots, q$ e que (x^*, v^*, w^*, z^*) satisfaz a *constraint qualification* de MF para o problema (2.8). Então, (x^*, v^*, w^*, z^*) é um ponto KKT de (2.7).*

Capítulo 3

Resolução Numérica do Problema MPEC

Nosso objetivo não é analisar algoritmos específicos para o MPEC e sim, discutir a aplicabilidade dos algoritmos gerais de PNL. Desta forma, o capítulo anterior nos conduz a utilizar algoritmos de PNL em que temos a garantia de convergência para pontos AGP. Existem bons métodos de PNL que satisfazem a condição AGP. Esta é uma questão de simples verificação em alguns casos. A condição é explicitamente provada em algoritmos de restauração inexata [15, 17, 18, 19], pode ser facilmente verificada em métodos de lagrangeano aumentado [7, 11, 16] e, provalvemente, é satisfeita por vários outros algoritmos clássicos de PNL. Em outros casos, como em [23], pode ser necessária uma análise um pouco mais detalhada. Neste capítulo, apresentaremos as estratégias, os algoritmos (métodos) utilizados e os resultados numéricos obtidos ao tentarmos resolver alguns problemas MPEC da literatura.

3.1 Algoritmo de Restauração Inexata

Um problema clássico da otimização é minimizar uma função qualquer sujeito a restrições não-lineares. Exceto em casos muito particulares, é possível encontrar

a solução analítica destes problemas. Portanto, métodos iterativos devem ser utilizados. Uma das idéias mais naturais é gerar uma sequência $\{x^k\}$ que satisfaça as restrições tal que o valor da função objetivo f decresça progressivamente. Desta forma, esperamos que o limite desta sequência seja a solução do problema. Entretanto, manter a factibilidade quando as restrições são não-lineares é difícil.

Geralmente, os métodos escolhem uma direção d^k para se mover a partir do ponto corrente x^k mas, se as restrições são não-lineares, os pontos da forma $x^k + \alpha d^k$ podem ser infactíveis, não importando quão pequeno seja o parâmetro α . Esta observação implica que, quando tivermos restrições não-lineares, o método deve estar preparado para perder e restaurar a factibilidade a cada iteração. O processo de trazer um ponto não factível para a factibilidade é chamada de restauração.

A principal diferença entre os métodos clássicos está na maneira em que é realizada a fase de restauração. Métodos factíveis, que realizam a restauração de forma exata, são muito importantes na prática porque em várias situações da vida real soluções não-ótimas factíveis tem um significado físico e, portanto, podem ser úteis enquanto que pontos não-factíveis não tem significado algum. Porém, o grande prejuízo dos métodos factíveis é que eles tendem a se comportar mal na presença de não-linearidades muito fortes. Nestes casos, não é possível dar grandes passos longe da solução porque a não-linearidade das restrições força com que a distância entre iterações factíveis consecutivas seja pequena.

Estes fatos, conduziram ao desenvolvimento de algoritmos que embora mantenham a infactibilidade sobre controle sejam tolerantes quando a iteração esteja longe da solução. Estes são chamados de algoritmos de restauração inexata.

Algoritmos de restauração inexata podem ser encontrados em [17, 18, 19]. Baseando-se nestas referências, apresentaremos a seguir, as idéias principais deste algoritmo. Além de termos a garantia de convergência para pontos que satisfaçam a condição AGP, a prova de convergência deste algoritmo não necessita da hipótese de regularidade. Logo, temos algumas razões adicionais para aplicar o

algoritmo restauração inexata ao problema MPEC.

3.1.1 Idéias Principais do Algoritmo

Consideremos o problema

$$\begin{aligned} \min f(x) \\ \text{s.a. } C(x) = 0, \quad x \in \Omega \end{aligned} \tag{3.1}$$

onde $f : \mathbb{R}^n \rightarrow \mathbb{R}$ e $C : \mathbb{R}^n \rightarrow \mathbb{R}^m$ são continuamente diferenciáveis e $\Omega \subset \mathbb{R}^n$ é fechado e convexo. Na prática, estaremos interessados no caso em que Ω é um politopo.

O objetivo é definir um modelo de algoritmo iterativo para resolver (3.1) tal que todas as iterações x^k sejam factíveis em relação a Ω e cada iteração consista, essencialmente, de duas fases: Restauração e Minimização.

Na fase de restauração, um movimento em direção a factibilidade é realizada, com uma pequena exigência relativa a redução de $\|C(x)\|_2$. Como resultado, um ponto intermediário $y^k \in \Omega$ é obtido. Imediatamente após a fase de restauração, construiremos um aproximação π_k da região factível que passa sobre y^k .

Na fase de minimização computaremos um ponto $z^{k,i} \in \pi_k$ tal que $f(z^{k,i}) \ll f(y^k)$ ($\ll$ significa “suficiente menor que”) e $\|z^{k,i} - y^k\| \leq \delta_{k,i}$ onde $\delta_{k,i}$ é o raio da região de confiança.

Entretanto, embora y^k seja “mais factível” que x^k e $z^{k,i}$ seja “mais ótimo” que y^k (no sentido que $f(z^{k,i}) < f(y^k)$), o ponto $z^{k,i}$ pode ser pior que x^k (menos factível, menos ótimo, ou ambos). Diante disto, somente aceitaremos $z^{k,i}$ como uma nova iteração se o valor de uma função de mérito não-diferenciável (penalização exata) em $z^{k,i}$ for suficientemente menor que o valor desta função avaliada em x^k . Se $z^{k,i}$ não for aceito, o raio da região de confiança é reduzido.

Uma característica muito importante dos algoritmos de restauração inexata é que temos a liberdade de escolha dos algoritmos utilizados na fase de restauração.

Isto nos permite explorar as características das restrições que podem não ser levadas em consideração por outros algoritmos de programação não-linear.

3.1.2 Descrição do Algoritmo

Antes de dar a descrição rigorosa do algoritmo, comentaremos algumas das suas principais características.

Notação: nesta seção $||\cdot||$ vai representar uma norma arbitrária e $|\cdot|$ denotará a norma euclidiana. Denotaremos a matriz jacobiana de $C(x)$ por $C'(x) \in \mathbb{R}^{m \times n}$ e $C'_j(x) = \nabla C_j(x)^T$ para todo $j = 1, \dots, m$.

Fase de Restauração

Como foi mencionado anteriormente, dado uma iteração $x^k \in \Omega$, o modelo de algoritmo computa um ponto intermediário $y^k \in \Omega$ “mais factível”. As condições que devem ser satisfeitas por y^k são:

$$|C(y^k)| \leq r|C(x^k)| \quad (3.2)$$

$$||y^k - x^k|| \leq \beta|C(x^k)| \quad (3.3)$$

onde $r \in [0, 1)$ e $\beta > 0$ são parâmetros dados independentes de k . A condição (3.2) determina a necessidade de obter um ponto intermediário pelo menos tão factível quanto x^k . A condição (3.3) impõe que y^k deve ser igual a x^k se a iteração corrente for factível.

Aproximação Linear da Região Factível

Após encontrar y^k satisfazendo as condições (3.2) e (3.3) construiremos a aproximação linear da região factível do problema (3.1) que contém o ponto intermediário y^k . Esta região auxiliar é dada por:

$$\pi_k = \{z \in \Omega \mid C'(y^k)(z - y^k) = 0\}.$$

Se y^k é um ponto factível então o conjunto π_k representa uma linearização do conjunto factível em y^k .

Fase de Minimização

O objetivo da fase de minimização é obter $z^{k,i} \in \pi_k \cap B_{k,i}$ tal que $f(z^{k,i}) \ll f(y^k)$, onde

$$B_{k,i} = \{x \in \mathbb{R}^n \mid \|x - y^k\| \leq \delta_{k,i}\},$$

e $\delta_{k,i} > 0$ é o raio da região de confiança.

A fase de minimização é precedida pelo cálculo da direção tangente de Cauchy (independente de i)

$$d^{k,tan} = P_k(y^k - \eta \nabla f(y^k)) - y^k, \quad (3.4)$$

onde $P_k(z)$ denota a projeção ortogonal de z em π_k e $\eta > 0$ é uma constante arbitrária independente de k . Daí segue que $d^{k,tan}$ é uma direção de descida factível em π_k . Assim, a norma de $d^{k,tan}$ vai ser usada para definir o critério de parada do algoritmo.

O ponto $y^k + d^{k,tan}$ pertence a π_k mas não necessariamente pertence a $B_{k,i}$. Então, definiremos o “breakpoint” $y^k + t_{(k,i,break)} d^{k,tan}$ onde

$$t_{(k,i,break)} = \sup\{t \in [0, 1] \mid [y^k, y^k + t d^{k,tan}] \subset B_{k,i}\}. \quad (3.5)$$

Mais ainda, o valor da função objetivo f em $y^k + t_{(k,i,break)} d^{k,tan}$ não é necessariamente menor que $f(y^k)$, portanto um valor suficientemente pequeno da função objetivo $f(y^k + t_{(k,i,dec)} d^{k,tan})$ deve ser obtido usando o clássico procedimento “backtracking”. Finalmente, $z^{k,i} \in \pi_k \cap B_{k,i}$ vai ser qualquer ponto tal que $f(z^{k,i}) \leq f(y^k + t_{(k,i,dec)} d^{k,tan})$. Alternativamente, $z^{k,i}$ pode ser qualquer ponto de $\pi_k \cap B_{k,i}$ tal que $f(y^k) - \tau_1 \delta_{k,i}$ ou $f(y^k) - \tau_2$, onde τ_1 e τ_2 são parâmetros não negativos do algoritmo.

Isto significa que para calcular $z^{k,i}$ de forma eficiente, podemos aplicar um algoritmo de minimização (com critério de parada razoável) para resolver o seguinte

problema

$$\min f(x) \text{ s.a. } x \in \pi_k \cap B_{k,i}. \quad (3.6)$$

Se $\|\cdot\|$ é a norma infinito então (3.6) representa um problema de minimização com restrições lineares.

Função de Mérito e Parâmetro de Penalização

A comparação entre $z^{k,i}$ e x^k envolve a avaliação de uma função de mérito em ambos os pontos. Utilizaremos a função de penalização exata que é dada por:

$$\psi(x, \theta) = \theta f(x) + (1 - \theta)|C(x)|$$

onde $\theta \in (0, 1]$ é o parâmetro de penalização utilizado para dar diferentes pesos para a função objetivo e para a factibilidade. A escolha do parâmetro θ em cada iteração depende de considerações práticas e teóricas. Por exemplo, se $|C(x)|$ é grande, o peso atribuído a $f(x)$ deve ser pequeno, pois não faz sentido se preocupar com o valor da função objetivo se o ponto corrente está longe da região factível.

Em síntese, desejaremos que a função de mérito avaliada no novo ponto seja menor que o valor da função de mérito avaliada em x^k . Ou seja, queremos $Ared_{k,i} > 0$ onde $Ared_{k,i}$, a redução real (“actual reduction”) da função de mérito, é definida por:

$$Ared_{k,i} = \psi(x^k, \theta_{k,i}) - \psi(z^{k,i}, \theta_{k,i}).$$

Então,

$$Ared_{k,i} = \theta_{k,i}[f(x^k) - f(z^{k,i})] + (1 - \theta_{k,i})[|C(x^k)| - |C(z^{k,i})|]. \quad (3.7)$$

Entretanto, como nos problemas sem restrições, uma redução da função de mérito não é suficiente para garantir a convergência. De fato, precisamos de uma redução suficiente, que vai ser definida como satisfatória se

$$Ared_{k,i} \geq 0.1Pred_{k,i}, \quad (3.8)$$

onde $Pred_{k,i}$ é a redução pretendida (“predicted reduction”) da função de mérito entre x^k e $z^{k,i}$. No nosso caso, definimos

$$Pred_{k,i} = \theta_{k,i}[f(x^k) - f(z^{k,i})] + (1 - \theta_{k,i})[|C(x^k)| - |C(y^k)|]. \quad (3.9)$$

A quantidade $Pred_{k,i}$ definida acima pode ser não-positiva dependendo do valor do parâmetro de penalização. Mas, se $\theta_{k,i}$ é suficientemente pequeno, $Pred_{k,i}$ é arbitrariamente próximo a $|C(x^k)| - |C(y^k)|$ que é necessariamente não negativo. Portanto, sempre poderemos escolher $\theta_{k,i} \in (0, 1]$ tal que

$$Pred_{k,i} \geq \frac{1}{2}[|C(x^k)| - |C(y^k)|].$$

Quando o critério (3.8) for satisfeito aceitaremos $z^{k,i}$ como uma nova iteração, ou seja, $x^{k+1} = z^{k,i}$. Caso contrário, reduziremos o raio da região de confiança.

Modelo de Algoritmo

Assuma que $\eta > 0$, $\beta > 0$, $r \in [0, 1)$, $\tau_1 > 0$, $\tau_2 > 0$ são parâmetros do algoritmo independente de k e $\sum_{k=0}^{\infty} w_k$ é uma série convergente de termos não-negativos. Suponha que $x^0 \in \Omega$ é uma aproximação inicial da solução e que $\theta_{-1} \in (0, 1)$ é um parâmetro de penalização inicial. Dados $x^k \in \Omega$, $\theta_{k-1} \in (0, 1]$ e $\delta_{k,0}$, os passos para calcular x^{k+1} ou para interromper o processo são dados pelo seguinte algoritmo.

Algoritmo 3.1.

Passo 1: *Inicializando o parâmetro de penalização*

Defina

$$\theta_{k,-1} = \min\{1, \min\{\theta_{-1}, \dots, \theta_{k-1}\} + w_k\}.$$

Passo 2: *Fase de Restauração*

Calcule $y^k \in \Omega$ tal que (3.2) e (3.3) sejam satisfeitas. Se isto não for possível, pare a execução do algoritmo declarando “falha ao melhorar a factibilidade”. Caso contrário, faça $i \leftarrow 0$.

Passo 3: *Direção Tangente de Cauchy*

Calcule $d^{k,tan}$ usando (3.4). Se $C(x^k) = 0$ e $d^{k,tan} = 0$ termine a execução do algoritmo declarando “convergência finita” e x^k como solução.

Passo 4: *Fase de Minimização*

Calcule $t_{(k,i,break)}$ usando (3.5). Defina $t_{(k,i,dec)}$ como sendo o primeiro termo da sequência $\{t_{k,1}, t_{k,2}, \dots\}$ tal que

$$f(y^k + td^{k,tan}) \leq f(y^k) + 0.1t\langle \nabla f(y^k), d^{k,tan} \rangle,$$

onde $\{t_{k,j}\}$ é definido por $t_{k,1} = t_{(k,i,break)}$ e $t_{k,j+1} \in [0.1t_{k,j}, 0.9t_{k,j}]$ para todo $j = 1, 2, \dots$. Calcule $z^{k,i} \in \pi_k \cap B_{k,i}$ tal que

$$f(z^{k,i}) \leq \max\{f(y^k + t_{(k,i,dec)}d^{k,tan}), f(y^k) - \tau_1\delta_{k,i}, f(y^k) - \tau_2\}.$$

Passo 5: *Escolha do parâmetro de penalização*

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

$$Pred_{k,i}(\theta) = \theta[f(x^k) - f(z^{k,i})] + (1 - \theta)[|C(x^k)| - |C(y^k)|].$$

Escolha $\theta_{k,i}$ o supremo dos valores de θ no intervalo $[0, \theta_{k,i-1}]$ tal que

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

Passo 5: *Aceitação ou Rejeição de $z^{k,i}$*

Defina $Ared_{k,i}$ e $Pred_{k,i}$ como em (3.7) e (3.9). Se o teste (3.8) for satisfeito, defina $x^{k+1} = z^{k,i}$, $\theta_k = \theta_{k,i}$ e volte ao Passo 1. Se o teste (3.8) não for satisfeito, escolha $\delta_{k,i+1} \in [0.1\delta_{k,i}, 0.9\delta_{k,i}]$, faça $i \leftarrow i + 1$ e volte ao Passo 4.

3.2 Estratégias

Nossa primeira estratégia é resolver o problema MPEC, dado na sua forma padrão, aplicando algoritmos gerais de PNL. Naturalmente, escolhemos algoritmos em que temos garantia de convergência para pontos AGP.

São eles: Algoritmo de Lagrangeano Aumentado, Algoritmo de Programação Quadrática Sequencial e o Algoritmo de Restauração Inexata.

O algoritmo de lagrangeano aumentado que utilizamos foi o EASY. Este software foi desenvolvido pelo grupo de otimização da Universidade de Campinas e está disponível na página www.ime.unicamp.br/~martinez/software.htm. Utilizamos também a rotina DNCONG da Biblioteca IMSL do Fortran para representar os algoritmos de programação quadrática sequencial. O algoritmo de restauração inexata é o descrito na seção anterior, ou seja, fizemos a implementação do Algoritmo 3.1. É interessante, então, fazermos alguns comentários sobre esta implementação.

A fase de restauração tem como objetivo encontrar um ponto intermediário y^k satisfazendo (3.2) e (3.3). Desde que este ponto satisfaça estas condições a prova de convergência deste algoritmo nos fornece total liberdade de escolha na maneira em que vamos realizar o Passo 2 do Algoritmo 3.1. Desta forma, definiremos o ponto intermediário y^k como sendo o *arg min* do seguinte problema:

$$\begin{aligned} \min \quad & \frac{1}{2} \|C(x)\|_2^2 \\ \text{s.a.} \quad & \|x - x^k\|_\infty \leq \beta \|C(x^k)\|_2, \\ & x \in \Omega. \end{aligned}$$

Para realizar o Passo 4 do Algoritmo 3.1 que corresponde a fase de minimização, definiremos $z^{k,i}$ como sendo o *arg min* do problema

$$\min f(x) \text{ s.a. } x \in \pi_k \cap B_{k,i}.$$

Antes de realizar a fase de minimização temos que calcular, no Passo 3 do Algoritmo 3.1, a direção tangente de cauchy definida por (3.4). Desta forma, sendo d^k o *arg min* do problema

$$\begin{aligned} \min \quad & \frac{1}{2} \|x - (y^k - \eta \nabla f(y^k))\|_2^2 \\ \text{s.a.} \quad & x \in \pi_k, \end{aligned}$$

podemos calcular $d^{k, \text{tan}}$ como sendo $d^k - y^k$.

Assim, podemos observar que a cada iteração deste algoritmo temos que resolver alguns subproblemas. Em nossa implementação utilizamos o próprio EASY para resolvê-los. Quanto ao aspecto de desempenho computacional, certamente, esta não é a melhor opção. Mas, não nos preocupamos com isso pois nosso objetivo, à princípio, é observar o comportamento dos algoritmos de restauração inexata quanto a qualidade de solução. Em outras palavras, o objetivo é descobrir se os algoritmos de restauração inexata podem ser aplicados com sucesso ou não ao MPEC. Melhorar o desempenho computacional deste algoritmo pode ser o objetivo de pesquisas futuras.

Uma outra observação que podemos fazer é que tanto o EASY quanto o algoritmo de restauração inexata que implementamos foram desenvolvidos para trabalhar somente com restrições de igualdade. Então, consideraremos a formulação (2.3) para o MPEC.

Motivados pelo Teorema 2.4.1, vamos propor uma outra alternativa para a resolução do problema MPEC. Esta alternativa, a qual denominaremos por estratégia da soma, consiste em trocar as restrições

$$w_i z_i = 0, \quad w_i \geq 0, \quad z_i \geq 0, \quad i = 1, \dots, q \quad (3.10)$$

pelas restrições

$$w_i z_i = 0, \quad w_i + z_i \geq 0, \quad i = 1, \dots, q. \quad (3.11)$$

Estas novas restrições não alteram o conjunto factível. Devido a restrição $w_i z_i = 0$ temos que pelo menos uma das coordenadas w_i ou z_i deve ser nula. Supondo que w_i é igual a zero então a restrição $w_i + z_i \geq 0$ vai garantir a não-negatividade de z_i , ou seja, $z_i \geq 0$ e vice-versa. A vantagem desta estratégia é que saímos de uma situação onde todos os pontos factíveis são não-regulares para uma outra situação em que pelo menos os pontos que cumprem a complementaridade estrita (uma das coordenadas w_i ou z_i é diferente de zero para todo $i = 1, \dots, q$) deixam de ser não-regulares.

Note que as restrições de complementaridade (3.10) podem ser escritas compactamente da seguinte forma:

$$w^T z = 0, \quad w \geq 0, \quad z \geq 0. \quad (3.12)$$

Infelizmente, as restrições

$$w^T z = 0, \quad w + z \geq 0 \quad (3.13)$$

não representam o mesmo conjunto factível correspondente as restrições (3.11). Por exemplo, consideremos $w = (1, 1)$, $z = (1, -1) \in \mathbb{R}^2$. Este par (w, z) é factível para as restrições acima mas não satisfazem as restrições (3.11). Logo, uma desvantagem da estratégia da soma é que a quantidade de restrições é consideravelmente maior, na proporção de $q - 1$ onde q é a dimensão de w e z .

Utilizaremos o EASY e a rotina do IMSL para verificar o comportamento desta estratégia. Assim, se considerarmos que as restrições ($w \geq 0$, $z \geq 0$) são fornecidas como limitantes (“bounds”) destes algoritmos, esta proporção aumenta para $2q - 1$. Uma outra desvantagem da estratégia da soma, em particular para o EASY, é o aumento do número de variáveis correspondentes as q variáveis de folga introduzidas a cada uma das restrições $w_i + z_i \geq 0$ (lembre-se que o EASY foi projetado para otimizar problemas com restrições de igualdade).

Resumindo, as estratégias que estamos propondo para resolução do problema MPEC são as seguintes:

- EASY - Algoritmo de Lagrangeano Aumentado.
- IMSL - Algoritmo de Programação Quadrática Sequencial.
- IR - Algoritmo de Restauração Inexata.
- E-SOMA - Estratégia da Soma com o EASY.
- I-SOMA - Estratégia da Soma com o IMSL.

3.3 Resultados Numéricos

Iniciaremos esta seção fornecendo algumas informações iniciais com respeito aos experimentos numéricos:

- ✓ Máquina: Athlon - 1Ghz - 256Mb;
- ✓ Sistema Operacional: Windows 98;
- ✓ Linguagem de Programação: Fortran.

Para realizar estes experimentos procuramos utilizar MPEC's da literatura. Nossas fontes foram as seguintes:

- ✓ MPECLib - MPEC World Home Page
<http://www.gamsworld.org/mpec/>
- ✓ Handbook of Test Problems in Local and Global Optimization
<http://titan.princeton.edu/TestProblems/>
- ✓ [8] F. Facchinei, H. Jiang and L. Qi
A smoothing method for mathematical programs with equilibrium constraints (1999).

Realizamos os experimentos com 15 problemas retirados das referências acima. Para cada problema foram realizados 20 testes associados à 20 pontos iniciais distintos gerados aleatoriamente.

As tabelas e gráficos, à seguir, irão fornecer uma idéia do comportamento de cada estratégia na resolução do problema MPEC. As tabelas apresentam a seguinte organização:

- Problema: temos um número associado (apenas para controle), o nome do problema na literatura e a informação se a solução deste problema satisfaz a complementaridade estrita ou não, ou seja, vai informar se o problema é degenerado (DG) ou não.

DG - a solução do problema não satisfaz a complementaridade estrita. Vamos atribuir a DG (na tabela) alguma das situações: S (sim) ou N (não).

- Tamanho: fornece a quantidade de variáveis e restrições necessárias para a implementação do problema teste,

n é a número de variáveis e m é o número de restrições.

- Resultado: ao realizarmos cada teste ocorria uma das seguinte situações:

S1 - obter como resultado final a solução do problema (segundo à referência fonte do mesmo);

S2 - obter como resultado final uma outra solução, ou seja, um outro ponto estacionário cujo valor objetivo da função no ponto é maior que a suposta solução do problema (é um possível minimizador local);

S3 - obter como resultado final um ponto estacionário do quadrado das restrições não factível, quando tratarmos dos algoritmos relacionados com o EASY ou, a declaração de execução de um erro fatal de algum dos tipos:

- (1) Search direction unphil;
- (2) Line search took more than 5 function calls;
- (3) Maximum number of iterations exceeded;
- (4) Search direction is close to zero;
- (5) The constraints for the QP subproblem are inconsistent;

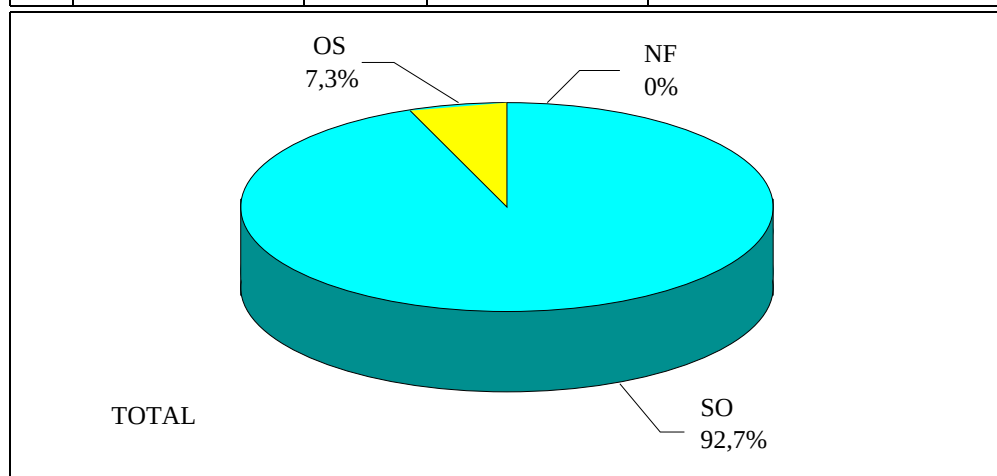
quando considerarmos os algoritmos da Biblioteca IMSL. Quando ocorrer a situação S3 diremos que a estratégia não funcionou. Considerando que para cada problema foram realizados 20 testes achamos interessante contabilizar a quantidade de testes em ocorreram cada uma das situações S1, S2 e S3. Desta forma, SO (SOlução), OS (Outra Solução) e NF (Não Funcionou) representam as seguintes situações:

SO - quantidade de testes em que ocorreram a situação S1;

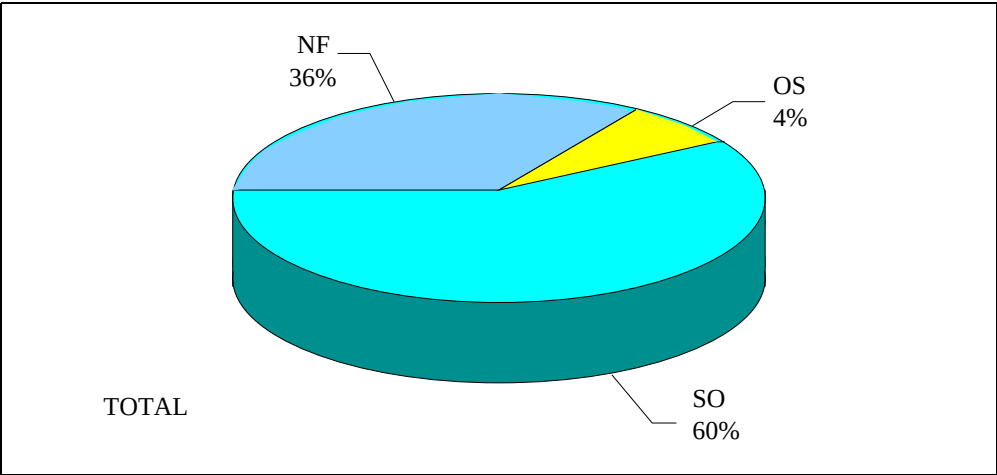
OS - quantidade de testes em que ocorreram a situação S2;

NF - quantidade de testes em que ocorreram a situação S3.

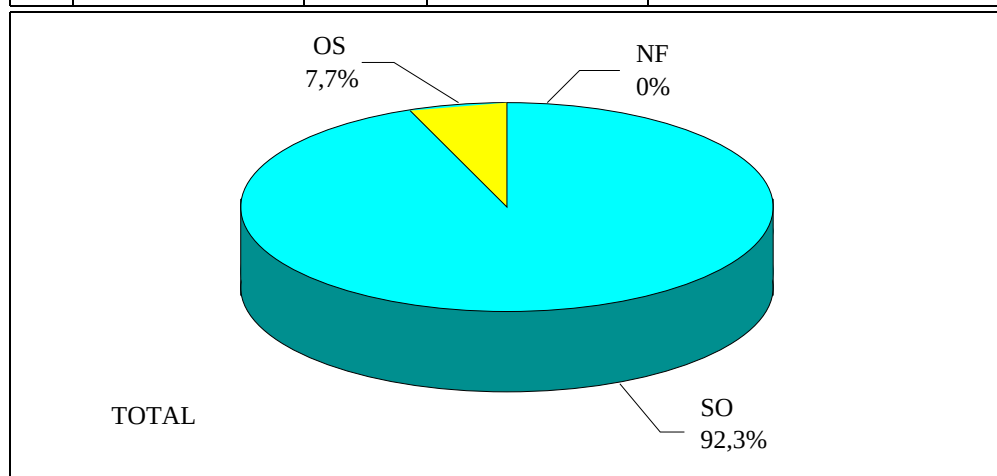
Estratégia: EASY							
Problema			Tamanho		Resultado		
	nome	DG	<i>n</i>	<i>m</i>	SO	OS	NF
1	fjq1	S	11	7	20	0	0
2	fjq2	N	11	7	20	0	0
3	fjq3	N	11	7	20	0	0
4	fjq4	N	11	7	20	0	0
5	desilva	S	8	5	20	0	0
6	hq1	N	3	2	20	0	0
7	fjq9	N	8	5	20	0	0
8	bard3	N	13	8	16	4	0
9	clark-west	N	10	6	14	6	0
10	falk-liu	S	12	7	20	0	0
11	shim-aiy	S	9	6	20	0	0
12	viswes1	N	17	10	17	3	0
13	viswes2	S	10	6	15	5	0
14	yezza1	N	8	6	16	4	0
15	yezza2	S	6	4	20	0	0
	TOTAL				278	22	0



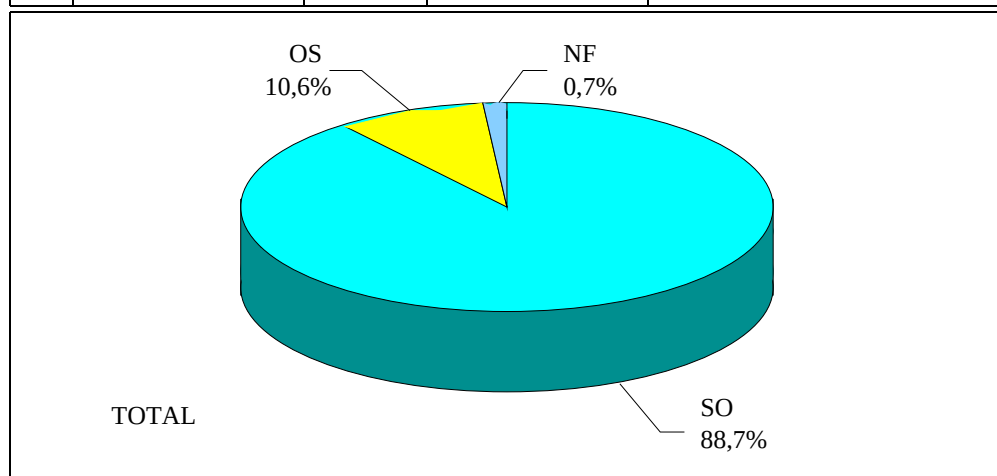
Estratégia: IMSL							
Problema			Tamanho		Resultado		
	nome	DG	<i>n</i>	<i>m</i>	SO	OS	NF
1	fjq1	S	11	7	11	0	9
2	fjq2	N	11	7	10	0	10
3	fjq3	N	11	7	8	0	12
4	fjq4	N	11	7	8	0	12
5	desilva	S	8	5	17	0	3
6	hq1	N	3	2	20	0	0
7	fjq9	N	8	5	19	0	1
8	bard3	N	13	8	7	0	13
9	clark-west	N	10	6	7	6	7
10	falk-liu	S	12	7	15	0	5
11	shim-aiy	S	9	6	2	0	18
12	viswes1	N	17	10	6	1	13
13	viswes2	S	10	6	16	1	3
14	yezza1	N	8	6	16	4	0
15	yezza2	S	6	4	18	0	2
	TOTAL				180	12	108



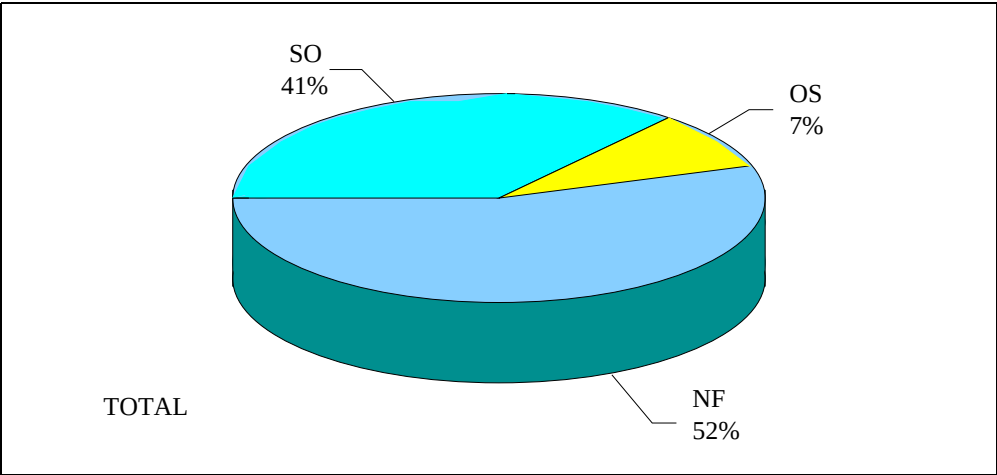
Estratégia: IR							
Problema			Tamanho		Resultado		
	nome	DG	<i>n</i>	<i>m</i>	SO	OS	NF
1	fjq1	S	11	7	20	0	0
2	fjq2	N	11	7	20	0	0
3	fjq3	N	11	7	20	0	0
4	fjq4	N	11	7	20	0	0
5	desilva	S	8	5	20	0	0
6	hq1	N	3	2	20	0	0
7	fjq9	N	8	5	20	0	0
8	bard3	N	13	8	16	4	0
9	clark-west	N	10	6	14	6	0
10	falk-liu	S	12	7	20	0	0
11	shim-aiy	S	9	6	20	0	0
12	viswes1	N	17	10	16	4	0
13	viswes2	S	10	6	15	5	0
14	yezza1	N	8	6	16	4	0
15	yezza2	S	6	4	20	0	0
	TOTAL				277	23	0



Estratégia: E-SOMA							
Problema			Tamanho		Resultado		
	nome	DG	<i>n</i>	<i>m</i>	SO	OS	NF
1	fjq1	S	15	14	20	0	0
2	fjq2	N	15	14	18	2	0
3	fjq3	N	15	14	20	0	0
4	fjq4	N	15	14	20	0	0
5	desilva	S	10	8	20	0	0
6	hq1	N	4	3	20	0	0
7	fjq9	N	10	8	20	0	0
8	bard3	N	17	15	7	11	2
9	clark-west	N	14	13	10	10	0
10	falk-liu	S	16	14	20	0	0
11	shim-aiy	S	12	11	20	0	0
12	viswes1	N	23	21	16	4	0
13	viswes2	S	14	13	19	1	0
14	yezza1	N	10	9	16	4	0
15	yezza2	S	8	7	20	0	0
	TOTAL				266	32	2



Estratégia: I-SOMA							
Problema			Tamanho		Resultado		
	nome	DG	<i>n</i>	<i>m</i>	SO	OS	NF
1	fjq1	S	11	14	6	0	14
2	fjq2	N	11	14	7	1	12
3	fjq3	N	11	14	8	0	12
4	fjq4	N	11	14	7	0	13
5	desilva	S	8	8	1	0	19
6	hq1	N	3	3	20	0	0
7	fjq9	N	8	8	16	0	4
8	bard3	N	13	15	7	6	7
9	clark-west	N	10	13	8	6	6
10	falk-liu	S	12	14	4	0	16
11	shim-aiy	S	9	11	2	0	18
12	viswes1	N	17	21	1	1	18
13	viswes2	S	10	13	13	3	4
14	yezza1	N	8	9	16	4	0
15	yezza2	S	6	7	7	0	13
	TOTAL				123	21	156



Como já foi dito anteriormente, o objetivo deste trabalho é discutir a aplicabilidade de algoritmos gerais de PNL ao problema MPEC. Discutimos a questão da falta de regularidade e o fato de todos os pontos factíveis do MPEC satisfazerem a condição de Fritz-John. Diante disto, chegamos até a questionar se os algoritmos de PNL poderiam convergir para qualquer ponto factível do MPEC. Mas, motivados pela condição AGP nos propusemos a verificar o comportamento dos algoritmos de PNL quando aplicados ao MPEC. Agora, com os dados experimentais temos melhores condições de discutir a aplicabilidade dos algoritmos de PNL ao MPEC.

Como foram realizados 20 testes para cada problema num total de 15 então temos um total de 300 testes para cada estratégia. Uma quantidade razoável de testes para podermos tirar algumas conclusões do trabalho. Os gráficos buscaram ilustrar, por meio de porcentagens, a quantidade de testes em que ocorreram cada uma das situações S1, S2 e S3 neste total de 300 testes. Com estes gráficos objetivamos, principalmente, analisar o desempenho de cada estratégia na resolução do problema MPEC relacionando o sucesso ou fracasso de cada uma delas à porcentagem de SO.

Assim, analisando os dados das tabelas e dos gráficos, podemos concluir que os algoritmos de lagrangeano aumentado (EASY) e os algoritmos de restauração inexata (IR) podem ser aplicados com sucesso na resolução do problema MPEC (dado na sua forma padrão), já que, ambas as estratégias encontraram as soluções dos problemas em mais de 90% dos 300 testes realizados. Os resultados do IMSL, apesar de não serem tão expressivos quanto do EASY e do IR, também podem ser considerados bons já que as soluções dos problemas foram encontradas em 60% dos casos.

O que reforça ainda mais o fato de que podemos aplicar com sucesso os algoritmos de PNL (em que temos a garantia de convergência para pontos AGP) ao MPEC (na forma padrão) são os resultados obtidos pelas estratégias da soma, mesmo em casos não-degenerados. Se a falta de regularidade tivesse tanta in-

fluência no comportamento destes algoritmos então a estratégia da soma deveria levar certa vantagem sobre as estratégias que resolvem os problemas dados na sua forma padrão. Mas, comparando os resultados do EASY com E-SOMA e do IMSL com I-SOMA podemos observar que esta vantagem não existiu. Em outras palavras, a condição de otimalidade AGP está suprimindo a não-regularidade dos pontos factíveis.

Conclusão

Discutimos este problema induzidos pela não-regularidade dos pontos factíveis do MPEC. Todos os pontos factíveis do MPEC satisfazem a condição de Fritz-John. Consequentemente, neste caso, a condição de Fritz-John é completamente inútil. Entretanto, ressaltamos que muitos algoritmos de PNL tem a propriedade de convergir para pontos que satisfazem uma condição de otimalidade mais forte do que a condição de Fritz-John, denominada condição AGP. O problema MPEC é um bom exemplo para diferenciar a condição de Fritz-John da condição AGP. Em síntese, a aplicabilidade da condição AGP ao problema MPEC diz que o conjunto dos pontos AGP para o problema MPEC também satisfazem a condição AGP para um problema associado onde as restrições de positividade são eliminadas. Neste novo problema nem todos os pontos factíveis satisfazem a condição de Fritz-John. Estas foram as principais razões que nos conduziram a esperar que a aplicação dos algoritmos de PNL que convergem para pontos AGP ao problema MPEC tivessem bons resultados.

Os experimentos numéricos corresponderam as nossas expectativas confirmando que, de fato, os algoritmos de PNL em que temos a garantia de convergência para pontos AGP podem ser aplicados com sucesso ao problema MPEC. Além disso, os resultados não muito expressivos da estratégia da soma reforçam este fato pois, se a falta de regularidade tivesse tanta influência no comportamento dos algoritmos utilizados, então a estratégia da soma deveria levar certa vantagem sobre as estratégias que resolvem os problemas dados na sua forma padrão, principalmente, nos casos não-degenerados. Mas, na prática, esta vantagem não existiu. Uma explicação simples e resumida de tudo isto está centrada na condição AGP que, por ser uma condição de otimalidade mais forte que a

condição de Fritz-John, está limitando o conjunto dos pontos candidatos a serem pontos limites dos algoritmos fazendo com que o fato de todos os pontos factíveis do MPEC serem não-regulares não represente uma ameaça à aplicabilidade dos algoritmos de PNL ao MPEC.

Referências Bibliográficas

- [1] Andreani, R., Friedlander, A., Martínez, J.M. *On the solution of finite-dimensional variational inequalities using smooth optimization with simple bounds*. Journal on Optimization Theory and Applications 94, pp. 635-657, (1997).
- [2] Andreani, R. and Martínez, J.M. *Solving complementarity problems by means of a new smooth constrained nonlinear solver*. In Reformulation - Nonsmooth, Piecewise Smooth, Semismooth and Smoothing methods. Edited by M. Fukushima and L. Qi. Kluwer, pp.1-24, (1999).
- [3] Andreani, R. and Martínez, J.M. *On the reformulation of nonlinear complementarity problems using the Fischer-Burmeister function*. In Applied Mathematics Letters . vol 12, (5), pp 7-12, (1999).
- [4] Andreani, R. and Martínez, J.M. *On the solution of bounded and unbounded mixed complementarity problems*. In Optimization . vol 50, (3-4), pp 265-278, (2001).
- [5] Andreani, R. and Martínez, J.M. *On the solution of mathematical programming problems with equilibrium constraints*. Mathematical Methods of Operations Research 54, pp. 345-358, (2001).
- [6] Bertsekas, D.P. *Nonlinear Programming*, 2nd Edition, Athena Scientific, Belmont, Massachusetts, (1999).

-
- [7] Diniz-Ehrhardt, M.A., Gomes-Ruggiero, M.A., Martínez, J.M. and Santos, S.A. *Augmented Lagrangian algorithms based on the spectral projected gradient for solving nonlinear programming problems*. To appear in Journal of Optimization Theory and Applications, (2000).
 - [8] Facchinei, F., Jiang, H. and Qi, L. *A smoothing method for mathematical programs with equilibrium constraints*. Mathematical Programming 85, pp. 107-134, (1999).
 - [9] Harker, P.T. and Pang, J.S. *Finite-dimensional variational inequality and nonlinear complementarity problem: a survey of theory, algorithms and applications*. Mathematical Programming 48, pp. 161-220, (1990).
 - [10] Haslinger, J. and Neittaanmäki, P. *Finite element approximation for optimal shape design: theory and applications*. Wiley, Chichester, (1998).
 - [11] Krejic, N., Martínez, J.M., Mello, M.P. and Pilotta, E.A. *Validation of an augmented Lagrangian algorithm with a Gauss-Newton Hessian approximation using a set of hard-spheres problems*. Computational Optimization and Applications 16, 247-263 (2000).
 - [12] Luenberger, D.G. *Linear and Nonlinear Programming*, Addison Wesley, (1986).
 - [13] Mangasarian, O.L. and Fromovitz, S. *Fritz-John necessary optimality conditions in presence of equality and inequality constraints*. Journal of Mathematical Analysis and Applications 17, pp.37, (1967).
 - [14] Marcote, P. *Network design problem with congestion effects: a class of bilevel programming*. Mathematical Programming 34, pp. 142-162, (1986).
 - [15] Martínez, J.M. *Two-phase model algorithm with global convergence for nonlinear programming*. Journal of Optimization Theory and Applications 96, pp. 397-436 (1998).

-
- [16] Martínez, J.M. *BOX-QUACAN and the implementation of Augmented Lagrangian algorithms for minimization with inequality constraints*. Computational and Applied Mathematics 19, pp. 31-56 (2000).
- [17] Martínez, J.M. *Inexact restoration method with Lagrangian tangent decrease and new merit function for nonlinear programming*. Journal of Optimization Theory and Applications 111, pp. 39-58 (2001).
- [18] Martínez, J.M. and Pilotta, E.A. *Inexact restoration algorithms for constrained optimization*. Journal of Optimization Theory and Applications 104, pp. 135-163 (2000).
- [19] Martínez, J.M. and Pilotta E.A. *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).
- [20] Martínez, J.M. and Santos, S.A. *Métodos Computacionais de Otimização*. 22o. Colóquio Brasileiro de Matemática, Impa - Rio de Janeiro, (1999).
- [21] Martínez, J.M. and Svaiter, B.F. *A practical optimality condition without constraint qualifications for nonlinear programming*. To appear in Journal of Optimization Theory and Applications, (2002).
- [22] Murphy, F.H., Sherali, H.D. and Soyster, A.L. *A mathematical programming approach for determining oligopolistic market equilibrium*. Mathematical Programming 24, pp. 92-106.
- [23] Spelluci, P. *An SQP method for general nonlinear programs using only equality constrained subproblems*, Mathematical Programming 82, pp. 413-447, (1998).