

Universidade Federal de Juiz de Fora
Instituto de Ciências Exatas
Programa de Pós-Graduação em Matemática

Mario Octavio Vera Campoverde

**Estudo de um Caso de Equilíbrio de Nash Usando Técnicas de Inequação
Variacional, Problemas de Complementaridade e Teoria de Otimização**

Juiz de Fora

2017

Mario Octavio Vera Campoverde

Estudo de um Caso de Equilíbrio de Nash Usando Técnicas de Inequação Variacional, Problemas de Complementaridade e Teoria de Otimização

Dissertação apresentada ao Programa de Pós-Graduação em Matemática da Universidade Federal de Juiz de Fora, na área de concentração em Matemática , como requisito parcial para obtenção do título de Mestre em Matemática.

Orientador: Sandro Rodrigues Mazorche

Juiz de Fora

2017

Ficha catalográfica elaborada através do Modelo Latex do CDC da UFJF
com os dados fornecidos pelo(a) autor(a)

Vera Campoverde, Mario.

Estudo de um Caso de Equilíbrio de Nash Usando Técnicas de Inequação Variacional, Problemas de Complementaridade e Teoria de Otimização / Mario Octavio Vera Campoverde. – 2017.

95 f. : il.

Orientador: Sandro Rodrigues Mazorche

Dissertação (Mestrado) – Universidade Federal de Juiz de Fora, Instituto de Ciências Exatas. Programa de Pós-Graduação em Matemática, 2017.

1. Inequação Variacional. 2. Problema de Complementaridade. 3. Equilíbrio de Nash. I. Rodrigues Mazorche, Sandro, orient. II. Título.

Mario Octavio Vera Campoverde

Estudo de um Caso de Equilíbrio de Nash Usando Técnicas de Inequação Variacional, Problemas de Complementaridade e Teoria de Otimização

Dissertação apresentada ao Programa de Pós-Graduação em Matemática da Universidade Federal de Juiz de Fora, na área de concentração em Matemática, como requisito parcial para obtenção do título de Mestre em Matemática.

Aprovada em:

BANCA EXAMINADORA

Prof. Dr. Sandro Rodrigues Mazorche - Orientador
Universidade Federal de Juiz de Fora

Professor Dr. Wilhelm Passarella Freire
Universidade Federal de Juiz de Fora

Professor Dr. Dênis Emanuel da Costa Vargas
IF SUDESTE MG Campus Rio Pomba

Dedico este trabalho a minha mãe, pois ela é a fonte de minha inspiração e motivação constante.

AGRADECIMENTOS

Agradeço ao professor Sandro Rodrigues por dedicar seu tempo, acreditar em mim e ter paciência para me orientar na realização deste trabalho, muito obrigado.

A minha querida família, mãe, pai, irmãs e sobrinhas pelo amor incondicional, seu apoio, confiança e incentivo.

Agradeço também ao corpo docente do departamento de Matemática da UFJF que contribuíram para que possa crescer como profissional.

Agradeço a meus amigos e colegas do mestrado por compartilhar horas de estudo comigo, os momentos de diversão, as conversas e troca de ideias e por fazer minha vida mais confortável. Muito obrigado a todas as pessoas que colaboraram me incentivando, seja direta ou indiretamente, obrigado pela experiência. Quero fazer um agradecimento especial a Pablo Rojas Oviedo por me ajudar com as ilustrações.

Agradeço a CAPES pelo apoio financeiro.

“Vale mas tener cicatrices por valiente, que la piel intacta por cobarde”

Bruce Lee.

RESUMO

Neste trabalho realizamos um estudo teórico das inequações variacionais (VI), mostrando condições necessárias e suficientes para a existência de soluções, assim mesmo apresentamos algumas classificações para vários casos especiais de VI, explicando a interconexão entre a VI e o problema de complementaridade (CP), bem como sua relação com um programa de otimização não linear e a teoria de jogos de estratégias que esta relacionada com os problemas de equilíbrio de Nash. Estabelecemos alguns resultados de equivalência entre uma VI e um CP, mostramos por exemplo que resolver uma VI associada a uma função definida num conjunto que tem uma determinada estrutura é equivalente a resolver um problema de complementaridade mista (MiCP); assim como também, vimos que sob certas hipóteses, um ponto de equilíbrio de Nash é solução de uma determinada VI. Finalmente, aplicamos a teoria para o estudo de um caso de problema de equilíbrio de Nash, modelado via as condições de KKT como um MiCP, ou criando uma família de programas convexos, onde usei alguns algoritmos para determinar numericamente a solução do ponto de equilíbrio de Nash.

Palavras-chave: Inequação Variacional. Problema de Complementaridade. Equilíbrio de Nash.

ABSTRACT

In this work, we present a theoretical study of the variational inequalities (VI), showing necessary and sufficient conditions for the existence of solutions. We also present some classifications for several special cases of LV, explaining the interconnection between LV and the complementarity problem. As well as its relationship with a nonlinear optimization program and the strategy game theory that is related to Nash equilibrium problems. We have established some equivalence results between a VI and a CP, for example we have shown that solving a VI associated with a function defined in a set having a given structure is equivalent to solving a mixed complementarity problem (MiCP); As well as, we have seen that under certain hypotheses, a Nash equilibrium is the solution of a given VI. Finally, we apply the theory to the study of a Nash equilibrium problem case, modeled via the KKT conditions as an MiCP, or creating a family of convex programs, where used some algorithms to numerically determine the equilibrium point solution of Nash.

LISTA DE ILUSTRAÇÕES

Figura 1 – $\max f(x) = -\min(-f(x))$	15
Figura 2 – As direções d^1 e d^2 são viáveis em relação ao conjunto K	17
Figura 3 – Exemplos de cones.	17
Figura 4 – O cone tangente de Bouligand $\mathcal{B}(\bar{x}, K)$	20
Figura 5 – O cone dual do cone $\mathcal{B}(\bar{x}, K)$ e as condições de otimalidade (2.10) e (2.11).	21
Figura 6 – Cone normal e o dual de cone tangente para un conjunto convexo.	29
Figura 7 – O cone $\mathcal{L}(\bar{x}, K)$ obtido pela linearização das restrições das desigualdades ativas em \bar{x}	35
Figura 8 – A direção \bar{d} satisfaz a condição de Mangasarian-Fromovitz.	37
Figura 9 – Interpretação geométrica das condições de KKT.	40
Figura 10 – Representação dos mínimos de f do exemplo 2.2.	41
Figura 11 – O cone normal e solução de uma VI.	57
Figura 12 – O Cone Tangente.	68
Figura 13 – Condições KKT de uma VI.	71

LISTA DE ABREVIATURAS E SIGLAS

VI	Inequação Variacional
QVI	Inequação Quase Variacional
AVI	Inequação Variacional Afim
CP	Problema de Complementaridade
LCP	Problema de Complementaridade Linear
NCP	Problema de Complementaridade Não Linear
MiCP	Problema de Complementaridade Mista
MLCP	Problema de Complementaridade Linear Mista
KKT	Karush Kuhn Tucker
CQ	Condição de Qualificação
PNL	Programa não Linear Padrão
EP	Problema de Equilíbrio

SUMÁRIO

1	INTRODUÇÃO	12
2	CONCEITOS PRELIMINARES	14
2.1	Noções de otimização	14
2.1.1	Definições e alguns fatos básicos	14
2.1.2	Condições de otimalidade para problemas com restrições.	16
2.1.3	Condições de otimalidade no caso de restrições de igualdade	23
2.1.4	Elementos de convexidade, o problema de minimização convexo	27
2.1.5	Problemas com restrições de igualdade e desigualdade	34
2.2	Introdução as inequações variacionais	40
2.2.1	Exemplos de Inequações variacionais	40
2.2.2	Teoremas sobre ponto fixo	42
2.2.3	A caracterização da projeção sobre um conjunto convexo	43
2.2.4	Inequações variacionais em \mathbb{R}^n	45
2.2.5	Problema que conduzem a uma inequação variacional	48
2.3	Introdução à teoria dos jogos	48
2.3.1	Jogos de estratégias	48
2.3.2	Equilíbrios de Nash	52
3	INEQUAÇÕES VARIACIONAIS E PROBLEMAS DE COM- PLEMENTARIDADE	56
3.1	Descrição do Problema	56
3.1.1	Problemas Afins	60
3.2	Relações entre classes de problemas	60
3.3	Integrabilidade e o sistema KKT	66
3.3.1	Problemas de otimização restrita	66
3.3.2	O Sistema Karush - Kuhn - Tucker	70
3.4	Problemas de equilíbrio de Nash	72
4	UM ESTUDO DE UM CASO DE EQUILÍBRIO DE NASH .	74
4.1	Definição do problema	74
4.2	Desenvolvimento matemático	75
4.3	Exemplo de equilíbrio do modelo de Nash - Cournot	91
5	CONCLUSÕES E TRABALHOS FUTUROS	93

REFERÊNCIAS	94
-----------------------	----

1 INTRODUÇÃO

As inequações variacionais tem sua origem no cálculo das variações associadas à minimização dos funcionais de dimensões infinita. O estudo sistemático começou no início da década de 1960 com o trabalho do matemático italiano Guido Stampacchia e seus colaboradores, que usaram a inequação variacional (VI) como ferramenta analítica para estudar problemas de fronteira livre definidos por operadores dielétricos parciais não-lineares decorrentes de problemas unilaterais em teoria da elasticidade e da plasticidade e na mecânica. Alguns dos primeiros trabalhos sobre inequações variacionais são [15], [16], [22], [23]. Em particular, o primeiro teorema da existência e unicidade da solução das VIs foi provado em [22].

O problema de complementaridade não linear (NCP) é um sistema de desigualdades finitas não-lineares em muitas variáveis finitas não-negativas, juntamente com uma equação especial que expressa a relação de complementaridade entre as variáveis e desigualdades correspondentes. Esta condição de complementaridade é a característica principal que distingue o NCP de um sistema de desigualdade geral. A inequação (ou desigualdade) variacional finita-dimensional, que como veremos, é uma generalização do NCP, fornece uma ampla configuração unificadora para o estudo de otimização e problemas de equilíbrio e serve como o principal quadro computacional para a solução prática de uma série de problemas na ciências matemáticas. O estudo do NCP e da VI finito-dimensional também começou no início da década de 1960; e em um período de quatro décadas, tornou-se uma disciplina muito frutífera no campo da programação matemática. Os desenvolvimentos incluem uma rica teoria matemática, uma série de algoritmos de soluções eficazes, uma infinidade de conexões interessantes com inúmeras disciplinas e uma ampla gama de aplicações importantes em engenharia e economia. Como resultado de suas amplas associações, a literatura da VI/CP beneficiou-se das contribuições feitas por matemáticos (puros, aplicados e computacionais), cientistas da computação, engenheiros de vários tipos (civis, químicos, elétricos, mecânicos e sistemas) e economistas de diversas especialidades (agrícola, computacional, energética, financeira e espacial).

Este trabalho apresenta um estudo teórico das inequações variacionais e problemas de complementaridade de dimensão finita, e sua aplicação num problema de equilíbrio de Nash. O trabalho está dividido da seguinte forma.

No capítulo 2, apresentamos algumas noções de otimização (ver seção 2.1), estudaremos por exemplo as condições de Karush - Kuhn - Tucker (KKT) para um problema de minimização com restrições de igualdade e desigualdade, na seção 2.2 introduzimos a definição de inequação variacional e damos alguns exemplos; para finalizar o capítulo, na seção 2.3 fazemos uma introdução à teoria dos jogos de estratégias onde apresentamos a definição de equilíbrio de Nash.

No capítulo 3 estudamos a interconexão entre as inequações variacionais e os casos especiais

de problemas de complementaridade, onde usamos as condições KKT para mostrar que sob certas hipóteses, resolver uma inequação variacional (VI) é equivalente a resolver um problema de complementaridade mista (MiCP). Assim como também, estabelecemos (ver proposição 3.4) a relação entre uma VI e os equilíbrios de Nash.

No capítulo 4 fornecemos um estudo de uma aplicação de um caso de equilíbrio de Nash, chamado o modelo de Nash-Cournot, onde dadas N empresas que fornecem um produto homogêneo de uma forma não cooperativa, o objetivo é determinar um ponto de equilíbrio de Nash que corresponde a um conjunto de níveis de produção para as N empresas, modelando o problema via as condições de KKT como um MiCP, como um NCP; ou, conforme o artigo [20], mediante a criação de uma família de programas de otimização não linear, onde usamos alguns algoritmos para exibir a solução de equilíbrio de Nash.

2 CONCEITOS PRELIMINARES

Neste capítulo seram apresentados alguns resultados que seram utilizados ao longo deste trabalho, na primeira seção veremos algumas noções de otimização como programação convexa e as condições de KKT para problemas de otimização com restrições mistas. Logo, introduzimos a definição de inequação variacional, estudando condições necessárias e suficientes para garantir a existencia de soluções; finalmente, veremos alguns resultados sobre teoria de jogos onde definimos os equilíbrios de Nash.

2.1 Noções de otimização

Nesta seção estudaremos alguns fatos básicos sobre os problemas de otimização, que serão necessários ao longo deste trabalho. Para maiores informações veja as referências [8], [21].

2.1.1 Definições e alguns fatos básicos

Sejam dados um conjunto $K \subseteq \mathbb{R}^n$ e uma função $f : \mathbb{R}^n \rightarrow \mathbb{R}$. O problema principal a ser considerado é o de achar um minimizador de f no conjunto K . Este problema será escrito como:

$$\begin{aligned} &\text{Minimizar} && f(x) \\ &\text{Sujeito a} && x \in K \end{aligned} \tag{2.1}$$

O conjunto K será chamado conjunto viável (factível) do problema, os pontos de K serão chamados pontos viáveis, e f será chamada função objetivo.

Definição 2.1. Dizemos que um ponto $\bar{x} \in K$ é:

a) minimizador global do problema (2.1), se

$$f(\bar{x}) \leq f(x) \quad \forall x \in K \tag{2.2}$$

b) minimizador local do problema (2.1), se existe $\delta > 0$ tal que:

$$f(\bar{x}) \leq f(x) \quad \forall x \in K \cap B[\bar{x}, \delta] \tag{2.3}$$

Pela definição, é claro que todo minimizador global também é minimizador local, mas não reciprocamente. Se para todo $x \neq \bar{x}$ a desigualdade (2.2) ou (2.3) é estrita, \bar{x} será chamado minimizador estrito de (2.1) (global ou local, respectivamente).

Definição 2.2. Dizemos que $\bar{v} \in [-\infty, +\infty)$ definido por:

$$\bar{v} = \inf_{x \in K} f(x)$$

é o valor ótimo do problema (2.1).

Uma função pode admitir vários minimizadores globais, mas o valor ótimo (global) do problema, naturalmente, sempre é o mesmo.

Quando existem uma vizinhança V de $\bar{x} \in K$ e um número $\beta > 0$ tais que

$$f(x) - f(\bar{x}) \geq \beta \|x - \bar{x}\| \quad \forall x \in K \cap V$$

dizemos que f satisfaz a condição de crescimento linear no conjunto K em torno de \bar{x} .

Dizemos que f satisfaz a condição de crescimento quadrático quando

$$f(x) - f(\bar{x}) \geq \beta \|x - \bar{x}\|^2 \quad \forall x \in K \cap V.$$

Observamos que qualquer uma destas duas condições implica em que \bar{x} é um minimizador local estrito do problema (2.1).

Não é difícil ver que qualquer problema de maximização

$$\begin{array}{ll} \text{Maximizar} & f(x) \\ \text{Sujeito a} & x \in K \end{array}$$

pode ser transformado em um problema de minimização equivalente:

$$\begin{array}{ll} \text{Minimizar} & -f(x) \\ \text{Sujeito a} & x \in K \end{array}$$

Em particular, as soluções locais e globais de ambos problemas são as mesmas, com sinais opostos para os valores ótimos.

Tipicamente, o conjunto viável de um problema é definido por um sistema de igualdades

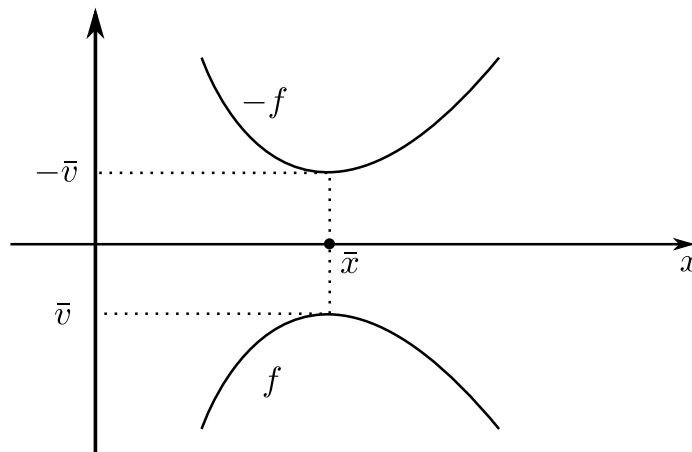


Figura 1 – $\max f(x) = -\min(-f(x))$.

e/ou desigualdades e/ou uma inclusão, como, por exemplo,

$$K = \{x \in \mathbb{R}^n : h_i(x) = 0, i = 1, \dots, l, g_j(x) \leq 0, j = 1, \dots, m\} \quad (2.4)$$

ou, em notação mais compacta,

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

onde $h : \mathbb{R}^n \rightarrow \mathbb{R}^l$ e $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$.

Quando $K = \mathbb{R}^n$, dizemos que o problema (2.1) é irrestrito, e quando $K \neq \mathbb{R}^n$ falamos de otimização com restrições.

Quando $l \neq 0$, $m = 0$, falamos sobre um problema com restrições de igualdade; finalmente, se $l \neq 0$, $m \neq 0$, falamos sobre um problema com restrições mistas.

Dizemos que um conjunto é poliedral quando ele pode ser representado como o conjunto das soluções de um sistema finito de equações e inequações lineares. Isto é:

$$K = \{x \in \mathbb{R}^n : Ax = b, Cx \leq d\} \quad (2.5)$$

onde $A \in \mathbb{R}^{l \times n}$, $C \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^l$ e $d \in \mathbb{R}^m$. Neste contexto, dizemos que $h : \mathbb{R}^n \rightarrow \mathbb{R}^l$, $h(x) = Ax - b$, é uma função afim. Assim como $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$, $g(x) = Cx - d$.

Uma função $f : \mathbb{R}^n \rightarrow \mathbb{R}$ definida por

$$f(x) = \langle Qx, x \rangle + \langle q, x \rangle \quad (2.6)$$

onde $Q \in \mathbb{R}^{n \times n}$, $q \in \mathbb{R}^n$, é chamada função quadrática. Neste contexto, podemos admitir (sem perda de generalidade) que a matriz Q é simétrica, já que

$$\langle Qx, x \rangle = \frac{1}{2} \langle (Q + Q^T)x, x \rangle.$$

Portanto, na definição de uma função quadrática sempre podemos substituir a matriz Q associada pela matriz simétrica $(Q + Q^T)/2$.

Uma classe especial de problemas de otimização se refere ao caso em que K é um conjunto poliedral. Se além disso f for quadrática, (2.1) se chama problema de programação quadrática, e se f for linear ($Q = 0$ em (2.6)), o problema é de programação linear.

2.1.2 Condições de otimalidade para problemas com restrições.

Agora consideramos um problema com restrições em formato geral

$$\begin{array}{ll} \text{Minimizar} & f(x) \\ \text{Sujeito a} & x \in K, \end{array} \quad (2.7)$$

onde K é um conjunto dado (não-vazio) cuja estrutura por enquanto não é especificada, e $f : \mathbb{R}^n \rightarrow \mathbb{R}$ é a função objetivo.

Definição 2.3. Dizemos que $d \in \mathbb{R}^n$ é uma direção viável em relação ao conjunto K no ponto $\bar{x} \in K$, quando existe $\epsilon > 0$ tal que

$$\bar{x} + td \in K \quad \forall t \in [0, \epsilon].$$

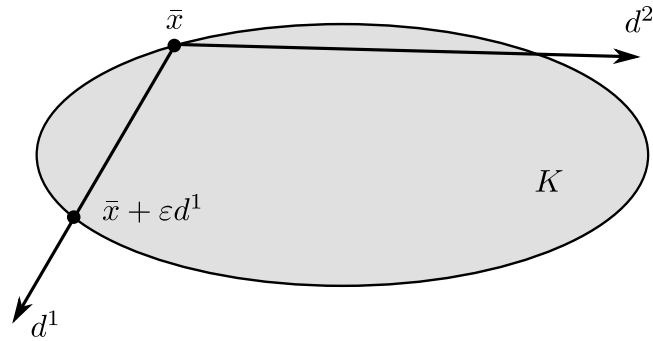


Figura 2 – As direções d^1 e d^2 são viáveis em relação ao conjunto K .

Denotamos por $\mathcal{V}(\bar{x}, K)$ o conjunto de todas as direções viáveis em relação ao conjunto K no ponto $\bar{x} \in K$.

Definição 2.4. Um conjunto $C \subseteq \mathbb{R}^n$ chama-se cone quando para qualquer $d \in C$ implica que $td \in C$ para todo $t \geq 0$.

Pela definição, se C é um cone não-vazio, necessariamente $0 \in C$. Alguns exemplos de cone são: o espaço \mathbb{R}^n , qualquer subespaço de \mathbb{R}^n , o ortante não-negativo \mathbb{R}_+^n . É fácil

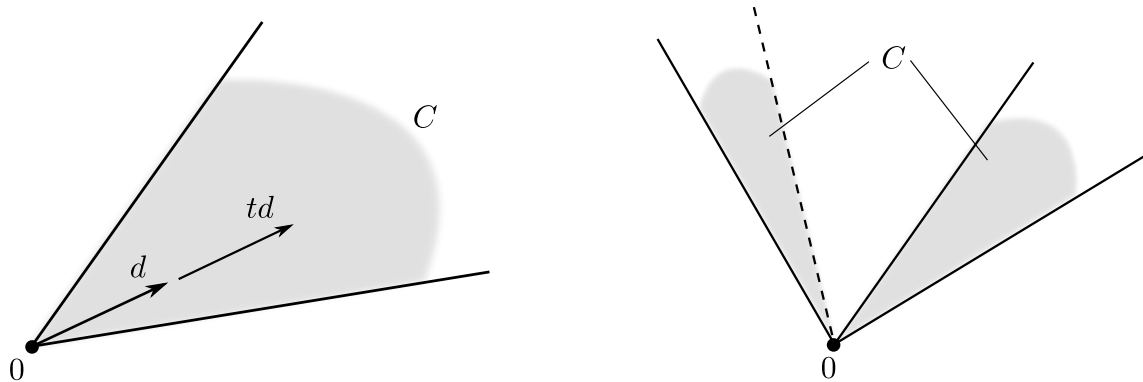


Figura 3 – Exemplos de cones.

ver que para $K \subseteq \mathbb{R}^n$ e $\bar{x} \in K$ quaisquer, o conjunto de direções viáveis $\mathcal{V}(\bar{x}, K)$ sempre é um cone, já que se $\bar{x} + td \in K$ para todo $t \in [0, \epsilon]$ então $\bar{x} + t(\alpha d) \in K$ para todo $\alpha > 0$ e todo $t \in [0, \epsilon/\alpha]$. Além disso, o cone $\mathcal{V}(\bar{x}, K)$ sempre é não-vazio (pelo menos, $0 \in \mathcal{V}(\bar{x}, K)$).

Em relação a comportamento da função objetivo, a seguinte noção é fundamental.

Definição 2.5. Dizemos que $d \in \mathbb{R}^n$ é uma direção de descida da função $f : \mathbb{R}^n \rightarrow \mathbb{R}$ no ponto $\bar{x} \in \mathbb{R}^n$, se existe $\epsilon > 0$ tal que

$$f(\bar{x} + td) < f(\bar{x}) \quad \forall t \in (0, \epsilon].$$

Denotamos por $\mathcal{D}(\bar{x}, f)$ o conjunto de todas as direções de descida da função f no ponto \bar{x} . É claro que o conjunto $\mathcal{D}(\bar{x}, f)$ pode ser vazio (por exemplo, este é o caso

quando \bar{x} é um minimizador irrestrito de f , mesmo que seja local). Quando $\mathcal{D}(\bar{x}, f)$ é não vazio, ele não é cone (porque $0 \notin \mathcal{D}(\bar{x}, f)$). No entanto, o conjunto $\mathcal{D}(\bar{x}, f) \cup \{0\}$ é um cone não-vazio.

Lema 2.1. *Seja $f : \mathbb{R}^n \rightarrow \mathbb{R}$ uma função diferenciável no ponto $\bar{x} \in \mathbb{R}^n$. Então,*

- (a) *Para todo $d \in \mathcal{D}(\bar{x}, f)$, tem-se que $\langle \nabla f(\bar{x}), d \rangle \leq 0$.*
- (b) *Se $d \in \mathbb{R}^n$ satisfaz $\langle \nabla f(\bar{x}), d \rangle < 0$, tem-se que $d \in \mathcal{D}(\bar{x}, f)$.*

Demonstração. A demonstração pode ser encontrada em [8]. □

Se $\bar{x} \in K$ é um minimizador local do problema (2.7), é obvio que não pode existir uma direção de descida de f no ponto \bar{x} que ao mesmo tempo seja viável em relação ao conjunto K , isto é, necessariamente temos

$$\mathcal{D}(\bar{x}, f) \cap \mathcal{V}(\bar{x}, K) = \emptyset.$$

Pelo lema 2.1, temos então que

$$\langle \nabla f(\bar{x}), d \rangle \geq 0 \quad \forall d \in \mathcal{V}(\bar{x}, K).$$

No entanto, esta condição (mesmo sendo matematicamente correta) nem sempre fornece informação útil, porque o conjunto de direções viáveis $\mathcal{V}(\bar{x}, K)$ pode conter só o vetor nulo em muitas situações naturais (e neste caso a condição acima vale trivialmente, isto é, independentemente da função f e da otimalidade ou não do ponto \bar{x}). Para obter uma descrição mais informativa, precisamos considerar as direções tangentes, que introduzimos a seguir. Definimos a distancia entre um ponto $x \in \mathbb{R}^n$ e o conjunto $K \subseteq \mathbb{R}^n$ como

$$\text{dist}(x, K) = \inf_{y \in K} \|x - y\|$$

Definição 2.6. Dizemos que $d \in \mathbb{R}^n$ é uma direção tangente em relação ao conjunto K no ponto $\bar{x} \in K$ quando

$$\text{dist}(\bar{x} + td, K) = o(t), \quad t \in \mathbb{R}_+.$$

Denotamos por $\mathcal{T}(\bar{x}, K)$ o conjunto de todas as direções tangentes ao conjunto K no ponto $\bar{x} \in K$. Claramente,

$$\mathcal{V}(\bar{x}, K) \subseteq \mathcal{T}(\bar{x}, K), \tag{2.8}$$

isto é, todas as direções viáveis são tangentes (para $d \in \mathcal{V}(\bar{x}, K)$ tem-se que $\text{dist}(\bar{x} + td, K) = 0$ para todo $t > 0$ suficientemente pequeno), mas não reciprocamente. Também é claro que

o conjunto $\mathcal{T}(\bar{x}, K)$ é um cone, chamado o cone tangente em relação ao conjunto K no ponto \bar{x} . Observamos que, de forma equivalente, o cone tangente pode ser definido como

$$\mathcal{T}(\bar{x}, K) = \left\{ d \in \mathbb{R}^n : \begin{array}{l} \forall (t_k) \subseteq \mathbb{R}_+, \lim_{k \rightarrow +\infty} t_k = 0, \\ \exists (d_k) \subseteq \mathbb{R}^n, \lim_{k \rightarrow +\infty} d_k = d, \text{ tal que} \\ \bar{x} + t_k d_k \in K \text{ para todo } k \in \mathbb{N} \end{array} \right\}.$$

Para ver isto, notamos que para toda $(t_k) \subseteq \mathbb{R}_+, \lim_{k \rightarrow +\infty} t_k = 0$, temos

$$\begin{aligned} \text{dist}(\bar{x} + t_k d, K) &\leq \|(\bar{x} + t_k d) - (\bar{x} + t_k d_k)\| \\ &= t_k \|d_k - d\| \\ &= o(t_k). \end{aligned}$$

Outra noção útil (um pouco mais geral do que a do cone tangente) é a do cone tangente de Bouligand:

Definição 2.7. O cone tangente de Bouligand em relação ao conjunto K no ponto $\bar{x} \in K$, denotado por $\mathcal{B}(\bar{x}, K)$, consiste de todas as direções $d \in \mathbb{R}^n$, chamadas direções tangentes a K no ponto \bar{x} , para as quais existe uma seqüência de vetores $(x_k) \subseteq K$ e uma seqüência de escalares positivos $(t_k) \subseteq \mathbb{R}$ tais que:

$$\lim_{k \rightarrow \infty} x_k = \bar{x}, \quad \lim_{k \rightarrow \infty} t_k = 0, \quad \text{y} \quad \lim_{k \rightarrow \infty} \frac{x_k - \bar{x}}{t_k} = d.$$

Ou, de forma equivalente, ver [8],

$$\mathcal{B}(\bar{x}, K) = \left\{ d \in \mathbb{R}^n : \begin{array}{l} \exists (t_k) \subseteq \mathbb{R}_+ \setminus \{0\}, \lim_{k \rightarrow +\infty} t_k = 0, \\ \exists (d_k) \subseteq \mathbb{R}^n, \lim_{k \rightarrow +\infty} d_k = d, \text{ tais que} \\ \bar{x} + t_k d_k \in K \text{ para todo } k \in \mathbb{N} \end{array} \right\}.$$

Comparando as definições acima, temos que, em geral,

$$\mathcal{T}(\bar{x}, K) \subseteq \mathcal{B}(\bar{x}, K). \quad (2.9)$$

Como vamos ver mais adiante, quando o conjunto K é definido por restrições funcionais é possível dar uma descrição construtiva (algébrica) das direções tangentes, quase sempre os cones tangente e de Bouligand coincidem. Observamos que o caso em que \bar{x} está no interior do conjunto K é o caso fácil de analisar: é obvio que

$$\bar{x} \in \text{int}K \quad \Rightarrow \quad \mathcal{T}(\bar{x}, K) = \mathcal{B}(\bar{x}, K) = \mathcal{V}(\bar{x}, K) = \mathbb{R}^n.$$

Também é claro que quando \bar{x} é um ponto isolado do conjunto K , tem-se que

$$\mathcal{T}(\bar{x}, K) = \mathcal{B}(\bar{x}, K) = \mathcal{V}(\bar{x}, K) = \{0\}.$$

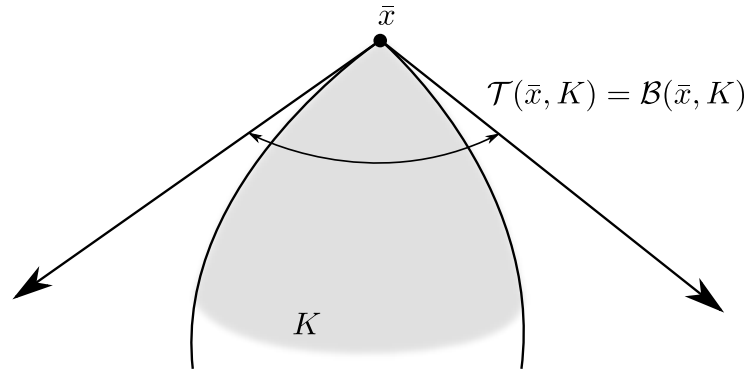


Figura 4 – O cone tangente de Bouligand $\mathcal{B}(\bar{x}, K)$.

A importancia do cone tangente de Bouligand deve-se ao fato de que é ele que aparece de maneira mais natural no desenvolvimento de condições necessárias e suficientes de otimalidade de primeira ordem, como veremos mais adiante. No entanto, o cone $\mathcal{T}(\bar{x}, K)$ é mais facil de se visualizar.

O resultado a seguir (ver [8]) mostra o papel do cone $\mathcal{B}(\bar{x}, K)$ nas condições de otimalidade.

Teorema 2.1. *Sejam $K \subseteq \mathbb{R}^n$ e $f : \mathbb{R}^n \rightarrow \mathbb{R}$ uma função diferenciável no ponto $\bar{x} \in K$. Se \bar{x} é um minimizador local do problema (2.7). Então,*

$$\langle \nabla f(\bar{x}), d \rangle \geq 0 \quad \forall d \in \mathcal{B}(\bar{x}, K). \quad (2.10)$$

Demonstração. Para $d = 0 \in \mathcal{B}(\bar{x}, K)$, a condição (2.10) vale trivialmente. Fixemos $d \in \mathcal{B}(\bar{x}, K) \setminus \{0\}$ arbitrário, pela definição do cone tangente de Bouligand existem seqüências $(t_k) \subseteq \mathbb{R}_+ \setminus \{0\}$ e $(d_k) \subseteq \mathbb{R}^n$ tais que $\lim_{k \rightarrow +\infty} t_k = 0$, $\lim_{k \rightarrow +\infty} d_k = d$ e $\bar{x} + t_k d_k \in K$ para todo $k \in \mathbb{N}$.

Como \bar{x} é um minimizador local de (2.7), $\bar{x} + t_k d_k \in K$ e $\lim_{k \rightarrow +\infty} (\bar{x} + t_k d_k) = \bar{x}$, para todo k suficientemente grande e pela diferenciabilidade de f no ponto \bar{x} temos que

$$\begin{aligned} 0 &\leq f(\bar{x} + t_k d_k) - f(\bar{x}) \\ &= t_k \langle \nabla f(\bar{x}), d_k \rangle + o(t_k \|d_k\|) \\ &= t_k \langle \nabla f(\bar{x}), d_k \rangle + o(t_k). \end{aligned}$$

Dividindo os dois lados da desigualdade acima por $t_k > 0$ e passando ao limite quando $k \rightarrow +\infty$, obtemos (2.10). \square

Observamos que se \bar{x} é um minimizador local de f em K , (2.10) implica a condição

$$\langle \nabla f(\bar{x}), d \rangle \geq 0 \quad \forall d \in \mathcal{T}(\bar{x}, K),$$

já que $\mathcal{T}(\bar{x}, K) \subseteq \mathcal{B}(\bar{x}, K)$. No entanto, esta condição é mais fraca, considerando que nem sempre os dois cones são iguais.

Definição 2.8. Dizemos que um ponto $\bar{x} \in \mathbb{R}^n$ é ponto estacionário do problema (2.7) se $\bar{x} \in K$ e vale (2.10).

O teorema 2.1 mostra que para um ponto ser uma solução local do problema (2.7), ele tem que ser estacionário no sentido da condição (2.10).

A condição de otimalidade no teorema 2.1 se chama condição primal, porque ela é baseada nas variáveis "originais" (também chamadas primais) do problema (2.7), isto é, as variáveis $x \in \mathbb{R}^n$. A condição de otimalidade em questão pode ser transformada em forma "primal-dual" utilizando a noção seguinte, ver [8].

Definição 2.9. O cone dual (às vezes, este cone também é chamado cone polar) de um cone $C \subseteq \mathbb{R}^n$ é definido por

$$C^* = \{y \in \mathbb{R}^n : \langle y, d \rangle \leq 0 \quad \forall d \in C\}.$$

Usando a noção de cone dual, a condição de otimalidade (2.10) é equivalente a

$$-\nabla f(\bar{x}) \in (\mathcal{B}(\bar{x}, K))^*. \quad (2.11)$$

Como veremos mais adiante, o cálculo do cone $(\mathcal{B}(\bar{x}, K))^*$ para conjuntos K com uma

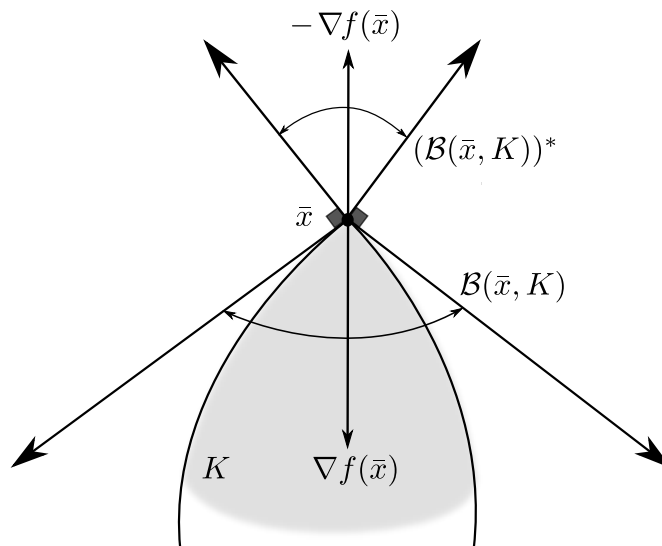


Figura 5 – O cone dual do cone $\mathcal{B}(\bar{x}, K)$ e as condições de otimalidade (2.10) e (2.11).

estrutura mais concreta faz uso de variáveis auxiliares chamadas variáveis duais. Por isso, a condição de otimalidade (2.11), bem como suas conseqüências, se chamam condições primais-duais.

O teorema 2.2 a seguir, que pode ser encontrado na referência [8], de novo mostra o papel importante do cone $\mathcal{B}(\bar{x}, K)$. De fato, neste resultado o cone $\mathcal{B}(\bar{x}, K)$ não pode ser substituído por $\mathcal{T}(\bar{x}, K)$.

Teorema 2.2. *Sejam $K \subseteq \mathbb{R}^n$ e $f : \mathbb{R}^n \rightarrow \mathbb{R}$ uma função diferenciável no ponto $\bar{x} \in K$. Então,*

$$\langle \nabla f(\bar{x}), d \rangle > 0 \quad \forall d \in \mathcal{B}(\bar{x}, K) \setminus \{0\} \quad (2.12)$$

se, e somente se, f satisfaz no conjunto K em torno de \bar{x} a condição de crescimento linear: existem uma vizinhança V de \bar{x} e um número $\beta > 0$ tais que

$$f(x) - f(\bar{x}) \geq \beta \|x - \bar{x}\| \quad \forall x \in K \cap V \quad (2.13)$$

Em particular, (2.13) implica que \bar{x} é um minimizador local estrito do problema (2.7).

Demonstração. Suponhamos que (2.12) valha, mas não (2.13). Então, existe uma seqüência $(x_k) \subseteq K \setminus \{\bar{x}\}$ tal que $\lim_{k \rightarrow +\infty} x_k = \bar{x}$ e

$$f(x_k) - f(\bar{x}) < \frac{1}{k} \|x_k - \bar{x}\| \quad \forall k \in \mathbb{N}.$$

Definimos $t_k = \|x_k - \bar{x}\| > 0$ e $d_k = (x_k - \bar{x}) / \|x_k - \bar{x}\|$. Temos que $\bar{x} + t_k d_k = x_k \in K$ para todo $k \in \mathbb{N}$. Como (d_k) é limitada, podemos supor (tomando uma subsequência, se for necessário) que ela converge a $d \in \mathbb{R}^n$, $\|d\| = 1$. Pela definição de cone de Bouligand $\mathcal{B}(\bar{x}, K)$, temos que $d \in \mathcal{B}(\bar{x}, K) \setminus \{0\}$. Logo, para todo $k \in \mathbb{N}$, temos

$$\begin{aligned} \frac{t_k}{k} \|d_k\| &= \frac{1}{k} \|x_k - \bar{x}\| \\ &> f(x_k) - f(\bar{x}) \\ &= \langle \nabla f(\bar{x}), x_k - \bar{x} \rangle + o(\|x_k - \bar{x}\|) \\ &= t_k \langle \nabla f(\bar{x}), d_k \rangle + o(t_k). \end{aligned}$$

Dividindo os dois lados por $t_k > 0$ e passando ao limite quando $k \rightarrow +\infty$, obtemos que

$$0 \geq \langle \nabla f(\bar{x}), d \rangle,$$

em contradição com (2.12).

Suponhamos agora que valha a condição de crescimento linear (2.13). Fixemos $d \in \mathcal{B}(\bar{x}, K) \setminus \{0\}$ arbitrário e as seqüências associadas $(t_k) \subseteq \mathbb{R}_+ \setminus \{0\}$ e $(d_k) \subseteq \mathbb{R}^n$ tais que $\lim_{k \rightarrow +\infty} t_k = 0$, $\lim_{k \rightarrow +\infty} d_k = d$ e $\bar{x} + t_k d_k \in K$ para todo $k \in \mathbb{N}$.

Para todo k suficientemente grande, $\bar{x} + t_k d_k \in K \cap V$. Portanto, utilizando (2.13), obtemos que

$$\begin{aligned} \beta t_k \|d_k\| &= \beta \|(\bar{x} + t_k d_k) - \bar{x}\| \\ &\leq f(\bar{x} + t_k d_k) - f(\bar{x}) \\ &= t_k \langle \nabla f(\bar{x}), d_k \rangle + o(t_k \|d_k\|). \end{aligned}$$

Dividindo os dois lados desta desigualdade por $t_k > 0$ e tomando o limite quando $k \rightarrow +\infty$, obtemos que

$$0 < \beta \|d\| \leq \langle \nabla f(\bar{x}), d \rangle,$$

o que prova (2.12). □

Cabe ressaltar que a condição (2.12) só pode ser satisfeita quando o cone $\mathcal{B}(\bar{x}, K)$ não contém subespaços (fora do subespaço trivial $\{0\}$). Este fato indica as limitações de aplicação do teorema 2.2, o que vai ficar bem claro mais adiante.

A questão do cálculo dos cones $\mathcal{B}(\bar{x}, K)$ e $(\mathcal{B}(\bar{x}, K))^*$ para várias classes de problemas é fundamental para se obter condições de otimalidade úteis tanto do ponto de vista computacional quanto do ponto de vista teórico.

A seguir, vemos que os cones $\mathcal{T}(\bar{x}, K)$ e $\mathcal{B}(\bar{x}, K)$ sempre são fechados, para qualquer conjunto K .

Proposição 2.1. *Sejam $K \subseteq \mathbb{R}^n$ um conjunto qualquer e $\bar{x} \in K$. Então, os cones $\mathcal{T}(\bar{x}, K)$ e $\mathcal{B}(\bar{x}, K)$ são fechados e tem-se que*

$$\{0\} \subseteq cl\mathcal{V}(\bar{x}, K) \subseteq \mathcal{T}(\bar{x}, K) \subseteq \mathcal{B}(\bar{x}, K)$$

Demonstração. A demonstração pode ser encontrada na referência [8]. □

2.1.3 Condições de otimalidade no caso de restrições de igualdade

Agora estudaremos o problema de otimização com restrições de igualdade:

$$\begin{aligned} & \text{Minimizar} && f(x) \\ & \text{Sujeito a} && x \in K, \end{aligned} \tag{2.14}$$

$$\begin{aligned} K &= \{x \in \mathbb{R}^n : h(x) = 0\} \\ &= \{x \in \mathbb{R}^n : h_i(x) = 0, i = 1, \dots, l\}, \end{aligned} \tag{2.15}$$

onde $f : \mathbb{R}^n \rightarrow \mathbb{R}$ e $h : \mathbb{R}^n \rightarrow \mathbb{R}^l$ são funções dadas. A partir da descrição analítica do cone tangente (de Bouligand) e do dual dele para os conjuntos com estrutura (2.15), obteremos as condições necessárias de otimalidade (de Lagrange) para o problema (2.14)-(2.15).

O primeiro resultado é fácil de obter: só as direções que pertencem ao subespaço

$$\begin{aligned} \ker h'(\bar{x}) &= \{d \in \mathbb{R}^n : h'(\bar{x})d = 0\} \\ &= \{d \in \mathbb{R}^n : \langle \nabla h_i(\bar{x}), d \rangle = 0, i = 1, \dots, l\} \end{aligned}$$

podem ser tangentes em relação ao conjunto K dado por (2.15). Ou seja, as direções tangentes tem que ser ortogonais a todos os gradientes das componentes de h no ponto \bar{x} .

Proposição 2.2. *Suponhamos que a função $h : \mathbb{R}^n \rightarrow \mathbb{R}^l$ seja diferenciável no ponto $\bar{x} \in K$, onde K é dado por (2.15). Então,*

$$\mathcal{B}(\bar{x}, K) \subseteq \ker h'(\bar{x}). \tag{2.16}$$

Demonstração. Obviamente, $0 \in \mathcal{B}(\bar{x}, K)$ e $0 \in \ker h'(\bar{x})$. Seja $d \in \mathcal{B}(\bar{x}, K) \setminus \{0\}$. Então, existem $(t_k) \subseteq \mathbb{R}_+ \setminus \{0\}$ e $(d_k) \subseteq \mathbb{R}^n$ tais que $\lim_{k \rightarrow +\infty} t_k = 0$, $\lim_{k \rightarrow +\infty} d_k = d$ e $(\bar{x} + t_k d_k) \in K$ (isto é, $h(\bar{x} + t_k d_k) = 0$) para todo $k \in \mathbb{N}$. Logo, pela diferenciabilidade de h em \bar{x} temos

$$\begin{aligned} 0 &= h(\bar{x} + t_k d_k) - h(\bar{x}) \\ &= t_k h'(\bar{x}) d_k + o(t_k \|d_k\|). \end{aligned}$$

Dividindo os dois lados desta igualdade por $t_k > 0$ e tomando o limite quando $k \rightarrow +\infty$, como $\lim_{k \rightarrow +\infty} d_k = d$, segue-se que

$$0 = h'(\bar{x})d,$$

isto é, $d \in \ker h'(\bar{x})$. □

Em geral, não há igualdade na relação (2.16), como pode ser visto no exemplo seguinte.

Exemplo 2.1. Seja $h : \mathbb{R} \rightarrow \mathbb{R}$, $h(x) = x^2$, $\bar{x} = 0$.

Evidentemente, $K = \{x \in \mathbb{R} : x^2 = 0\} = \{\bar{x}\}$. Portanto, $\mathcal{B}(\bar{x}, K) = \{0\}$. Mas como $h'(\bar{x}) = 0$, tem-se que $\ker h'(\bar{x}) = \mathbb{R}$.

Concluimos que neste caso $\mathcal{B}(\bar{x}, K) \neq \ker h'(\bar{x})$.

Para conseguir uma descrição precisa (isto é, uma igualdade em (2.16)), necessitamos algumas hipóteses adicionais. Em particular, provaremos essa igualdade quando qualquer uma das seguintes duas condições é satisfeita:

$$\begin{aligned} &\{\nabla h_i(\bar{x}), i = 1, \dots, l\} \\ &\text{é um conjunto linearmente independente,} \end{aligned} \tag{2.17}$$

ou

$$\begin{aligned} &h \text{ é uma função afim, isto é,} \\ &h(x) = Ax - b, \text{ onde } A \in \mathbb{R}^{l \times n}, b \in \mathbb{R}^l. \end{aligned} \tag{2.18}$$

As condições que garantem a igualdade em (2.16) se chamam condições de regularidade das restrições do problema (2.14) (ou ainda, condições de qualificação das restrições) como veremos a seguir, (2.17) e (2.18) são condições de regularidade nesse sentido. Note que a condição de independência linear dos gradientes das restrições (2.17) é equivalente a

$$\text{rank } h'(\bar{x}) = l,$$

isto é, a transformação linear $h'(\bar{x})$ é sobrejetiva:

$$\text{im } h'(\bar{x}) = \mathbb{R}^l.$$

Usando resultados da Álgebra Linear, ver a referência [13], outras condições equivalentes são

$$\ker (h'(\bar{x}))^T = (\text{im } h'(\bar{x}))^\perp = \{0\}.$$

Finalmente, quando a condição (2.17) é satisfeita, necessariamente temos que $l \leq n$, isto é, o número das restrições é menor ou igual ao número das variáveis (no espaço \mathbb{R}^n não podem existir mais de n vetores linearmente independentes).

Teorema 2.3. *Suponhamos que $h : \mathbb{R}^n \rightarrow \mathbb{R}^l$ seja diferenciável numa vizinhança do ponto $\bar{x} \in K$, onde K é dado por (2.15), e que a derivada de h seja contínua em \bar{x} . Então,*

(a) *Quando a condição de regularidade das restrições (2.17) de independência linear dos gradientes das restrições é satisfeita, tem-se que*

$$\mathcal{T}(\bar{x}, K) = \mathcal{B}(\bar{x}, K) = \ker h'(\bar{x}).$$

(b) *Quando as restrições são lineares, isto é, vale (2.18), tem-se que*

$$\mathcal{V}(\bar{x}, K) = \mathcal{T}(\bar{x}, K) = \mathcal{B}(\bar{x}, K) = \ker h'(\bar{x}).$$

Demonstração. A demonstração pode ser encontrada em [8]. □

Teorema 2.4. *Suponhamos que $f : \mathbb{R}^n \rightarrow \mathbb{R}$ seja diferenciável no ponto $\bar{x} \in \mathbb{R}^n$ e que $h : \mathbb{R}^n \rightarrow \mathbb{R}^l$ seja diferenciável numa vizinhança deste ponto, com derivada contínua em \bar{x} . Suponhamos também que \bar{x} seja um minimizador local do problema (2.14)-(2.15).*

Se vale uma das condições de regularidade das restrições (de independência linear dos gradientes (2.17) ou de linearidade (2.18)). Então,

$$\langle \nabla f(\bar{x}), d \rangle = 0 \quad \forall d \in \ker h'(\bar{x}).$$

Demonstração. Pelos teoremas 2.1 e 2.3, tem-se que

$$\langle \nabla f(\bar{x}), d \rangle \geq 0 \quad \forall d \in \mathcal{B}(\bar{x}, K) = \ker h'(\bar{x}).$$

Como $\ker h'(\bar{x})$ é um subespaço, para todo $d \in \ker h'(\bar{x})$ temos que $-d \in \ker h'(\bar{x})$ e portanto,

$$\langle \nabla f(\bar{x}), d \rangle \geq 0, \quad -\langle \nabla f(\bar{x}), d \rangle \geq 0,$$

o que implica a afirmação. □

A condição necessária de primeira ordem para o problema (2.14)-(2.15), dada no teorema 2.4, pode ser escrita em outra forma equivalente:

$$-\nabla f(\bar{x}) \in (\ker h'(\bar{x}))^\perp = \text{im} (h'(\bar{x}))^\top,$$

onde o sinal negativo foi introduzido por razões essencialmente históricas. Para fazer conexão com a noção de cone dual vemos o resultado seguinte, que pode ser encontrado em [8].

Lema 2.2. Para qualquer matriz $A \in \mathbb{R}^{l \times n}$ tem-se

$$(\ker A)^* = \operatorname{im} A^\top = (\ker A)^\perp.$$

Demonstração. Veja a referência [8]. □

Portanto, sob as hipóteses do teorema 2.4, temos que

$$-\nabla f(\bar{x}) \in (\ker h'(\bar{x}))^* = (\ker h'(\bar{x}))^\perp = \operatorname{im} (h'(\bar{x}))^\top.$$

Em particular, chegamos à seguinte condição necessária de otimalidade para o problema (2.14)-(2.15), sob uma das condições de regularidade das restrições ((2.17) ou (2.18)):

$$-\nabla f(\bar{x}) \in \operatorname{im} (h'(\bar{x}))^\top,$$

isto é,

$$\exists \bar{\lambda} \in \mathbb{R}^l \text{ tal que } \nabla f(\bar{x}) = - (h'(\bar{x}))^\top \bar{\lambda}.$$

Esta é a condição clássica de Lagrange. É comum formular este resultado em termos da Lagrangiana do problema (2.14)-(2.15):

$$\begin{aligned} L : \mathbb{R}^n \times \mathbb{R}^l &\longrightarrow \mathbb{R} \\ (x, \lambda) &\longrightarrow L(x, \lambda) = f(x) + \langle \lambda, h(x) \rangle \\ &= f(x) + \sum_{i=1}^l \lambda_i h_i(x). \end{aligned}$$

Observamos que

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

Onde $h'(x)$ é a matriz jacobiana da função h no ponto x , definida por

$$h'(x) = \begin{pmatrix} \frac{\partial h_1(x)}{\partial x_1} & \cdots & \frac{\partial h_1(x)}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial h_l(x)}{\partial x_1} & \cdots & \frac{\partial h_l(x)}{\partial x_n} \end{pmatrix} = \begin{pmatrix} (\nabla h_1(x))^\top \\ \vdots \\ (\nabla h_l(x))^\top \end{pmatrix}$$

Teorema 2.5. Suponhamos que $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ seja diferenciável no ponto $\bar{x} \in \mathbb{R}^n$ e que $h : \mathbb{R}^n \longrightarrow \mathbb{R}^l$ seja diferenciável numa vizinhança deste ponto, com derivada contínua em \bar{x} . Suponhamos também que \bar{x} seja um minimizador local do problema (2.14)-(2.15).

Se vale uma das condições de regularidade das restrições (de independência linear dos gradientes (2.17) ou de linearidade (2.18)). Então, existe $\bar{\lambda} \in \mathbb{R}^l$ tal que

$$\nabla_x L(\bar{x}, \bar{\lambda}) = 0. \tag{2.19}$$

Sob a condição (2.17), $\bar{\lambda} \in \mathbb{R}^l$ satisfazendo (2.19) é único.

Demonstração. A existência de um $\bar{\lambda} \in \mathbb{R}^l$ satisfazendo (2.19) já foi verificada acima. Mostramos então a unicidade para o caso (2.17). Para algum outro $\hat{\lambda} \in \mathbb{R}^l$ satisfazendo $\nabla_x L(\bar{x}, \hat{\lambda}) = 0$, temos

$$\begin{aligned} 0 &= \nabla_x L(\bar{x}, \bar{\lambda}) - \nabla_x L(\bar{x}, \hat{\lambda}) \\ &= (h'(\bar{x}))^\top (\bar{\lambda} - \hat{\lambda}) \\ &= \sum_{i=1}^l (\bar{\lambda}_i - \hat{\lambda}_i) \nabla h_i(\bar{x}). \end{aligned}$$

Pela condição (2.17), isto implica que $\bar{\lambda} - \hat{\lambda} = 0$, isto é, $\bar{\lambda}$ satisfazendo (2.19) é único. \square

No teorema 2.5, assim como no teorema 2.3, a condição de regularidade das restrições é crucial: sem esta condição as afirmações do teorema não podem ser garantidas.

Definição 2.10. Dizemos que um ponto $\bar{x} \in \mathbb{R}^n$ é ponto estacionário do problema (2.14)-(2.15), se $\bar{x} \in K$ e (2.19) é satisfeito para algum $\bar{\lambda} \in \mathbb{R}^l$. Este $\bar{\lambda}$ se chama o multiplicador de Lagrange, associado ao ponto estacionário \bar{x} .

Portanto, qualquer solução local do problema (2.14)-(2.15) (onde as funções são diferenciáveis e as derivadas são contínuas) que satisfaz a condição de regularidade das restrições, é um ponto estacionário.

O sistema de equações

$$\nabla_x L(x, \lambda) = 0, \quad h(x) = 0$$

em relação de $(x, \lambda) \in \mathbb{R}^n \times \mathbb{R}^l$, que caracteriza os pontos estacionários do problema (2.14)-(2.15) e os multiplicadores de Lagrange associados, se chama o sistema de Lagrange.

2.1.4 Elementos de convexidade, o problema de minimização convexo

Nesta subseção estudamos conjuntos convexos e funções convexas. Convexidade é uma noção muito importante na teoria de otimização. Com hipóteses de convexidade, as condições necessárias de otimalidade passam a ser suficientes. Em outras palavras, todo ponto estacionário torna-se uma solução do problema. Além disso, no caso convexo podemos desenvolver a teoria de dualidade em sua forma mais completa, isto é, associar ao problema original (primal) um outro problema, chamado dual, que sob certas hipóteses é equivalente ao original e às vezes é mais fácil de resolver. Finalmente, as ferramentas da Análise Convexa serão necessárias para a caracterização do cone dual do cone tangente no caso de restrições mistas (de igualdade e desigualdade), o que resulta nas condições de otimalidade primais-duais de Karush-Kuhn-Tucker. Para uma visão mais detalhada veja as referências [8], [21].

Definição 2.11. Um conjunto $K \subseteq \mathbb{R}^n$ é chamado conjunto convexo se para quaisquer $x, y \in K$ e $\alpha \in [0, 1]$, tem-se que $\alpha x + (1 - \alpha)y \in K$.

O ponto $\alpha x + (1 - \alpha)y$, onde $\alpha \in [0, 1]$, se chama combinação convexa de x e y (com parâmetro α). A seguir mostramos que a dualização de um cone qualquer (mesmo não-convexo) sempre produz um cone convexo, para mais detalhes ver [21], [8].

Proposição 2.3. *Para todo cone $\emptyset \neq C \subseteq \mathbb{R}^n$, o cone dual C^* sempre é convexo e fechado.*

Demonstração. Sejam $x, y \in C^*$, isto é, $\langle x, d \rangle \leq 0$ e $\langle y, d \rangle \leq 0$ para todo $d \in C$. Seja $\alpha \in [0, 1]$. Para qualquer $d \in C$ temos que

$$\langle \alpha x + (1 - \alpha)y, d \rangle = \alpha \langle x, d \rangle + (1 - \alpha) \langle y, d \rangle \leq 0,$$

isto é, $\alpha x + (1 - \alpha)y \in C^*$. Portanto, C^* é convexo.

Seja $(y_k) \subseteq C^*$ tal que $\lim_{k \rightarrow +\infty} y_k = y$. Fixando $d \in C$ arbitrário, e passando ao limite quando $k \rightarrow +\infty$ na relação $\langle y_k, d \rangle \leq 0$, obtemos que $\langle y, d \rangle \leq 0$. Portanto, como $d \in C$ era arbitrário, $y \in C^*$. Isto mostra que C^* é fechado. \square

Definição 2.12. Dado um conjunto $K \subseteq \mathbb{R}^n$ qualquer, o fecho cônico de K , denotado por cone K , é o menor cone convexo em \mathbb{R}^n que contém K (ou, equivalentemente, a interseção de todos os cones convexos em \mathbb{R}^n que contém K).

A seguir veremos que se $K \subseteq \mathbb{R}^n$ é um conjunto convexo. Então, o cone tangente de Bouligand $\mathcal{B}(\bar{x}, K)$ é convexo e fechado, a demonstração dos seguintes resultados pode ser encontrada em [8].

Proposição 2.4. *Seja $K \subseteq \mathbb{R}^n$ um conjunto convexo. Tem-se que*

$$\text{cone } K = \{d \in \mathbb{R}^n : d = \alpha x, x \in K, \alpha \in \mathbb{R}_+\}.$$

Demonstração. Ver [8]. \square

Teorema 2.6. *Sejam $K \subseteq \mathbb{R}^n$ um conjunto convexo, $\bar{x} \in K$. Então,*

$$\begin{aligned} \mathcal{V}(\bar{x}, K) &= \text{cone}(K - \{\bar{x}\}), \\ \mathcal{T}(\bar{x}, K) &= \mathcal{B}(\bar{x}, K) = \text{cl}\mathcal{V}(\bar{x}, K) = \text{cl}\text{cone}(K - \{\bar{x}\}). \end{aligned}$$

Em particular, $\mathcal{T}(\bar{x}, K)$ é convexo e fechado.

Demonstração. Veja a referência [8]. \square

Proposição 2.5. *Seja $C \subseteq \mathbb{R}^n$ um cone qualquer. Então, $\text{cl}C$ é um cone e tem-se que*

$$C^* = (\text{cl}C)^*$$

Em particular, se $K \subseteq \mathbb{R}^n$ é um conjunto convexo e $\bar{x} \in K$, tem-se que

$$(\mathcal{T}(\bar{x}, K))^* = (\mathcal{B}(\bar{x}, K))^* = (\mathcal{V}(\bar{x}, K))^* = (\text{cone}(K - \{\bar{x}\}))^*.$$

Demonstração. A demonstração pode ser encontrada na referência [8]. \square

Outra noção útil em Análise Convexa é a de cone normal.

Definição 2.13. Sejam $K \subseteq \mathbb{R}^n$ um conjunto convexo e $\bar{x} \in K$. O cone normal (cone de direções normais) no ponto \bar{x} em relação ao conjunto K é dado por

$$\mathcal{N}(\bar{x}, K) = \{d \in \mathbb{R}^n : \langle d, x - \bar{x} \rangle \leq 0 \quad \forall x \in K\}.$$

A seguir mostramos que, no caso convexo, o dual de cone tangente é exatamente o cone normal definido acima.

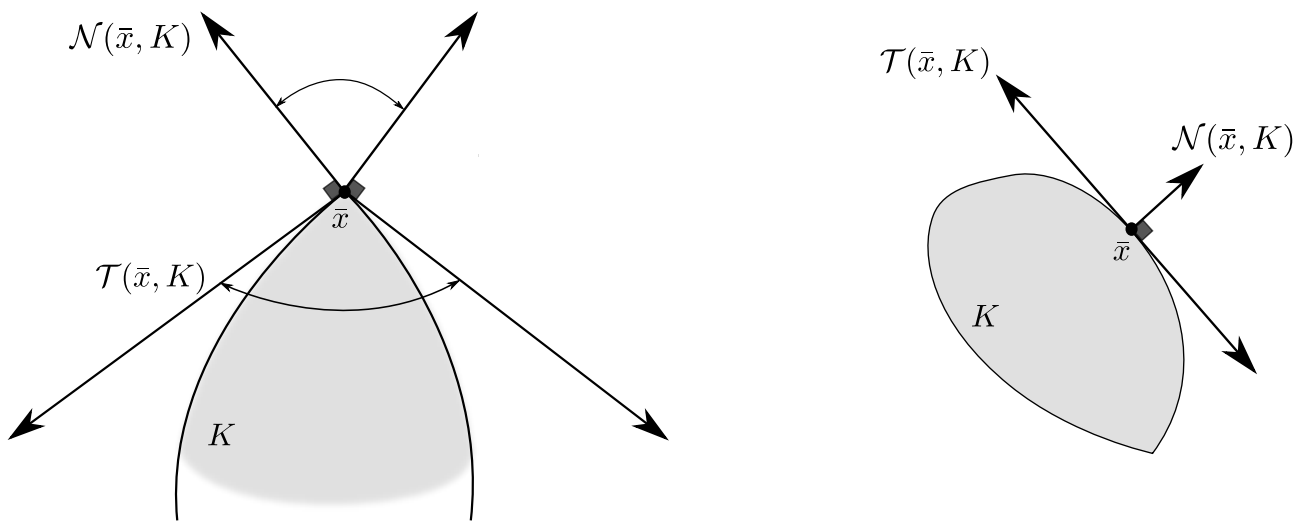


Figura 6 – Cone normal e o dual de cone tangente para um conjunto convexo.

Teorema 2.7. Sejam $K \subseteq \mathbb{R}^n$ um conjunto convexo e $\bar{x} \in K$. Então,

$$(\mathcal{T}(\bar{x}, K))^* = (\mathcal{B}(\bar{x}, K))^* = (\mathcal{V}(\bar{x}, K))^* = (\text{cone}(K - \{\bar{x}\}))^* = \mathcal{N}(\bar{x}, K).$$

Demonstração. As primeiras três igualdades já foram provadas (veja a proposição 2.5). Suponhamos que $y \in (\text{cone}(K - \{\bar{x}\}))^*$, isto é, $\langle y, d \rangle \leq 0$ para todo $d \in \text{cone}(K - \{\bar{x}\})$. Em particular, $\langle y, x - \bar{x} \rangle \leq 0$ para todo $x \in K$, isto implica que $y \in \mathcal{N}(\bar{x}, K)$. Reciprocamente, suponhamos que $y \in \mathcal{N}(\bar{x}, K)$. Temos que $\langle y, x - \bar{x} \rangle \leq 0$ para todo $x \in K$, isto é, $\langle y, d \rangle \leq 0$ para todo $d \in (K - \{\bar{x}\})$. Logo, $\langle y, d \rangle \leq 0$ para todo $d \in \text{cone}(K - \{\bar{x}\})$. Concluimos que $y \in (\text{cone}(K - \{\bar{x}\}))^*$. \square

Como consequência da caracterização do cone tangente e do seu dual, obtemos as seguintes condições de otimalidade para um problema com conjunto viável convexo.

Teorema 2.8. Sejam $K \subseteq \mathbb{R}^n$ um conjunto convexo e $f : \mathbb{R}^n \rightarrow \mathbb{R}$ uma função diferenciável no ponto $\bar{x} \in K$.

Se \bar{x} é um minimizador local de f no conjunto K . Então,

$$\langle \nabla f(\bar{x}), x - \bar{x} \rangle \geq 0 \quad \forall x \in K, \quad (2.20)$$

ou equivalentemente,

$$-\nabla f(\bar{x}) \in \mathcal{N}(\bar{x}, K). \quad (2.21)$$

Demonstração. Pelo teorema 2.1 tem-se que $\langle \nabla f(\bar{x}), d \rangle \geq 0$ para todo $d \in \mathcal{B}(\bar{x}, K)$. Pelo teorema 2.6, temos que

$$\{d \in \mathbb{R}^n : d = x - \bar{x}, x \in K\} \subseteq \text{cone}(K - \{\bar{x}\}) \subseteq \mathcal{B}(\bar{x}, K),$$

o que implica (2.20). Pela definição 2.13, (2.20) e (2.21) são equivalentes. \square

Definição 2.14. Se $K \subseteq \mathbb{R}^n$ é um conjunto convexo, diz-se que a função $f : K \rightarrow \mathbb{R}$ é convexa em K quando para quaisquer $x, y \in K$ e $\alpha \in [0, 1]$ tem-se que

$$f(\alpha x + (1 - \alpha)y) \leq \alpha f(x) + (1 - \alpha)f(y).$$

A função f diz-se estritamente convexa quando a desigualdade acima é estrita para todos $x \neq y$ e $\alpha \in (0, 1)$.

Dizemos que o problema

$$\begin{array}{ll} \text{Minimizar} & f(x) \\ \text{Sujeito a} & x \in K \end{array} \quad (2.22)$$

é um problema de minimização convexo quando $K \subseteq \mathbb{R}^n$ é um conjunto convexo e $f : K \rightarrow \mathbb{R}$ é uma função convexa no conjunto K . A importancia da convexidade já pode ser vista no resultado seguinte.

Teorema 2.9. *Sejam $K \subseteq \mathbb{R}^n$ um conjunto convexo e $f : K \rightarrow \mathbb{R}$ uma função convexa em K .*

Então, todo minimizador local do problema (2.22) é global. Além disso, o conjunto de minimizadores é convexo.

Se f é estritamente convexa, não pode haver mais de um minimizador.

Demonstração. A demonstração pode ser encontrada na referência [8]. \square

Definição 2.15. Se $K \subseteq \mathbb{R}^n$ é um conjunto convexo, dizemos que $f : K \rightarrow \mathbb{R}$ é uma função côncava em K , quando a função $-f$ é convexa em K .

As afirmações do teorema 2.9 são verdadeiras se substituimos minimização de uma função convexa num conjunto convexo por maximização de uma função côncava num conjunto convexo.

A seguir mostramos que a interseção de conjuntos convexos é um conjunto convexo.

Proposição 2.6. *Sejam $K_i \subseteq \mathbb{R}^n$, $i \in I$, conjuntos convexos, onde I é um conjunto qualquer (possivelmente infinito).*

Então, a interseção $K = \bigcap_{i \in I} K_i$ também é um conjunto convexo.

Demonstração. Sejam $x, y \in K$. Então, $x, y \in K_i$ para todo $i \in I$. Como os conjuntos K_i , $i \in I$ são convexos, $\alpha x + (1 - \alpha)y \in K_i$ para qualquer $\alpha \in [0, 1]$ e todo $i \in I$. Logo, $\alpha x + (1 - \alpha)y \in K$. Portanto, K é convexo. \square

Corolário 2.10. *Um conjunto poliedral em \mathbb{R}^n é convexo.*

Demonstração. Um conjunto poliedral é o conjunto de soluções de um sistema finito de equações e inequações lineares, isto é, uma interseção de semi-espacos e hiperplanos que são conjuntos convexos. Portanto, pela proposição 2.6, um conjunto poliedral é convexo. \square

Agora estudaremos algumas propriedades das funções convexas. Primeiro mostramos que uma soma de múltiplos não negativos de um número finito de funções convexas é uma função convexa.

Proposição 2.7. *Sejam $K \subseteq \mathbb{R}^n$ um conjunto convexo e $f_i : K \rightarrow \mathbb{R}$, $i = 1, \dots, p$, funções convexas em K . Então, para quaisquer $\mu_i \in \mathbb{R}_+$, $i = 1, \dots, p$, a função*

$$f : K \rightarrow \mathbb{R}, \quad f(x) = \sum_{i=1}^p \mu_i f_i(x)$$

é convexa em K .

Demonstração. Para $x, y \in K$ e $\alpha \in [0, 1]$ quaisquer, temos que

$$\begin{aligned} f(\alpha x + (1 - \alpha)y) &= \sum_{i=1}^p \mu_i f_i(\alpha x + (1 - \alpha)y) \\ &\leq \sum_{i=1}^p \mu_i (\alpha f_i(x) + (1 - \alpha)f_i(y)) \\ &= \alpha \sum_{i=1}^p \mu_i f_i(x) + (1 - \alpha) \sum_{i=1}^p \mu_i f_i(y) \\ &= \alpha f(x) + (1 - \alpha)f(y), \end{aligned}$$

onde a desigualdade segue da convexidade de f_i e do fato de que $\mu_i \geq 0$, $i = 1, \dots, p$. \square

Proposição 2.8. *Sejam $g : \mathbb{R}^n \rightarrow \mathbb{R}$ uma função convexa, $\psi : \mathbb{R} \rightarrow \mathbb{R}$ uma função convexa e não-decrescente. Então, a função*

$$f : \mathbb{R}^n \rightarrow \mathbb{R}, \quad f(x) = \psi(g(x))$$

é convexa.

Demonstração. Para $x, y \in \mathbb{R}^n$ e $\alpha \in [0, 1]$ quaisquer, pela convexidade de g , temos que

$$g(\alpha x + (1 - \alpha)y) \leq \alpha g(x) + (1 - \alpha)g(y).$$

Agora,

$$\begin{aligned} f(\alpha x + (1 - \alpha)y) &= \psi(g(\alpha x + (1 - \alpha)y)) \\ &\leq \psi(\alpha g(x) + (1 - \alpha)g(y)) \\ &\leq \alpha \psi(g(x)) + (1 - \alpha)\psi(g(y)) \\ &= \alpha f(x) + (1 - \alpha)f(y), \end{aligned}$$

onde a primeira desigualdade segue do fato de que ψ é não decrescente, e a segunda decorre da convexidade de ψ . \square

Teorema 2.11. *Suponhamos que o conjunto $K \subseteq \mathbb{R}^n$ seja convexo e a função $f : K \rightarrow \mathbb{R}$ seja convexa em K . Então, o conjunto de nível*

$$L_{f,K}(c) = \{x \in K : f(x) \leq c\}$$

é convexo para todo $c \in \mathbb{R}$.

Demonstração. Seja $c \in \mathbb{R}$ arbitrário. Se $L_{f,K}(c) = \emptyset$, a conclusão segue (o conjunto vazio é convexo trivialmente).

Sejam $x, y \in L_{f,K}(c)$, isto é, $x, y \in K$ e $f(x) \leq c$, $f(y) \leq c$. Pela convexidade de K , $\alpha x + (1 - \alpha)y \in K$. Agora, pela convexidade de f em K ,

$$\begin{aligned} f(\alpha x + (1 - \alpha)y) &\leq \alpha f(x) + (1 - \alpha)f(y) \\ &\leq \alpha c + (1 - \alpha)c = c, \end{aligned}$$

o que mostra que $(\alpha x + (1 - \alpha)y) \in L_{f,K}(c)$. \square

Definição 2.16. Seja $K \subseteq \mathbb{R}^n$ um conjunto convexo. Dizemos que $g : K \rightarrow \mathbb{R}^m$ é convexa em K se todas as funções $g_i : K \rightarrow \mathbb{R}$, $i = 1, \dots, m$, são convexas em K .

Como consequência do teorema 2.11, obtemos uma condição suficiente para garantir a convexidade de um conjunto definido por restrições funcionais.

Corolário 2.12. *Seja $\Omega \subseteq \mathbb{R}^n$ um conjunto convexo. Sejam $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$ uma função convexa e $h : \mathbb{R}^n \rightarrow \mathbb{R}^l$ uma função afim. Então, o conjunto*

$$K = \{x \in \Omega : h(x) = 0, g(x) \leq 0\}$$

é convexo.

Demonstração. Como é fácil de observar, tem-se que

$$K = K_1 \cap K_2 \cap K_3,$$

onde

$$\begin{aligned} K_1 &= \bigcap_{i=1}^l \{x \in \Omega : h_i(x) \leq 0\} = \bigcap_{i=1}^l L_{h_i, \Omega}(0), \\ K_2 &= \bigcap_{i=1}^l \{x \in \Omega : -h_i(x) \leq 0\} = \bigcap_{i=1}^l L_{-h_i, \Omega}(0), \\ K_3 &= \bigcap_{j=1}^m \{x \in \Omega : g_j(x) \leq 0\} = \bigcap_{i=1}^l L_{g_j, \Omega}(0). \end{aligned}$$

Lembramos que quando h é afim, h e $-h$ são convexas. Portanto, todos os conjuntos de nível acima são convexas (teorema 2.11) e a interseção deles também é um conjunto convexo (proposição 2.6). \square

Quando uma função é diferenciável, a convexidade admite várias caracterizações que são muito úteis para determinar se uma função é convexa ou não.

Teorema 2.13. *Sejam $K \subseteq \mathbb{R}^n$ um conjunto aberto e convexo e $f : K \rightarrow \mathbb{R}$ uma função diferenciável em K . Então, as propriedades seguintes são equivalentes:*

(a) *A função f é convexa em K .*

(b) *Para todo $x, y \in K$,*

$$f(y) \geq f(x) + \langle \nabla f(x), y - x \rangle.$$

(c) *Para todo $x, y \in K$,*

$$\langle \nabla f(y) - \nabla f(x), y - x \rangle \geq 0.$$

Quando f é duas vezes diferenciável em K , as propriedades acima também são equivalentes a

(d) *A matriz Hessiana de f é semi-definida positiva em todo ponto de K :*

$$\langle \nabla^2 f(x)d, d \rangle \geq 0 \quad \forall x \in K, \quad \forall d \in \mathbb{R}^n.$$

Demonstração. A demonstração pode ser encontrada nas referências [8], [21]. \square

2.1.5 Problemas com restrições de igualdade e desigualdade

Consideremos um problema com restrições mistas de igualdade e desigualdade

$$\begin{aligned} & \text{Minimizar} && f(x) \\ & \text{Sujeito a} && x \in K, \end{aligned} \tag{2.23}$$

$$K = \{x \in \mathbb{R}^n : h(x) = 0, g(x) \leq 0\} \tag{2.24}$$

onde $f : \mathbb{R}^n \rightarrow \mathbb{R}$, $h : \mathbb{R}^n \rightarrow \mathbb{R}^l$ e $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$ são funções dadas.

Definição 2.17. Seja $\bar{x} \in K$, onde K é definido por (2.24). Dizemos que a restrição de desigualdade que corresponde ao índice $i \in \{1, \dots, m\}$ é ativa no ponto \bar{x} quando $g_i(\bar{x}) = 0$, isto é, quando a desigualdade é satisfeita como igualdade.

O conjunto dos índices das restrições ativas no ponto $\bar{x} \in K$ será denotado por

$$I(\bar{x}) = \{i = 1, \dots, m : g_i(\bar{x}) = 0\}.$$

Não incluímos nesta definição as restrições de igualdade porque todas elas sempre são ativas em todo ponto viável.

Um minimizador local \bar{x} do problema (2.23)-(2.24) também é minimizador local do problema

$$\begin{aligned} & \text{Minimizar} && f(x) \quad \text{sujeito a: } x \in \tilde{K}, \\ & \tilde{K} = \{x \in \mathbb{R}^n : && h(x) = 0, g_i(x) = 0, i \in I(\bar{x})\}, \end{aligned}$$

onde removemos todas as restrições que são inativas em \bar{x} . As condições de otimalidade para o problema (2.23)-(2.24) podem ser obtidas através das condições de otimalidade para o problema acima. No entanto, esta abordagem também produz resultados relativamente fracos. Primeiro, a condição de regularidade neste caso seria a seguinte:

$$\begin{aligned} & \{\nabla h_i(\bar{x}) : i = 1, \dots, l\} \cup \{\nabla g_i(\bar{x}) : i \in I(\bar{x})\} \\ & \text{é um conjunto linearmente independente.} \end{aligned} \tag{2.25}$$

Esta condição é mais forte da que nós usaremos. E o que é talvez ainda mais importante, usando o problema modificado não é possível provar que os multiplicadores de Lagrange associados às restrições de desigualdade devem ser não negativos, o que é uma propriedade fundamental.

Supondo que h e g sejam diferenciáveis em $\bar{x} \in K$, introduzimos o cone obtido pela linearização (no ponto \bar{x}) das restrições de igualdade e das restrições ativas de desigualdade:

$$\begin{aligned} \mathcal{L}(\bar{x}, K) &= \left\{ d \in \mathbb{R}^n : \begin{aligned} & \langle \nabla h_i(\bar{x}), d \rangle = 0, \quad i = 1, \dots, l, \\ & \langle \nabla g_i(\bar{x}), d \rangle \leq 0 \quad \forall i \in I(\bar{x}) \end{aligned} \right\} \\ &= \{d \in \ker h'(\bar{x}) : \langle \nabla g_i(\bar{x}), d \rangle \leq 0 \quad \forall i \in I(\bar{x})\} \end{aligned} \tag{2.26}$$

Como veremos a seguir, só as direções que pertencem ao cone $\mathcal{L}(\bar{x}, K)$ podem ser tangentes no ponto $\bar{x} \in K$ em relação ao conjunto K .

Proposição 2.9. *Seja $K \subseteq \mathbb{R}^n$ definido por (2.24). Suponhamos que $h : \mathbb{R}^n \rightarrow \mathbb{R}^l$ e $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$ sejam diferenciáveis em $\bar{x} \in K$. Então,*

$$\mathcal{B}(\bar{x}, K) \subseteq \mathcal{L}(\bar{x}, K),$$

onde $\mathcal{B}(\bar{x}, K)$ é o cone tangente de Bouligand e o cone $\mathcal{L}(\bar{x}, K)$ é definido por (2.26).

Demonstração. É claro que $0 \in \mathcal{B}(\bar{x}, K) \cap \mathcal{L}(\bar{x}, K)$. Seja então $d \in \mathcal{B}(\bar{x}, K) \setminus \{0\}$. Portanto, existem seqüências $(t_k) \subseteq \mathbb{R}_+ \setminus \{0\}$ e $(d_k) \subseteq \mathbb{R}^n$ tais que

$$\lim_{k \rightarrow +\infty} t_k = 0, \quad \lim_{k \rightarrow +\infty} d_k = d \quad \text{e} \quad \bar{x} + t_k d_k \in K,$$

isto é, $h(\bar{x} + t_k d_k) = 0$ e $g(\bar{x} + t_k d_k) \leq 0$ para todo $k \in \mathbb{N}$. Logo,

$$\begin{aligned} 0 &= h(\bar{x} + t_k d_k) - h(\bar{x}) \\ &= t_k (h'(\bar{x})d_k + o(t_k \|d_k\|) / t_k) \end{aligned}$$

e para todo $i \in I(\bar{x})$,

$$\begin{aligned} 0 &\geq g_i(\bar{x} + t_k d_k) \\ &= g_i(\bar{x} + t_k d_k) - g_i(\bar{x}) \\ &= t_k (\langle \nabla g_i(\bar{x}), d_k \rangle + o(t_k \|d_k\|) / t_k). \end{aligned}$$

Dividindo em ambos os lados das relações acima por $t_k > 0$ e tomando o limite quando $k \rightarrow +\infty$, tem-se que $d \in \ker h'(\bar{x})$ e $\langle \nabla g_i(\bar{x}), d \rangle \leq 0 \quad \forall i \in I(\bar{x})$. Portanto, $d \in \mathcal{L}(\bar{x}, K)$. \square

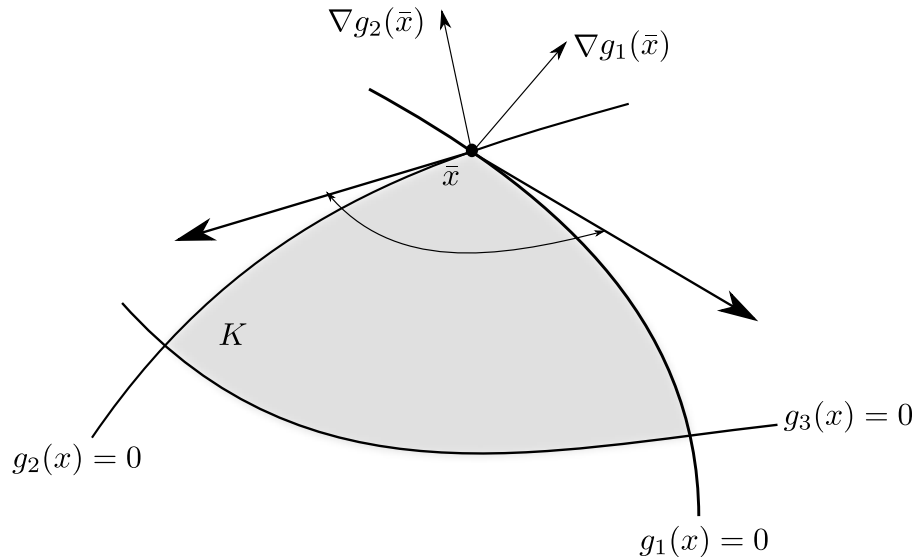


Figura 7 – O cone $\mathcal{L}(\bar{x}, K)$ obtido pela linearização das restrições das desigualdades ativas em \bar{x} .

Observamos que, diferentemente do caso de restrições de igualdade, no caso em que há restrições de desigualdade, o conjunto $\mathcal{L}(\bar{x}, K)$ obtido pela linearização das restrições pode não conter nenhum subespaço não-trivial. Por isso, podemos combinar a proposição 2.9 e o teorema 2.2 para afirmar o seguinte (para mais detalhes ver [8]).

Teorema 2.14. *Suponhamos que $f : \mathbb{R}^n \rightarrow \mathbb{R}$, $h : \mathbb{R}^n \rightarrow \mathbb{R}^l$ e $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$ sejam diferenciáveis no ponto $\bar{x} \in K$. Seja $\mathcal{L}(\bar{x}, K)$ o cone dado por (2.26). Se*

$$\langle \nabla f(\bar{x}), d \rangle > 0 \quad \forall d \in \mathcal{L}(\bar{x}, K) \setminus \{0\},$$

então f satisfaz no conjunto K em torno de \bar{x} a condição de crescimento linear: existem uma vizinhança V de \bar{x} e um número $\beta > 0$ tais que

$$f(x) - f(\bar{x}) \geq \beta \|x - \bar{x}\| \quad \forall x \in K \cap V.$$

Em particular, \bar{x} é um minimizador local estrito do problema (2.23)-(2.24).

Como foi feito para o caso de restrições de igualdade, necessitamos algumas hipóteses adicionais para obter uma descrição precisa do cone tangente (de Bouligand), isto é, para satisfazer a igualdade $\mathcal{B}(\bar{x}, K) = \mathcal{L}(\bar{x}, K)$. De novo, uma das condições deste tipo é a linearidade das restrições:

$$\begin{aligned} h \text{ e } g \text{ são funções afins, isto é,} \\ h(x) = Ax - b, \quad A \in \mathbb{R}^{l \times n}, \quad b \in \mathbb{R}^l \\ g(x) = Cx - d, \quad C \in \mathbb{R}^{m \times n}, \quad d \in \mathbb{R}^m. \end{aligned} \tag{2.27}$$

A seguir introduzimos algumas outras condições de regularidade, apropriadas para o caso não-linear.

Dizemos que o ponto $\bar{x} \in K$ satisfaz a condição de regularidade de Mangasarian-Fromovitz, quando

$$\begin{aligned} \{\nabla h_i(\bar{x}) : i = 1, \dots, l\} \\ \text{é um conjunto linearmente independente e} \\ \exists \bar{d} \in \ker h'(\bar{x}) \text{ tal que } \langle \nabla g_i(\bar{x}), \bar{d} \rangle < 0 \quad \forall i \in I(\bar{x}). \end{aligned} \tag{2.28}$$

Observamos que a condição de regularidade de independência linear dos gradientes das restrições ativas (2.25) implica na condição de Mangasarian-Fromovitz (2.28) (ver as referências [21], [8]).

O fato de que a condição de Mangasarian-Fromovitz (2.28) não implica em independência linear dos gradientes (2.25) fica claro na proposição 2.10 a seguir, que introduz uma caracterização dual equivalente (dada por (2.29)) à condição de Mangasarian-Fromovitz (2.28). Com efeito, basta observar que (2.25) significa que o sistema de equações em (2.29) só tem a solução nula, sem impor a condição $\mu_i \geq 0$, $i \in I(\bar{x})$. Logo, a condição (2.29) (que é equivalente a (2.28)) é menos restritiva de que a condição de independência linear (2.25).

Proposição 2.10. *Seja $\bar{x} \in K$, onde K é dado por (2.24). O ponto \bar{x} satisfaz a condição de regularidade de Mangasarian-Fromovitz (2.28) se, e somente se, o seguinte sistema nas variáveis $(\lambda, \mu) \in \mathbb{R}^l \times \mathbb{R}^{|I(\bar{x})|}$:*

$$0 = \sum_{i=1}^l \lambda_i \nabla h_i(\bar{x}) + \sum_{i \in I(\bar{x})} \mu_i \nabla g_i(\bar{x}), \quad \mu_i \geq 0, \quad i \in I(\bar{x}) \tag{2.29}$$

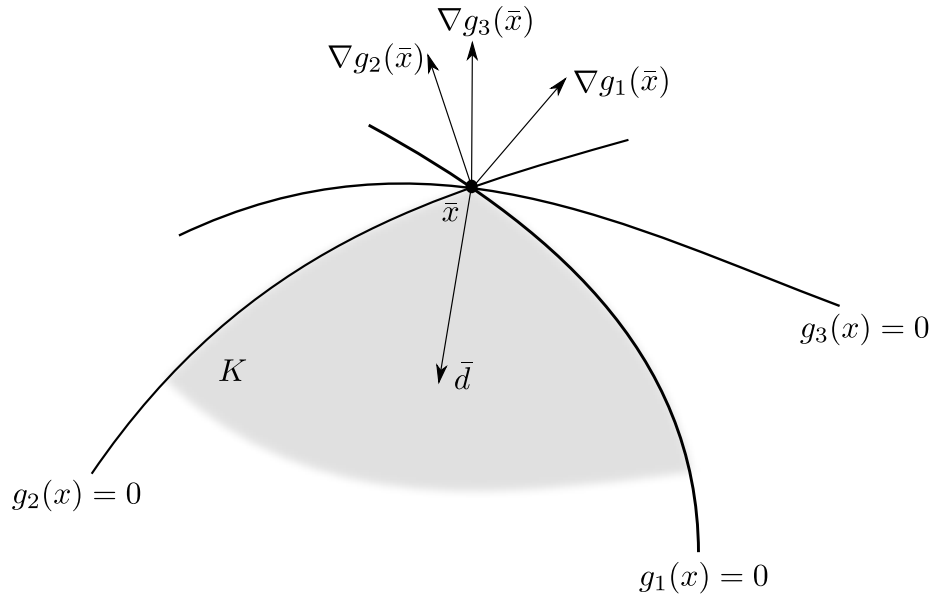


Figura 8 – A direção \bar{d} satisfaz a condição de Mangasarian-Fromovitz.

tem unicamente a solução nula.

Demonstração. Veja a referência [8]. □

Outra condição de qualificação das restrições é a condição de Slater:

$$\begin{aligned} h \text{ é uma função afim, } g \text{ é uma função convexa, e} \\ \exists \hat{x} \in \mathbb{R}^n \text{ tal que } h(\hat{x}) = 0, g_i(\hat{x}) < 0, i = 1, \dots, m \end{aligned} \quad (2.30)$$

Teorema 2.15. *Seja K definido por (2.24). Suponhamos que $h : \mathbb{R}^n \rightarrow \mathbb{R}^l$ seja diferenciável numa vizinhança do ponto $\bar{x} \in K$, com derivada contínua em \bar{x} . Suponhamos que $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$ seja diferenciável em \bar{x} . Seja $\mathcal{L}(\bar{x}, K)$ o cone de linearização das restrições ativas definido por (2.26). Então:*

- (a) *Se a condição de regularidade das restrições (2.28) de Mangasarian-Fromovitz é satisfeita, tem-se que*

$$\mathcal{T}(\bar{x}, K) = \mathcal{B}(\bar{x}, K) = \mathcal{L}(\bar{x}, K).$$

- (b) *Se a condição de regularidade das restrições (2.30) de Slater é satisfeita, tem-se que*

$$cl\mathcal{V}(\bar{x}, K) = \mathcal{T}(\bar{x}, K) = \mathcal{B}(\bar{x}, K) = \mathcal{L}(\bar{x}, K).$$

- (c) *Se as restrições são lineares, isto é, (2.27) é satisfeita, tem-se que*

$$\mathcal{V}(\bar{x}, K) = \mathcal{T}(\bar{x}, K) = \mathcal{B}(\bar{x}, K) = \mathcal{L}(\bar{x}, K).$$

Demonstração. A demonstração pode ser encontrada na referência [8]. □

Pelo visto nos teoremas 2.1 e 2.15, temos o resultado seguinte.

Teorema 2.16. *Suponhamos que $f : \mathbb{R}^n \rightarrow \mathbb{R}$ e $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$ sejam diferenciáveis no ponto $\bar{x} \in \mathbb{R}^n$ e que $h : \mathbb{R}^n \rightarrow \mathbb{R}^l$ seja diferenciável numa vizinhança deste ponto, com derivada contínua em \bar{x} .*

Se \bar{x} é um minimizador local do problema (2.23)-(2.24) e se alguma das condições de regularidade das restrições do teorema 2.15 é satisfeita (isto é, ou de linearidade (2.27) ou de Mangasarian-Fromovitz (2.28) ou de Slater (2.30)). Então, tem-se que

$$\langle \nabla f(\bar{x}), d \rangle \geq 0 \quad \forall d \in \mathcal{L}(\bar{x}, K)$$

onde o cone de linearização das restrições ativas $\mathcal{L}(\bar{x}, K)$ vem dado por (2.26).

A condição necessária de primeira ordem do teorema 2.16 pode ser escrita em forma primal-dual:

$$-\nabla f(\bar{x}) \in (\mathcal{L}(\bar{x}, K))^* \quad (2.31)$$

As condições clássicas de Karush-Kuhn-Tucker são obtidas via a computação explícita do cone dual acima. Precisaremos então do seguinte resultado sobre o cone dual.

Lema 2.3 (Lema de Farkas). *Seja*

$$K = \{x \in \mathbb{R}^n : A_1 x = 0, A_2 x \leq 0\}$$

onde $A_1 \in \mathbb{R}^{m_1 \times n}$ e $A_2 \in \mathbb{R}^{m_2 \times n}$ são arbitrárias. Então,

$$K^* = \left\{ x \in \mathbb{R}^n : \begin{array}{l} x = A_1^\top y_1 + A_2^\top y_2, \\ y_1 \in \mathbb{R}^{m_1}, y_2 \in \mathbb{R}_+^{m_2} \end{array} \right\} \quad (2.32)$$

Demonstração. Veja a demonstração na referência [8]. □

Suponhamos agora que $\bar{x} \in K$ seja um minimizador local do problema (2.23)-(2.24), onde f é diferenciável em \bar{x} e as funções h e g satisfazem as hipóteses do teorema 2.15. Sob qualquer uma das condições de regularidade das restrições, pelo teorema 2.16 (ver também (2.31) e o Lema de Farkas 2.3), obtemos a seguinte condição necessária primal-dual de otimalidade de primeira ordem: existem $\bar{\lambda} \in \mathbb{R}^l$ e $\bar{\mu}_i \geq 0$, $i \in I(\bar{x})$, tais que

$$-\nabla f(\bar{x}) = (h'(\bar{x}))^\top \bar{\lambda} + \sum_{i \in I(\bar{x})} \bar{\mu}_i \nabla g_i(\bar{x}).$$

Para escrever esta condição no formato conhecido como o teorema de Karush-Kuhn-Tucker, introduzimos a Lagrangiana do problema (2.23)-(2.24):

$$\begin{aligned} L : \mathbb{R}^n \times \mathbb{R}^l \times \mathbb{R}^m &\longrightarrow \mathbb{R} \\ L(x, \lambda, \mu) &= f(x) + \langle \lambda, h(x) \rangle + \langle \mu, g(x) \rangle. \end{aligned}$$

Temos que

$$\begin{aligned} \nabla_x L(x, \lambda, \mu) &= \nabla f(x) + (h'(x))^\top \lambda + (g'(x))^\top \mu, \\ x \in \mathbb{R}^n, \lambda \in \mathbb{R}^l, \mu \in \mathbb{R}^m. \end{aligned}$$

Teorema 2.17. *Sejam $f : \mathbb{R}^n \rightarrow \mathbb{R}$ e $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$ funções diferenciáveis no ponto $\bar{x} \in \mathbb{R}^n$, e seja $h : \mathbb{R}^n \rightarrow \mathbb{R}^l$ uma função diferenciável numa vizinhança do ponto \bar{x} , com derivada contínua neste ponto. Seja \bar{x} um minimizador local do problema (2.23)-(2.24). Então, sob qualquer uma das três condições de regularidade das restrições do teorema 2.15 (ou de linearidade (2.27) ou de Mangasarian-Fromovitz (2.28) ou de Slater (2.30)), tem-se que existem $\bar{\lambda} \in \mathbb{R}^l$ e $\bar{\mu} \in \mathbb{R}_+^m$ tais que*

$$\nabla_x L(\bar{x}, \bar{\lambda}, \bar{\mu}) = 0 \quad (2.33)$$

$$\bar{\mu}_i g_i(\bar{x}) = 0, \quad i = 1, \dots, m \quad (2.34)$$

Sob a condição de independência linear dos gradientes das restrições ativas (2.25), $(\bar{\lambda}, \bar{\mu}) \in \mathbb{R}^l \times \mathbb{R}_+^m$ que satisfaz (2.33)-(2.34) é único.

Demonstração. A demonstração pode ser encontrada em [8]. □

Definição 2.18. Dizemos que $\bar{x} \in \mathbb{R}^n$ é um ponto estacionário do problema (2.23)-(2.24) (ou ainda, um ponto de KKT), quando $\bar{x} \in K$ e existem $\bar{\lambda} \in \mathbb{R}^l$ e $\bar{\mu} \in \mathbb{R}_+^m$ tais que as condições (2.33) e (2.34) são satisfeitas. Estes elementos $\bar{\lambda}$ e $\bar{\mu}$ serão chamados os multiplicadores de Lagrange associados ao ponto estacionário \bar{x} .

O sistema de equações e inequações

$$\nabla_x L(x, \lambda, \mu) = 0, \quad h(x) = 0, \quad \mu \geq 0, \quad g(x) \leq 0, \quad \langle \mu, g(x) \rangle = 0$$

nas variáveis $(x, \lambda, \mu) \in \mathbb{R}^n \times \mathbb{R}^l \times \mathbb{R}^m$, caracteriza os pontos estacionários do problema (2.23)-(2.24) e os multiplicadores de Lagrange associados. Este sistema se chama sistema de Karush-Kuhn-Tucker (KKT) e pode se reescrever da seguinte forma:

$$\nabla f(x) + \sum_{i=1}^l \lambda_i \nabla h_i(x) + \sum_{i=1}^m \mu_i \nabla g_i(x) = 0 \quad (2.35)$$

$$h(x) = 0 \quad (2.36)$$

$$\mu_i g_i(x) = 0 \quad \forall i = 1, \dots, m \quad (2.37)$$

$$\mu_i \geq 0 \quad \forall i = 1, \dots, m \quad (2.38)$$

$$g_i(x) \leq 0 \quad \forall i = 1, \dots, m \quad (2.39)$$

No caso de um problema de programação convexo, as condições de otimalidade de KKT tornam-se suficientes.

Teorema 2.18. *Seja $\bar{x} \in \mathbb{R}^n$ um ponto estacionário do problema (2.23)-(2.24), no sentido da definição 2.18. Sejam f e g funções convexas e seja h uma função afim. Então, \bar{x} é um minimizador do problema (2.23)-(2.24).*

Demonstração. Veja a demonstração na referencia [8]. □

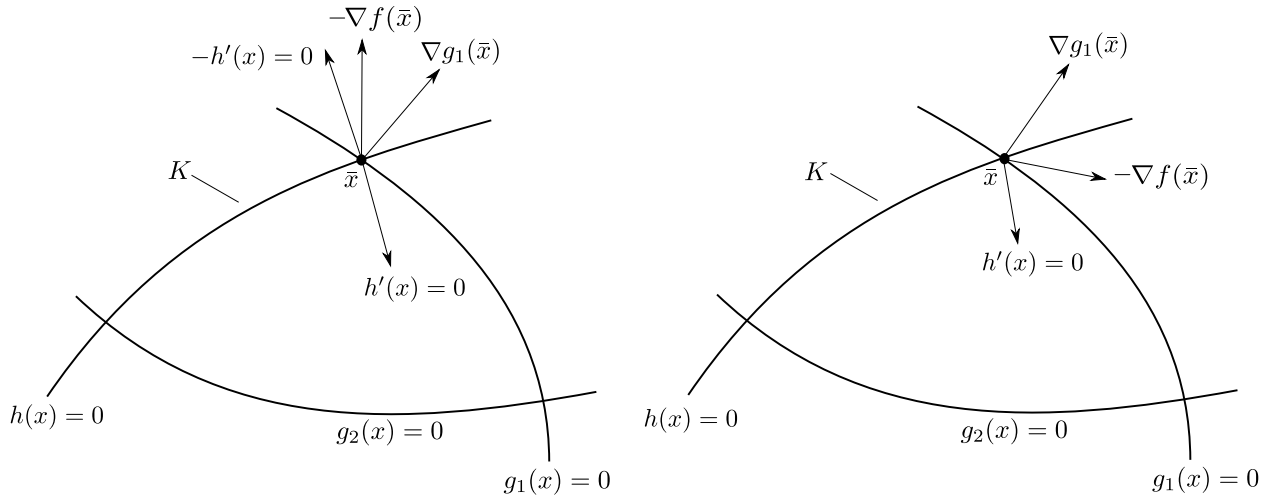


Figura 9 – Interpretação geométrica das condições de KKT.

2.2 Introdução as inequações variacionais

Nesta seção veremos alguns resultados sobre inequações variacionais, estudaremos condições necessárias e suficientes para garantir a existência de soluções de uma inequação variacional (VI), assim como também damos alguns exemplos e a definição formal de VI; para mais detalhes dos resultados que foram tratados nesta seção consultar as seguintes referências [10], [1], [11], [12].

2.2.1 Exemplos de Inequações variacionais

Nesta subseção pretendemos apresentar os elementos das inequações variacionais com alguns exemplos, oferecemos uma descrição da natureza de uma inequação variacional sem tentar formular ainda uma definição precisa. Para maiores informações veja a referência [10].

Exemplo 2.2. Seja $f : I = [a, b] \rightarrow \mathbb{R}$ uma função real diferenciável definida no intervalo fechado I . Seja $x_0 \in I$ tal que:

$$f(x_0) = \min_{x \in I} f(x).$$

Então, três casos podem ocorrer:

1. Se $a < x_0 < b$, então $f'(x_0) = 0$.
2. Se $x_0 = a$, então $f'_+(x_0) \geq 0$.
3. Se $x_0 = b$, então $f'_-(x_0) \leq 0$.

Estas afirmações podem ser resumidas, escrevendo:

$$f'(x_0)(x - x_0) \geq 0 \quad \forall x \in I. \quad (2.40)$$

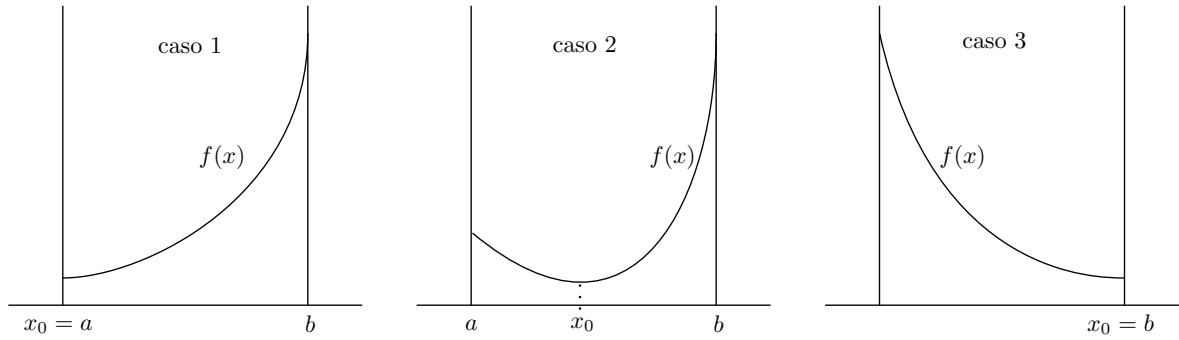


Figura 10 – Representação dos mínimos de f do exemplo 2.2.

A desigualdade (2.40) é conhecida como inequação variacional (VI) e dizemos que x_0 satisfaz ou é solução da inequação variacional (2.40).

Exemplo 2.3. Seja $K \subseteq \mathbb{R}^n$ um conjunto convexo e fechado, $f : K \rightarrow \mathbb{R}$ uma função diferenciável, e $x_0 \in K$ tais que:

$$f(x_0) = \min_{x \in K} f(x).$$

Assim, x_0 é um ponto onde o mínimo da função f é atingido; considero $x \in K$ arbitrário. Então, desde que K é convexo tem-se que $x_0 + t(x - x_0) \in K \quad \forall t \in [0, 1]$. Logo, a função $\phi : [0, 1] \rightarrow \mathbb{R}$ definida por $\phi(t) = f(x_0 + t(x - x_0))$ atinge seu mínimo em $t = 0$. Conseqüentemente, como no exemplo 2.2, temos que $\phi'(0) = \langle \nabla f(x_0), x - x_0 \rangle \geq 0$ para qualquer $x \in K$. Portanto, o ponto x_0 satisfaz a inequação variacional:

$$x_0 \in K : \quad \langle \nabla f(x_0), x - x_0 \rangle \geq 0 \quad \forall x \in K.$$

Observação 2.1. No exemplo 2.3, se K é limitado, a existência de ao menos um x_0 é imediata.

Exemplo 2.4. Seja $\Omega \subseteq \mathbb{R}^n$ um domínio limitado com fronteira $\partial\Omega$ e seja ψ uma função definida em $\bar{\Omega} = \Omega \cup \partial\Omega$, satisfazendo:

$$\max_{\Omega} \psi \geq 0 \quad \text{e} \quad \psi \leq 0 \quad \text{em} \quad \partial\Omega,$$

definimos

$$K = \left\{ v \in C^1(\bar{\Omega}) : v \geq \psi \text{ em } \Omega \text{ e } v = 0 \text{ em } \partial\Omega \right\}$$

Assim, dados $u, v \in K$ e $t \in [0, 1]$ temos que $(1-t)u + tv = 0$ em $\partial\Omega$. Além disso, como $u, v \geq \psi$ em Ω ; então, $(1-t)u \geq (1-t)\psi$ e $tv \geq t\psi$. Logo, $(1-t)u + tv \geq (1-t)\psi + t\psi = \psi$ em Ω . Então, $(1-t)u + tv \in K$ e portanto K é um conjunto convexo; suponha que K é não vazio. Definimos a função $f : K \rightarrow \mathbb{R}$ da seguinte maneira:

$$f(v) = \int_{\Omega} |\nabla v|^2 dx,$$

procuramos uma função $u \in K$ tal que:

$$f(u) = \int_{\Omega} |\nabla u|^2 dx = \min_{v \in K} \int_{\Omega} |\nabla v|^2 dx = \min_{v \in K} f(v)$$

assumindo que tal u existe, como K é convexo, de forma analoga ao exemplo 2.3, segue-se que para qualquer $v \in K$, tem-se que $u + t(v - u) \in K$, assim a função $\phi : [0, 1] \rightarrow \mathbb{R}$ definida por:

$$\phi(t) = \int_{\Omega} |\nabla (u + t(v - u))|^2 dx = f(u + t(v - u)) \quad 0 \leq t \leq 1$$

atinge seu mínimo em $t = 0$, isto implica que $\phi'(0) \geq 0$, que leva à seguinte VI:

$$u \in K : \int_{\Omega} \nabla u \cdot \nabla (v - u) dx \geq 0 \quad \forall v \in K$$

2.2.2 Teoremas sobre ponto fixo

Vários problemas de análise não linear podem ser resolvidos por meio do teorema do ponto fixo. Os resultados tratados nesta subseção podem ser vistos com mais detalhes nas referências [12], [1], [10], [11].

Seja $F : A \rightarrow A$ uma aplicação definida num conjunto A em si mesmo. Um ponto $x \in A$ é chamado ponto fixo de F , se e somente se $F(x) = x$. O primeiro teorema sobre ponto fixo, chamado o teorema da contração é a formulação abstracta do método de Picard de aproximações sucessivas.

Seja S um espaço métrico, com métrica d .

Definição 2.19. Uma aplicação $F : S \rightarrow S$ é uma contração, se e somente se

$$d(F(x), F(y)) \leq \alpha d(x, y) \quad \forall x, y \in S,$$

para algum $\alpha \in [0, 1)$, quando permitimos $\alpha = 1$, F é chamada não expansiva.

Teorema 2.19. *Seja S um espaço métrico completo e $F : S \rightarrow S$ uma contração. Então, existe um único ponto fixo de F .*

Demonstração. A demonstração pode ser encontrada nas referências [12], [11]. □

Teorema 2.20. *(Brouwer) Seja $B \subseteq \mathbb{R}^n$ uma bola fechada e $F : B \rightarrow B$ uma aplicação continua. Então, F possui ao menos um ponto fixo.*

No teorema de Brouwer 2.20, a bola B pode ser substituída por um subconjunto convexo compacto de \mathbb{R}^n , para maiores informações veja [10].

2.2.3 A caracterização da projeção sobre um conjunto convexo

Lema 2.4. *Seja K um subconjunto fechado e convexo de um espaço de Hilbert H . Então, para cada $x \in H$ existe um único $y \in K$ tal que:*

$$\|x - y\| = \inf_{\eta \in K} \|x - \eta\| \quad (2.41)$$

Demonstração. Seja $d = \inf_{\eta \in K} \|x - \eta\|$, pela definição de ínfimo, existe $\eta_k \in K$ tal que:

$$d \leq \|x - \eta_k\| \leq d + \frac{1}{k},$$

tomando limite quando $k \rightarrow +\infty$ obtemos:

$$\lim_{k \rightarrow \infty} \|x - \eta_k\| = d = \inf_{\eta \in K} \|x - \eta\|. \quad (2.42)$$

Agora, dados $k, h \in \mathbb{N}$, aplicamos a lei do paralelogramo aos vetores $x - \eta_k$ e $x - \eta_h$ e obtemos o seguinte:

$$2\|x - \eta_k\|^2 + 2\|x - \eta_h\|^2 = \|2x - (\eta_k + \eta_h)\|^2 + \|\eta_k - \eta_h\|^2.$$

Então, temos:

$$\|\eta_k - \eta_h\|^2 = 2\|x - \eta_k\|^2 + 2\|x - \eta_h\|^2 - 4\left\|x - \frac{1}{2}(\eta_k + \eta_h)\right\|^2. \quad (2.43)$$

Assim, como K é convexo, tem-se $\frac{1}{2}(\eta_k + \eta_h) \in K$. Então,

$$d \leq \left\|x - \frac{1}{2}(\eta_k + \eta_h)\right\|;$$

consequentemente,

$$d^2 \leq \left\|x - \frac{1}{2}(\eta_k + \eta_h)\right\|^2.$$

Assim, de (2.43) temos que:

$$0 \leq \|\eta_k - \eta_h\|^2 \leq 2\|x - \eta_k\|^2 + 2\|x - \eta_h\|^2 - 4d^2,$$

e a partir de (2.42) concluímos que:

$$\lim_{k, h \rightarrow \infty} \|\eta_k - \eta_h\| = 0;$$

então $(\eta_k) \subseteq K$ é uma seqüência de Cauchy e portanto ela é convergente. Logo, como K é fechado, existe $y \in K$ tal que $\lim_{k \rightarrow \infty} \eta_k = y$. Além disso por (2.42) tem-se:

$$\|x - y\| = \lim_{k \rightarrow \infty} \|x - \eta_k\| = d = \inf_{\eta \in K} \|x - \eta\|.$$

Para mostrar que y é único, observe que qualquer par de elementos $y, y' \in K$ os quais satisfazem (2.41), podem ser inseridos em (2.43) em lugar de η_k e η_h , isto é:

$$\|y - y'\|^2 = 2\|x - y\|^2 + 2\|x - y'\|^2 - 4\left\|x - \frac{1}{2}(y + y')\right\|^2 \leq 4d^2 - 4d^2 = 0.$$

Assim, $y = y'$ e portanto $y \in K$ é único. □

Observação 2.2. O ponto y satisfazendo (2.41) é chamado a projeção de x em K e escrevemos $y = Pr_K x$; por outro lado note que $Pr_K x = x \quad \forall x \in K$.

Teorema 2.21. *Seja K um subconjunto convexo e fechado de um espaço de Hilbert H . Então, $y = Pr_K x$ se e somente se,*

$$y \in K : \quad \langle y, \eta - y \rangle \geq \langle x, \eta - y \rangle \quad \forall \eta \in K. \quad (2.44)$$

Demonstração. Seja $x \in H$ e suponha que $y = Pr_K x$; então $y \in K$. Além disso, como K é convexo, dado qualquer $\eta \in K$, tem-se que $y + t(\eta - y) \in K$ para todo $0 \leq t \leq 1$. Portanto, por (2.41), a função $\phi : [0, 1] \rightarrow \mathbb{R}$, definida por:

$$\phi(t) = \|x - (y + t(\eta - y))\|^2 = \|x - y\|^2 - 2t \langle x - y, \eta - y \rangle + t^2 \|\eta - y\|^2$$

atinge seu mínimo em $t = 0$. Logo, $\phi'(0) \geq 0$, isto é, $\langle x - y, \eta - y \rangle \leq 0$ para todo $\eta \in K$ ou equivalentemente, tem-se que:

$$\langle y, \eta - y \rangle \geq \langle x, \eta - y \rangle \quad \forall \eta \in K.$$

Reciprocamente, suponha que valha (2.44). Então,

$$\begin{aligned} 0 \leq \langle y - x, \eta - y \rangle &= \langle y - x, \eta - x + x - y \rangle = \langle y - x, x - y \rangle + \langle y - x, \eta - x \rangle \\ &= -\|x - y\|^2 + \langle y - x, \eta - x \rangle. \end{aligned}$$

Logo, $\|y - x\|^2 \leq \langle y - x, \eta - x \rangle \leq \|y - x\| \|\eta - x\|$. Assim, tem-se: $\|y - x\| \leq \|\eta - x\|$ para todo $\eta \in K$. Consequentemente, $\|x - y\| = \inf_{\eta \in K} \|x - \eta\|$. Portanto, $y = Pr_K x$. \square

Corolário 2.22. *Seja K um conjunto convexo e fechado de um espaço de Hilbert H . Então, o operador $Pr_K : H \rightarrow K$ é não expansivo, isto é:*

$$\|Pr_K x - Pr_K x'\| \leq \|x - x'\| \quad \forall x, x' \in H. \quad (2.45)$$

Demonstração. Dados $x, x' \in H$, sejam $y = Pr_K x$ e $y' = Pr_K x'$. Então, pelo teorema 2.21 temos que:

$$\begin{aligned} y \in K : \quad \langle y, \eta - y \rangle &\geq \langle x, \eta - y \rangle \quad \forall \eta \in K, \\ y' \in K : \quad \langle y', \eta - y' \rangle &\geq \langle x', \eta - y' \rangle \quad \forall \eta \in K, \end{aligned}$$

escolhendo $\eta = y'$ na primeira desigualdade e $\eta = y$ na segunda desigualdade, obtemos:

$$\begin{aligned} \langle y, y' - y \rangle &\geq \langle x, y' - y \rangle, \\ \langle y', y - y' \rangle &\geq \langle x', y - y' \rangle. \end{aligned}$$

Agora, somando as duas desigualdades anteriores tem-se que:

$$\langle y, y' - y \rangle + \langle y', y - y' \rangle \geq \langle x, y' - y \rangle + \langle x', y - y' \rangle,$$

isto é equivalente a:

$$\langle y, y' - y \rangle - \langle y', y' - y \rangle \geq \langle x, y' - y \rangle - \langle x', y' - y \rangle.$$

Assim, temos que: $\langle y - y', y' - y \rangle \geq \langle x - x', y' - y \rangle$ e multiplicando por -1 a ambos lados da desigualdade, obtemos:

$$\langle y - y', y - y' \rangle \leq \langle x - x', y - y' \rangle.$$

Logo:

$$\|y - y'\|^2 = \langle y - y', y - y' \rangle \leq \langle x - x', y - y' \rangle \leq \|x - x'\| \|y - y'\|$$

Portanto, $\|y - y'\| \leq \|x - x'\|$. □

Teorema 2.23. *Seja $K \subseteq \mathbb{R}^n$ um conjunto convexo e compacto e seja $F : K \rightarrow K$ uma função contínua. Então, F possui um ponto fixo.*

Demonstração. Como K é limitado, existe $r > 0$ tal que $K \subseteq B[0, r] = B$. Logo, pelo corolário 2.22 a Pr_K é contínua. Assim, temos que a aplicação $F \circ Pr_K : B \rightarrow B$ é contínua. Logo, pelo teorema 2.20 a função $F \circ Pr_K$ possui pelo menos um ponto fixo; isto é, existe $x \in K$ tal que $(F \circ Pr_K)(x) = x$, e desde que $Pr_K x = x$, tem-se que $F(x) = x$. □

Teorema 2.24. *Seja $K \subseteq \mathbb{R}^n$ um conjunto compacto e convexo e seja $F : K \rightarrow \mathbb{R}^n$ uma função contínua. Então, existe $x \in K$ tal que:*

$$\langle F(x), y - x \rangle \geq 0 \quad \forall y \in K. \tag{2.46}$$

Demonstração. Como a função Pr_K é contínua, denotando I a função identidade, temos que a aplicação $Pr_K \circ (I - F) : K \rightarrow K$ é contínua. Então, pelo teorema 2.23 esta aplicação possui um ponto fixo; isto é, existe $x \in K$ tal que $x = (Pr_K \circ (I - F))(x)$. Logo, pelo teorema 2.21, temos que:

$$\langle x, y - x \rangle \geq \langle x - F(x), y - x \rangle \quad \forall y \in K$$

Portanto, $\langle F(x), y - x \rangle \geq 0 \quad \forall y \in K$. □

2.2.4 Inequações variacionais em \mathbb{R}^n

Nesta subseção daremos a definição formal de inequação variacional (VI) e veremos alguns teoremas que garantem a existencia de soluções, ver as referências [10], [3]. Na sua forma geral, uma VI é formalmente definida abaixo.

Definição 2.20. Dado um conjunto $K \subseteq \mathbb{R}^n$ e uma aplicação $F : K \rightarrow \mathbb{R}^n$, a inequação variacional denotada por $VI(K, F)$, consiste em encontrar um vetor $x \in K$ tal que

$$\langle F(x), y - x \rangle \geq 0 \quad \forall y \in K. \tag{2.47}$$

Ao longo deste trabalho estamos interessados apenas na situação em que o conjunto K é fechado e a função F é contínua.

Se K é um conjunto compacto e convexo, e a função $F : K \rightarrow \mathbb{R}^n$ contínua, temos provado a existência de uma solução da VI (2.47) no teorema 2.24. Por outro lado, note que a VI nem sempre possui uma solução. Por exemplo, se $K = \mathbb{R}$, $f(x)(y - x) \geq 0 \quad \forall y \in K$ não tem solução para $f(x) = e^x$.

De fato. Se existe $x_0 \in \mathbb{R}$ tal que $f(x_0)(y - x_0) \geq 0 \quad \forall y \in \mathbb{R}$. Então,

$$e^{x_0}(y - x_0) \geq 0 \quad \forall y \in \mathbb{R},$$

e como $e^{x_0} > 0$, temos que $y \geq x_0 \quad \forall y \in \mathbb{R}$ a qual é uma contradição.

O seguinte teorema dá uma condição necessária e suficiente para a existência de soluções. Dado um conjunto fechado e convexo $K \subseteq \mathbb{R}^n$ e $R > 0$, seja $K_R = K \cap B[0, R]$, note que K_R é um conjunto compacto e convexo. Logo, como $F : K_R \rightarrow \mathbb{R}^n$ é contínua, pelo teorema 2.24 temos que existe $x_R \in K_R$ tal que:

$$\langle F(x_R), w - x_R \rangle \geq 0 \quad \forall w \in K_R. \quad (2.48)$$

Teorema 2.25. *Seja $K \subseteq \mathbb{R}^n$ um conjunto fechado e convexo e $F : K \rightarrow \mathbb{R}^n$ contínua, uma condição necessária e suficiente para que exista uma solução da VI (2.47), é que exista $R > 0$ tal que, uma solução $x \in K_R$ de (2.48) satisfaz $\|x\| < R$.*

Demonstração. Suponha que existe uma solução $x \in K$ da VI (2.47). Então, considero um $R > 0$ tal que $x \in B(0, R)$ de tal forma que $K_R = K \cap B[0, R] \neq \emptyset$. Logo, $x \in K_R$ é uma solução da inequação variacional (2.48) e $\|x\| < R$.

Reciprocamente, suponha que existe $R > 0$ tal que uma solução $x \in K_R$ de (2.48), satisfaz $\|x\| < R$. Então, como a bola $B(0, R)$ é aberta e $x \in B(0, R)$ temos que existe $\delta > 0$ tal que $B(x, \delta) \subseteq B(0, R)$. Por outro lado, dado $y \in K$ com $y \neq x$, considero $\tau > 0$ tal que:

$$0 < \tau < \min \left\{ 1, \frac{\delta}{\|y - x\|} \right\}$$

Agora, desde que $x, y \in K$ e K é convexo temos que $w = x + \tau(y - x) \in K$. Além disso, $\|w - x\| = \tau\|y - x\| < \delta$. Então, $w \in B(x, \delta) \subseteq B(0, R)$; logo, $w \in K_R = K \cap B[0, R]$, e como $x \in K_R$ é solução da inequação variacional (2.48), temos que:

$$0 \leq \langle F(x), w - x \rangle = \langle F(x), \tau(y - x) \rangle = \tau \langle F(x), y - x \rangle,$$

Isto é,

$$x \in K_R : \quad \langle F(x), y - x \rangle \geq 0 \quad \forall y \in K$$

O qual significa que $x \in K$ é solução da VI (2.47). □

A continuação estabelecemos outra condição suficiente para a existência de soluções de uma inequação variacional.

Corolário 2.26. *Seja $K \subseteq \mathbb{R}^n$ um conjunto fechado e convexo e $F : K \rightarrow \mathbb{R}^n$ uma função contínua, se existe $x_0 \in K$ tal que:*

$$\lim_{\|x\| \rightarrow \infty} \frac{\langle F(x) - F(x_0), x - x_0 \rangle}{\|x - x_0\|} = +\infty$$

Então, existe uma solução da VI (2.47).

Demonstração. Considero $M > \|F(x_0)\| \geq 0$. Logo, pela definição de limite, existe $\delta > 0$ tal que:

$$\frac{\langle F(x) - F(x_0), x - x_0 \rangle}{\|x - x_0\|} \geq M, \text{ sempre que } \|x\| > \delta \text{ e } x \in K.$$

Agora, considero $R > \max\{\delta, \|x_0\|\}$. Assim, temos que:

$$\langle F(x) - F(x_0), x - x_0 \rangle \geq M\|x - x_0\|; \text{ sempre que } \|x\| \geq R, x \in K.$$

Daqui, obtemos o seguinte:

$$\langle F(x), x - x_0 \rangle \geq M\|x - x_0\| + \langle F(x_0), x - x_0 \rangle; \text{ sempre que } \|x\| \geq R, x \in K$$

Mas, pela desigualdade de Cauchy-Schwarz; temos que, $|\langle F(x_0), x - x_0 \rangle| \leq \|F(x_0)\| \|x - x_0\|$. Logo, se $\|x\| \geq R$ e $x \in K$. Então,

$$\langle F(x), x - x_0 \rangle \geq M\|x - x_0\| - \|F(x_0)\| \|x - x_0\| = (M - \|F(x_0)\|) \|x - x_0\|$$

e como $M > \|F(x_0)\|$. Então, $\|x\| \geq R > \|x_0\|$ com $x \in K$ implica que:

$$\langle F(x), x - x_0 \rangle \geq (M - \|F(x_0)\|) \|x - x_0\| > 0,$$

Isto é:

$$\langle F(x), x - x_0 \rangle > 0 \text{ para } \|x\| \geq R, x \in K. \quad (2.49)$$

Por outro lado, como K é fechado e convexo. Então, $K_R = K \cap B[0, R]$ é compacto e convexo. Além disso, sabemos que $R > \|x_0\|$; então, $x_0 \in K_R$. Assim, $K_R \neq \emptyset$. Logo, desde que F é contínua, pelo teorema 2.24 temos que existe $x_R \in K_R$ tal que:

$$\langle F(x_R), w - x_R \rangle \geq 0 \quad \forall w \in K_R \quad (2.50)$$

Agora, considero $w = x_0$ na desigualdade (2.50) e obtemos:

$$\langle F(x_R), x_R - x_0 \rangle = -\langle F(x_R), x_0 - x_R \rangle \leq 0.$$

Além disso, Suponha que $\|x_R\| = R$; então, desde que $x_R \in K$, pela desigualdade (2.49) temos que $\langle F(x_R), x_R - x_0 \rangle > 0$, isto é uma contradição. Logo, como $\|x_R\| \leq R$, devemos ter que $\|x_R\| < R$. Portanto, pelo teorema 2.25, $x_R \in K$ é solução da VI (2.47). \square

2.2.5 Problema que conduzem a uma inequação variacional

Seja $K \subseteq \mathbb{R}^n$ um conjunto fechado e convexo e $f : K \rightarrow \mathbb{R}$ uma função de classe C^1 ; defino $F(x) = \nabla f(x)$.

Proposição 2.11. *Suponha que existe $x \in K$ tal que $f(x) = \min_{y \in K} f(y)$. Então, x é uma solução da inequação variacional*

$$x \in K : \quad \langle F(x), y - x \rangle \geq 0 \quad \forall y \in K.$$

Demonstração. Seja $y \in K$, como K é convexo. Então, $z = x + t(y - x) \in K$ para todo $t \in [0, 1]$. Portanto a função $\varphi(t) = f(x + t(y - x))$ com $0 \leq t \leq 1$, atinge seu mínimo quando $t = 0$, isto implica que $0 \leq \varphi'(0) = \langle \nabla f(x), y - x \rangle = \langle F(x), y - x \rangle$. \square

A recíproca cumpre-se se f é uma função convexa.

Proposição 2.12. *Se $f : K \rightarrow \mathbb{R}$ é uma função convexa e x satisfaz:*

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

Então, $f(x) = \min_{y \in K} f(y)$.

Demonstração. Desde que f é uma função convexa, pelo teorema 2.13, temos que: $f(y) \geq f(x) + \langle \nabla f(x), y - x \rangle \quad \forall y \in K$. Logo, $f(y) \geq f(x) + \langle F(x), y - x \rangle \quad \forall y \in K$; mas, pela hipótese $\langle F(x), y - x \rangle \geq 0 \quad \forall y \in K$. Assim, temos que $f(y) \geq f(x) \quad \forall y \in K$. Portanto, $f(x) = \min_{y \in K} f(y)$. \square

2.3 Introdução à teoria dos jogos

A teoria dos jogos é a teoria da tomada de decisões independente e interdependente de dois ou mais participantes autônomos (jogadores), onde nenhum jogador tem um controle total sobre os resultados, para maiores informações veja as referências [2], [9].

2.3.1 Jogos de estratégias

Um jogo de estratégias consiste em um conjunto de jogadores, onde para cada um deles há um conjunto de alternativas e para cada resultado (ou combinação de alternativas) do jogo, há um pagamento para cada jogador. Vamos começar fixando as ideias com um exemplo.

Exemplo 2.5. Suponhamos que British Airways (BA) e Virgin Atlantic (VA) estão pensando em fixar uma tarifa para a rota Londres-Paris. Se ambas aerolíneas cobrarem um preço de 800 \$ cada uma, o lucro de BA seria de 60 milhões e o lucro de VA de 90 milhões. Se BA cobra 800 \$ e VA cobra 500 \$, então BA teria uma perda de 30 milhões e

VA um lucro de 150 milhões. Se no entanto, BA fixa um preço de 500 \$ e VA cobra 800 \$, então o lucro de BA seria de 90 milhões e a de VA 30 milhões, e se ambas cobram um preço de 500 \$, então ambas aerolinhas acabam com perdas de 20 milhões. Esta informação pode ser representada na forma de uma tabela como mostramos abaixo.

Aerolinhas	Virgin Atlantic		
	Preços	800 \$	500 \$
British Airways	800 \$	60, 90	-30, 150
	500 \$	90, 30	-20, -20

No exemplo anterior há dois jogadores (aerolinhas), cada um tem estratégias dadas pelos preços e cada escolha conjunta que façam os jogadores tem um pagamento (um lucro ou uma perda). Vale ressaltar que seria melhor para ambas companhias aéreas coordenar as mudanças de preços porque, sem essa coordenação, as empresas aéreas acabariam fazendo perdas bastante graves. Em situações desse tipo, os seguintes três elementos sempre parecem estar presentes:

1. Dois ou mais participantes (jogadores).
2. Cada participante tem um conjunto de alternativas para eleger (estratégias).
3. Para cada escolha possível há um pagamento que cada jogador recebe.

Estes são os ingredientes essenciais que constituem o que é chamado um jogo de estratégias. Em linguagem mais formal, os jogos de estratégias consistem de um conjunto de jogadores, para cada jogador há um conjunto de estratégias, e para cada resultado (ou combinação de estratégias) do jogo há um pagamento para cada jogador.

Definição 2.21. Um jogo de estratégias consiste de n indivíduos denotados por $1, 2, \dots, n$ e chamados os jogadores tal que para cada $i = 1, 2, \dots, n$ o jogador i possui:

1. Um conjunto de alternativas S_i (também chamado como o conjunto de estratégias do jogador i), os elementos deste conjunto são chamados as estratégias do jogador i .
2. Uma função de pagamento $u_i : S_1 \times S_2 \times \dots \times S_n \rightarrow \mathbb{R}$.

Um jogo com estas características é denotado como uma $2n$ -tupla:

$$G = (S_1, S_2, \dots, S_n, u_1, u_2, \dots, u_n).$$

Exemplo 2.6. Este é um exemplo de um jogo de estratégias com três jogadores 1,2,3. Os conjuntos de estratégias dos jogadores são:

$$S_1 = S_2 = S_3 = [0, 1].$$

Suas funções de pagamento são dadas por

$$u_1(x, y, z) = x + y - z, \quad u_2(x, y, z) = x - yz \quad \text{e} \quad u_3(x, y, z) = xy - z,$$

onde, por simplicidade, deixamos $s_1 = x$, $s_2 = y$ e $s_3 = z$, com $s_i \in S_i$, para cada $i = 1, 2, 3$. Se os jogadores escolheram as estratégias $x = \frac{1}{2}$, $y = 0$ e $z = \frac{1}{4}$, então os seus pagamentos serão

$$u_1\left(\frac{1}{2}, 0, \frac{1}{4}\right) = \frac{1}{4}, \quad u_2\left(\frac{1}{2}, 0, \frac{1}{4}\right) = \frac{1}{2} \quad \text{e} \quad u_3\left(\frac{1}{2}, 0, \frac{1}{4}\right) = -\frac{1}{4}.$$

Observe que a tupla de estratégias $(1, 1, 0)$ (ver [2]) dá um melhor pagamento para cada jogador.

Existem dois tipos de jogos, os jogos cooperativos e os jogos não-cooperativos; Dizemos que um jogo é cooperativo quando existe uma certa comunicação ou algum tipo de coordenação entre os jogadores, os quais podem eleger a melhor opção para ambos. Os jogos não-cooperativos se caracterizam pela não existência de comunicação entre os jogadores. Um caso particular de jogos de estratégias são os chamados jogos de matrizes, eles são formalmente definidos da seguinte forma.

Definição 2.22. Um *jogo de matrizes* é um jogo de estratégias de dois jogadores tal que:

- (i) O jogador 1 tem um conjunto finito de estratégias S_1 com m elementos,
- (ii) O jogador 2 tem um conjunto finito de estratégias S_2 com n elementos, e
- (iii) Os pagamentos que recebem os jogadores são funções $u_1, u_2 : S_1 \times S_2 \rightarrow \mathbb{R}$ das estratégias conjuntas $(s_1, s_2) \in S_1 \times S_2$.

Se $S_1 = \{s_1^1, s_2^1, \dots, s_m^1\}$ e $S_2 = \{s_1^2, s_2^2, \dots, s_n^2\}$. Então, denotando $a_{ij} = u_1(s_i^1, s_j^2)$ e $b_{ij} = u_2(s_i^1, s_j^2)$ podemos representar estas estratégias e pagamentos numa matriz $m \times n$.

		Jogador 2			
		Estratégias	s_1^2	s_2^2	\dots
Jogador 1	s_1^1	a_{11}, b_{11}	a_{12}, b_{12}	\dots	a_{1n}, b_{1n}
	s_2^1	a_{21}, b_{21}	a_{22}, b_{22}	\dots	a_{2n}, b_{2n}
	\vdots	\vdots	\vdots	\dots	\vdots
	s_m^1	a_{m1}, b_{m1}	a_{m2}, b_{m2}	\dots	a_{mn}, b_{mn}

Definição 2.23. Sejam $G = (S_1, S_2, \dots, S_n, u_1, u_2, \dots, u_n)$ um jogo, e \bar{s}_i, \hat{s}_i duas estratégias do jogador i . Dizemos que a estratégia \bar{s}_i é *estritamente dominada* por \hat{s}_i , se para qualquer combinação de estratégias usados pelos outros jogadores, o pagamento para o jogador i ao usar a estratégia \bar{s}_i é menor que o pagamento ao usar a estratégia \hat{s}_i , isto é,

$$u_i(s_1, \dots, s_{i-1}, \bar{s}_i, s_{i+1}, \dots, s_n) < u_i(s_1, \dots, s_{i-1}, \hat{s}_i, s_{i+1}, \dots, s_n), \quad (2.51)$$

para qualquer $(s_1, \dots, s_{i-1}, s_{i+1}, \dots, s_n) \in S_1 \times \dots \times S_{i-1} \times S_{i+1} \times \dots \times S_n$.

Uma estratégia \hat{s}_i é dita estritamente dominante, se a desigualdade (2.51) vale para toda estratégia $\bar{s}_i \in S_i$.

Um jogador racional não vai jogar a estratégia dominada porque se ele escolhe uma estratégia estritamente dominada \bar{s}_i , ele poderia encontrar uma outra estratégia \hat{s}_i de modo que possa obter um pagamento maior.

Exemplo 2.7. [Dilema do prisioneiro] Dois indivíduos que cometem um crime tem a escolha de confessar o crime ou ficar em silêncio. No caso de que um deles confessar e o outro ficar em silêncio, então aquele que confessou não vai à prisão, enquanto aquele que não confessou obtém uma sentença de dez anos. No caso de ambos confessarem, cada um recebe uma sentença de cinco anos. Se ambos não confessarem, ambos obtêm uma condenação de um ano. Ao examinar o jogo, observa-se as seguintes características:

Jogadores	Prisioneiro 2		
	Estratégias	Calar	Confessar
Prisioneiro 1	Calar	-1, -1	-10, 0
	Confessar	0, -10	-5, -5

- Ambos jogadores tem interesse em manter silêncio, já que neste caso ambos recebem uma condenação de um ano cada um.
- Uma vez que um jogador vai manter silêncio, o outro jogador tem um incentivo para confessar o crime.

Estes são precisamente o tipo de paradoxos que são tão inerentes aos jogos. A questão central não é apenas sobre a escolha que um jogador faz, mas também sobre as escolhas dos outros jogadores. Um exame aprofundado do jogo mostra que, se o jogador 1 usa a estratégia "confessar"(co), então ele ganha um melhor pagamento para cada escolha que o jogador 2 faz. Para ver isso, deixe u_1 indicar a função de pagamento do jogador 1 e observe que, se o jogador 2 escolhe "calar"(ca), então

$$u_1(co, ca) = 0 > -1 = u_1(ca, ca),$$

enquanto que se o jogador 2 escolhe "confessar"(co), então

$$u_1(co, co) = -5 > -10 = u_1(ca, co)$$

Ou seja, não importa qual seja a escolha do jogador 2, é melhor para o jogador 1 jogar a estratégia confessar. Dizemos que a estratégia do jogador 1 confessar é uma estratégia estritamente dominante (ver definição 2.23). Um exame semelhante das estratégias do jogador 2 revela que a estratégia confessar é uma estratégia estritamente dominante para o jogador 2. Na ausência de qualquer comunicação ou qualquer esquema de coordenação,

os jogadores racionais devem desempenhar suas estratégias estritamente dominantes, Uma solução para o "Dilema do Prisioneiro", poderia, portanto, acabar por ser (confessar, confessar). Esta é a solução usando estratégias estritamente dominantes.

2.3.2 Equilíbrios de Nash

Quando um jogo de estratégias é jogado, o objetivo de um jogador é maximizar seu pagamento. Um jogador individual pode, se ele ou ela conhece as escolhas dos outros jogadores, optar por maximizar seu pagamento dado as escolhas dos outros. Mas, então, todos os outros jogadores querem fazer o mesmo. Na verdade, parece bastante natural procurar um resultado que resulte da maximização simultânea de pagamentos individuais. Esse perfil de estratégias é geralmente chamado um equilíbrio de Nash e é definido como segue.

Definição 2.24. Um equilíbrio de Nash (ou simplesmente uma solução) de um jogo de estratégias

$$G = (S_1, S_2, \dots, S_n, u_1, u_2, \dots, u_n)$$

é um perfil de estratégias $(s_1^*, s_2^*, \dots, s_n^*) \in S_1 \times S_2 \times \dots \times S_n$, de modo que para cada jogador $i = 1, \dots, n$, cumpre-se

$$u_i(s_1^*, \dots, s_{i-1}^*, s_i^*, s_{i+1}^*, \dots, s_n^*) \geq u_i(s_1^*, \dots, s_{i-1}^*, s_i, s_{i+1}^*, \dots, s_n^*) \quad \forall s_i \in S_i.$$

De forma equivalente, para cada $i = 1, \dots, n$, s_i^* resolve o problema de otimização

$$\begin{aligned} & \text{Maximizar} && u_i(s_1^*, \dots, s_{i-1}^*, s_i, s_{i+1}^*, \dots, s_n^*) \\ & \text{sujeito a} && s_i \in S_i. \end{aligned}$$

Observe que um ponto de equilíbrio de Nash é equivalente a maximizar n funções. Além disso, um equilíbrio de Nash é uma tupla de estratégias conjuntas do jogo, tal que nenhum dos jogadores terá um incentivo a mudar de estratégia uma vez jogou o jogo. Assim, se todos os jogadores sabem que todos os outros estão comprometidos em jogar um equilíbrio de Nash, então todo mundo vai concordar em jogar sua estratégia de equilíbrio de Nash, pela simples razão de que é ideal para todos eles, no sentido que se um jogador muda de estratégia e os outros não, então o pagamento que receberá será menor ou igual ao correspondente ao que é dado pelo equilíbrio de Nash. Os equilíbrios de Nash tem sido amplamente utilizado em aplicações da teoria dos jogos. Uma razão para esta popularidade é que quando você olha para uma estratégia que não é um equilíbrio de Nash, então há pelo menos um jogador que pode melhorar o seu pagamento em detrimento de outro jogador. Um jogo pode ter um equilíbrio de Nash, nenhum ou mais de um equilíbrio de Nash. No exemplo 2.7 do dilema do prisioneiro a estratégia (confessar, confessar) é o unico equilíbrio de Nash (para mais detalhes ver [2]) e seus pagamentos respectivos são $(-5, -5)$. A seguir mostramos um exemplo onde existem mais de um equilíbrio de Nash.

		Jogador 2			
		Estratégias	α	β	γ
Jogador 1	a	1, 0	1, 1	2, 1	3, 1
	b	4, 2	0, 2	1, 1	2, 2
	c	-1, 1	1, 1	1, 2	2, 0
	d	3, 2	-1, 1	2, 2	-3, 2
	e	2, 1	-2, 2	1, -1	1, 5

Exemplo 2.8. No seguinte jogo de matrizes, cada jogador $i = 1, 2$ tem um conjunto finito de estratégias com 5 e 4 elementos respectivamente. Usando o método de estratégias estritamente dominadas tem-se que (b, α) , (a, β) , (a, γ) , (d, γ) e (a, δ) são 5 equilíbrios de Nash do Jogo de matrizes.

A seguir vemos um exemplo onde não existe nenhum ponto de equilíbrio de Nash.

Exemplo 2.9. Considere o seguinte jogo de dois jogadores com conjuntos de estratégias $S_1 = S_2 = [0, 1]$ e funções de pagamento $u_1(x, y) = \left(y - \frac{1}{2}\right)x$ e $u_2(x, y) = |x - y|$. Este jogo não tem nenhum equilíbrio de Nash. De fato. Suponhamos que existe um equilíbrio de Nash $(x^*, y^*) \in S_1 \times S_2$. Então, pela definição 2.24 de equilíbrio de Nash o ponto (x^*, y^*) satisfaz as duas desigualdades seguintes:

$$u_1(x^*, y^*) \geq u_1(x, y^*) \quad \forall x \in [0, 1], \quad (2.52)$$

$$u_2(x^*, y^*) \geq u_2(x^*, y) \quad \forall y \in [0, 1]. \quad (2.53)$$

A partir de (2.52) e usando a definição de u_1 tem-se que

$$\left(y^* - \frac{1}{2}\right)x^* \geq \left(y^* - \frac{1}{2}\right)x \quad \forall x \in [0, 1],$$

podem ocorrer três casos:

- (i) Se $y^* = \frac{1}{2}$, então o ponto $\left(x^*, \frac{1}{2}\right)$ é um candidato a equilíbrio de Nash, para todo $x^* \in [0, 1]$.
- (ii) Se $y^* > \frac{1}{2}$, então $x^* \geq x \quad \forall x \in [0, 1]$, isto implica que $x^* = 1$. Logo, o ponto $(1, y^*)$ é um candidato a equilíbrio de Nash, para todo $y^* \in \left(\frac{1}{2}, 1\right]$.
- (iii) Se $y^* < \frac{1}{2}$, então $x^* \leq x \quad \forall x \in [0, 1]$, isto implica que $x^* = 0$. Logo, o ponto $(0, y^*)$ é um candidato a equilíbrio de Nash, para todo $y^* \in \left[0, \frac{1}{2}\right)$.

Além disso, estes pontos candidatos a equilíbrios de Nash devem satisfazer a desigualdade (2.53) a qual é equivalente a:

$$|x^* - y^*| \geq |x^* - y| \quad \forall y \in [0, 1] \quad (2.54)$$

Mas nenhum deles satisfazem a relação (2.54). Portanto, não existe nenhum ponto de equilíbrio de Nash.

Existe um critério útil para encontrar um equilíbrio de Nash de um jogo de estratégias quando os conjuntos de estratégias são intervalos abertos de números reais (ver [2]). Nesse caso, é fácil ver que se uma tupla de estratégias (s_1^*, \dots, s_n^*) é o equilíbrio de Nash do jogo, então deve ser uma solução do sistema de equações.

$$\frac{\partial u_i(s_1^*, \dots, s_n^*)}{\partial s_i} = 0, \quad \forall i = 1, 2, \dots, n. \quad (2.55)$$

Portanto, o equilíbrio de Nash está entre as soluções do sistema (2.55). Quando o sistema (2.55) tem uma solução única, então esse ponto é o único equilíbrio de Nash do jogo. Este é essencialmente o teste para determinar o equilíbrio de Nash em jogos de estratégias, cujos conjuntos de estratégias são intervalos abertos. Para maiores detalhes veja a referência [2]. Em termos matemáticos precisos, esta é formulada da seguinte forma:

Proposição 2.13. *Seja G um jogo de estratégias, cujos conjuntos de estratégias são intervalos abertos e com funções pagamento duas vezes diferenciáveis. Suponha que uma tupla de estratégias (s_1^*, \dots, s_n^*) satisfaça:*

1. $\frac{\partial u_i(s_1^*, \dots, s_n^*)}{\partial s_i} = 0$, para cada jogador $i = 1, 2, \dots, n$.
2. Cada s_i^* é o ponto estacionário da função

$$u_i(s_1^*, \dots, s_{i-1}^*, s, s_{i+1}^*, \dots, s_n^*), \quad s \in S_i.$$

3. $\frac{\partial^2 u_i(s_1^*, \dots, s_n^*)}{\partial^2 s_i} < 0$ para cada $i = 1, \dots, n$.

Então, (s_1^*, \dots, s_n^*) é um equilíbrio de Nash do jogo G .

Na prática, geralmente descobrimos a solução do sistema (2.55) e usamos outras considerações econômicas para verificar se a solução é o equilíbrio de Nash do jogo. Aqui está um exemplo que ilustra o teste de Equilíbrio de Nash.

Exemplo 2.10. Considere um jogo de estratégias de três pessoas em que cada jogador tenha um conjunto de estratégias igual ao intervalo aberto $(0, +\infty)$. As funções pagamento dos jogadores são dadas por

$$\begin{aligned} u_1(x, y, z) &= 2xz - x^2y \\ u_2(x, y, z) &= \sqrt{12(x+y+z)} - y \\ u_3(x, y, z) &= 2z - xyz^2. \end{aligned}$$

Para encontrar o equilíbrio de Nash do jogo, devemos resolver o sistema de equações

$$\frac{\partial u_1}{\partial x} = 0, \quad \frac{\partial u_2}{\partial y} = 0 \quad \text{e} \quad \frac{\partial u_3}{\partial z} = 0.$$

Tomando derivadas, obtemos

$$\frac{\partial u_1}{\partial x} = 2z - 2xy, \quad \frac{\partial u_2}{\partial y} = \sqrt{\frac{3}{x+y+z}} - 1 \quad \text{e} \quad \frac{\partial u_3}{\partial z} = 2 - 2xyz.$$

Ou, simplificando as equações,

$$\begin{aligned} z &= xy \\ x + y + z &= 3 \\ xyz &= 1. \end{aligned}$$

Assim, a única solução do sistema de equações é $x = y = z = 1$.

$$\begin{aligned} \frac{\partial^2 u_1}{\partial x^2} &= -2y < 0, \\ \frac{\partial^2 u_2}{\partial y^2} &= -\frac{\sqrt{3}}{2} (x+y+z)^{-3/2} < 0, \\ \frac{\partial^2 u_3}{\partial z^2} &= -2xy < 0. \end{aligned}$$

Para todas as escolhas $x > 0$, $y > 0$ e $z > 0$. O teste de equilíbrio de Nash garante que $(1, 1, 1)$ é o único equilíbrio de Nash do jogo.

3 INEQUAÇÕES VARIACIONAIS E PROBLEMAS DE COMPLEMENTARIDADE

Neste capítulo, definimos formalmente os problemas que formam o tema principal deste trabalho: a inequação (ou desigualdade) variacional (VI) e o problema de complementaridade (CP). O corpo principal do presente capítulo consiste em duas partes. Na primeira parte, apresentamos algumas classificações básicas e terminologia associada para vários casos especiais desses problemas. Nós explicamos a interconexão entre a VI e o CP, bem como sua relação com um programa de otimização não-linear padrão. Na segunda parte, apresentamos a relação entre VI e / ou CP com os equilíbrios de Nash. Para maiores informações ver as referências [3], [10].

3.1 Descrição do Problema

O exemplo mais simples de uma inequação variacional é o problema clássico de resolver um sistema de equações não-lineares. De fato, como veremos em breve, esse problema pode ser considerado como uma VI sem restrições. Como vimos no capítulo anterior (ver definição 2.20), dado um conjunto K de \mathbb{R}^n e uma aplicação $F : K \rightarrow \mathbb{R}^n$, a inequação variacional denotada por $VI(K, F)$, é encontrar um vetor $x \in K$ tal que

$$\langle F(x), y - x \rangle \geq 0 \quad \forall y \in K. \quad (3.1)$$

O conjunto de soluções deste problema é denotado por $SOL(K, F)$. Como já mencionamos, nós estamos interessados na situação em que o conjunto K é fechado e a função F é contínua. Na maioria dos casos de $VI(K, F)$ discutidas no trabalho, o conjunto K é convexo. No entanto, matematicamente, alguns resultados não exigem a convexidade de K . Assim, não fazemos a convexidade de K uma suposição geral.

Como K é fechado e F é contínua, segue-se que $SOL(K, F)$ é sempre um conjunto fechado (embora possa estar vazio). Compreender outras propriedades do conjunto de soluções de uma VI é um tema importante que tem significado tanto teórico como prático.

Observação 3.1. Uma interpretação geométrica de uma VI, e mais especificamente da inequação definida por (3.1) é que um ponto $x \in K$ é uma solução da $VI(K, F)$, se e somente se, $F(x)$ forma um ângulo não obtuso com cada vetor da forma $y - x$ para todo $y \in K$.

Podemos formalizar esta observação usando o conceito de cone normal, como vimos na definição 2.13. Dado o conjunto $K \subseteq \mathbb{R}^n$ e qualquer vetor $x \in K$, o cone normal a K em x ; é o conjunto:

$$\mathcal{N}(x, K) = \{d \in \mathbb{R}^n : \langle d, y - x \rangle \leq 0 \quad \forall y \in K\}. \quad (3.2)$$

Observação 3.2. A desigualdade (3.1), diz claramente que um vetor $x \in K$ resolve a $VI(K, F)$ se e somente se $-F(x)$ é um vetor normal para K em x ; ou, de forma equivalente,

$$0 \in F(x) + \mathcal{N}(x, K).$$

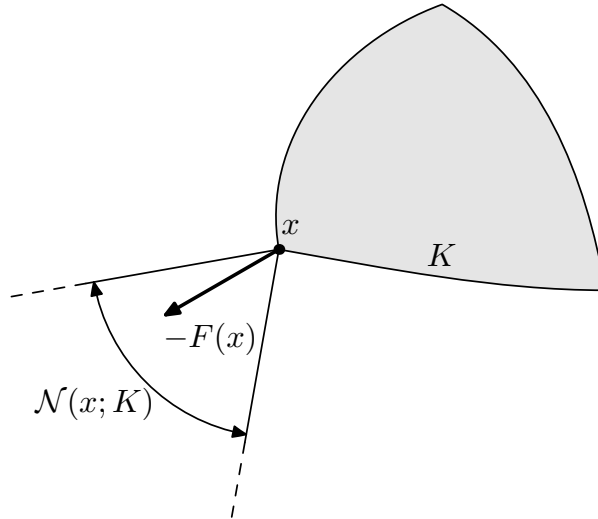


Figura 11 – O cone normal e solução de uma VI.

Observação 3.3. Se $K = \mathbb{R}^n$ então um vetor $x \in SOL(\mathbb{R}^n, F)$ se e somente se $F(x) = 0$. De fato: suponha que $x \in SOL(\mathbb{R}^n, F)$. Então, $\langle F(x), y - x \rangle \geq 0 \quad \forall y \in \mathbb{R}^n$, em particular considere $y = x - F(x)$. Temos que $\langle F(x), -F(x) \rangle \geq 0$. Consequentemente, $\|F(x)\|^2 \leq 0$. Logo, $F(x) = 0$. Reciprocamente se $F(x) = 0$, é claro que $x \in SOL(\mathbb{R}^n, F)$.

O argumento acima se aplica de forma mais geral a uma solução da $VI(K, F)$ que pertence ao interior topológico de $K \subseteq \mathbb{R}^n$. Especificamente, se x é uma solução da VI (3.1) e x pertence a $\text{int}(K)$; então, $F(x) = 0$. De fato, desde que $x \in \text{int}(K)$, existe $\delta > 0$ tal que $B(x, \delta) \subseteq K$; por outro lado, suponha que $F(x) \neq 0$, considere $\tau > 0$ tal que:

$$0 < \tau < \frac{\delta}{\|F(x)\|}.$$

Logo, fazendo $y = x - \tau F(x)$, temos que $\|y - x\| = \tau \|F(x)\| < \delta$. Assim, $y \in B(x, \delta) \subseteq K$ e como $x \in SOL(K, F)$ usando (3.1) tem-se:

$$-\tau \langle F(x), F(x) \rangle \geq 0$$

Então, $\|F(x)\|^2 \leq 0$ e portanto $F(x) = 0$.

Em todas as realizações interessantes de uma VI, é invariavelmente o caso de que nenhum dos zeros de F , se houver, são soluções da $VI(K, F)$. Em outras palavras, a VI é uma generalização genuinamente não trivial do problema clássico da resolução de equações. No

entanto, como vemos ao longo do trabalho, a teoria e os métodos para resolver equações são fundamentais para a análise e solução de uma VI.

Quando K é um cone (ver definição 2.4), a VI admite uma forma equivalente conhecida como um problema de complementaridade.

Definição 3.1. Dado um cone $K \subseteq \mathbb{R}^n$ e uma aplicação $F : K \rightarrow \mathbb{R}^n$ o problema de complementaridade, denotado por $CP(K, F)$ consiste em encontrar um vetor $x \in \mathbb{R}^n$ que satisfaz as seguintes condições:

$$(i) \ x \in K \quad (ii) \ -F(x) \in K^* \quad (iii) \ \langle F(x), x \rangle = 0$$

Onde K^* é o cone dual de K como vimos na definição 2.9.

A conexão precisa entre a $VI(K, F)$ e o $CP(K, F)$ quando K é um cone é descrito no seguinte resultado, ver [3].

Proposição 3.1. *Seja K um cone de \mathbb{R}^n . Um vetor x resolve a $VI(K, F)$ se e somente se, x resolve o $CP(K, F)$.*

Demonstração. Suponha que x resolve a $VI(K, F)$, é claro que $x \in K$ e desde que K é um cone tem-se que $0 \in K$, tomando $y = 0$ em (3.1), obtemos:

$$\langle F(x), x \rangle \leq 0$$

Além disso, como K é um cone e $x \in K$, segue-se que $2x \in K$. Logo, considerando $y = 2x$ em (3.1), obtemos:

$$\langle F(x), x \rangle \geq 0$$

Agora, das duas desigualdades acima, deduzimos que $\langle F(x), x \rangle = 0$. Por sua vez, de (3.1) tem-se:

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

Mas, $\langle F(x), x \rangle = 0$. Logo, $\langle F(x), y \rangle \geq 0 \quad \forall y \in K$; então, $-F(x) \in K^*$. Portanto, x resolve o $CP(K, F)$.

Reciprocamente se x resolve o $CP(K, F)$, então $x \in K$, $-F(x) \in K^*$ e $\langle F(x), x \rangle = 0$. Assim, dado qualquer $y \in K$, temos que:

$$\langle F(x), y - x \rangle = \langle F(x), y \rangle - \langle F(x), x \rangle = \langle F(x), y \rangle \geq 0,$$

pois $-F(x) \in K^*$. Portanto x resolve a $VI(K, F)$. □

Uma palavra sobre a notação $CP(K, F)$: quando a sigla CP é anexada ao par (K, F) , entende-se que K é um cone.

O $CP(K, F)$ é definido por três condições: (i) $x \in K$ (ii) $-F(x) \in K^*$ e (iii) $\langle F(x), x \rangle = 0$. Introduzimos alguns conceitos associados a vetores que satisfazem as duas primeiras condições.

Definição 3.2. Seja um vetor $x \in \mathbb{R}^n$:

(i) Dizemos que $x \in \mathbb{R}^n$ é factível (ou viável) ao $CP(K, F)$, se e somente se:

$$x \in K \quad \text{e} \quad -F(x) \in K^*. \quad (3.3)$$

(ii) Dizemos que $x \in \mathbb{R}^n$ é estritamente factível ao $CP(K, F)$, se e somente se $x \in K$ e $-F(x) \in \text{int}(K^*)$.

(iii) Dizemos que o $CP(K, F)$ é (estritamente) factível, se e somente se tem um vetor (estritamente) factível.

(iv) A região factível de o $CP(K, F)$ é o conjunto de todos os vetores factíveis e é denotada por $FEA(K, F)$. É claro que $SOL(K; F) \subseteq FEA(K, F)$.

Para uma função não-linear F , encontrar um ponto viável ao $CP(K, F)$ ou determinar se nenhum desses pontos existe não é necessariamente mais fácil computacionalmente do que resolver o próprio problema de complementaridade. No entanto, a compreensão da viabilidade do CP é importante porque, afinal, a viabilidade é uma condição necessária para a solvabilidade. Quando F é uma função afim, existem pelo menos dois casos em que a determinação da viabilidade do $CP(K, F)$ é presumivelmente mais fácil do que resolver o problema em si. O primeiro caso é quando K é um cone poliedral. Neste caso, a condição de viabilidade (3.3) é essencialmente um sistema de desigualdades lineares; Portanto, sua viabilidade pode ser determinada pela resolução de um programa linear (LP). O outro caso é quando K é um cone especial, conhecido como o cone de matrizes semi-finitas.

Definimos o ortante não negativo de \mathbb{R}^n , como o conjunto:

$$\mathbb{R}_+^n = \{x = (x_1, \dots, x_n) \in \mathbb{R}^n : x_i \geq 0 \quad \forall i = 1, \dots, n\}$$

Muitos casos especiais do problema de complementaridade são muito importantes na modelagem. Agora apresentamos dois dos mais importantes. Quando K é o ortante não negativo de \mathbb{R}^n , o $CP(K, F)$ é conhecido como o problema de complementaridade não linear e denotado por $NCP(F)$. Assim, temos a seguinte definição.

Definição 3.3. Dada uma aplicação $F : \mathbb{R}_+^n \rightarrow \mathbb{R}^n$ o $NCP(F)$ consiste em encontrar um vetor $x \in \mathbb{R}^n$, que satisfaça

$$0 \leq x \perp F(x) \geq 0. \quad (3.4)$$

Ao expressar a condição de ortogonalidade $\langle F(x), x \rangle = 0$ em termos de produtos de componente (o que é justificado porque x e $F(x)$ são ambos vetores não negativos), obtemos a seguinte formulação equivalente do $NCP(F)$:

$$\begin{aligned} 0 \leq x, \quad F(x) \geq 0 \\ x_i F_i(x) = 0, \quad \forall i = 1, \dots, n. \end{aligned}$$

A generalização do NCP é o problema de complementaridade mista, abreviado como MiCP. Isso corresponde ao caso do $CP(K, F)$ onde K é o cone especial $\mathbb{R}^{n_1} \times \mathbb{R}_+^{n_2}$, com $n_1 + n_2 = n$. Assim, chegamos à seguinte definição.

Definição 3.4. Sejam $G : \mathbb{R}^{n_1} \times \mathbb{R}_+^{n_2} \rightarrow \mathbb{R}^{n_1}$ e $H : \mathbb{R}^{n_1} \times \mathbb{R}_+^{n_2} \rightarrow \mathbb{R}^{n_2}$ duas aplicações. O $MiCP(G, H)$ consiste em encontrar um par de vetores $(u, v) \in \mathbb{R}^{n_1} \times \mathbb{R}_+^{n_2}$ tal que:

$$\begin{aligned} G(u, v) &= 0, \quad u \text{ livre} \\ 0 &\leq v \perp H(u, v) \geq 0. \end{aligned}$$

3.1.1 Problemas Afins

O $CP(K, F)$ é um caso especial da $VI(K, F)$ onde o conjunto K é um cone. No que se segue, apresentamos vários outros casos especiais da $VI(K, F)$ onde K ou F possui algumas outras estruturas interessantes. Para começar, seja F uma função afim dada por:

$$F(x) = q + Mx \quad \forall x \in \mathbb{R}^n \quad (3.5)$$

para algum vetor $q \in \mathbb{R}^n$ e uma matriz $M \in \mathbb{R}^{n \times n}$. Neste caso escrevemos $VI(K, q, M)$ em vez de $VI(K, F)$. O conjunto solução da $VI(K, q, M)$ é denotado por $SOL(K, q, M)$. Se além disso K é um conjunto poliedral (ver (2.5)), a $VI(K, q, M)$ denotada por $AVI(K, q, M)$ será chamada VI afim. Finalmente se K é um conjunto poliedral, mas F não é necessariamente afim, usamos o adjetivo "restrições lineares" para descrever a $VI(K, F)$. A diferença da AVI onde tanto o conjunto da definição e a função tem estrutura afim, a $VI(K, q, M)$ e a VI com restrições lineares tem só um membro afim no par (K, F) . Uma importante VI com restrições lineares é a caixa restrita $VI(K, F)$ onde o conjunto K é um retângulo fechado, dado por:

$$K = \{x \in \mathbb{R}^n : a_i \leq x_i \leq b_i, \quad i = 1, \dots, n\} \quad (3.6)$$

Em particular, permitimos que alguns limites inferiores a_i sejam iguais a $-\infty$ e alguns limites superiores b_i iguais a $+\infty$, este marco estendido inclui como um caso particular o $NCP(F)$, o qual corresponde fazer a igual zero e b o vetor de todas as componentes que são iguais a $+\infty$.

Similarmente o MiCP é um caso especial da VI caixa restrita, um $MiCP(G, H)$ onde G e H são ambas funções afins é chamado um problema de complementaridade linear mista, abreviado como MLCP. Um NCP com uma função afim definida é chamado um problema de complementaridade linear, abreviado como LCP. Notacionalmente se F é a função afim dada por (3.5), o $NCP(F)$ se escreve como $LCP(q, M)$.

3.2 Relações entre classes de problemas

Resumimos as várias classes de problemas e suas conexões no diagrama abaixo. A notação $P \Rightarrow Q$ significa que podemos derivar o problema Q , a partir do problema P ; a

notação $P \Leftrightarrow Q$ significa que os dois problemas P e Q são "equivalentes".

$$\begin{array}{ccccc}
 VI(K, q, M) & & & & VI(K, q, M) \\
 \uparrow & & & & \downarrow \\
 VI(K, F) & \Rightarrow & VI \text{ restrições lineares} & \Rightarrow & AVI(K, q, M) \\
 \downarrow & & \Downarrow & & \Downarrow \\
 CP(K, F) & \Rightarrow & MiCP(F) & \Rightarrow & MLCP \\
 & & \downarrow & & \downarrow \\
 & & NCP(F) & \Rightarrow & LCP(q, M)
 \end{array}$$

Todas, exceto duas relações no diagrama acima seguem facilmente das definições dos respectivos problemas. As exceções são a equivalência entre uma VI com restrições lineares e um MiCP e a equivalência entre uma AVI e um MLCP. Na verdade, a última equivalência é uma consequência imediata do primeiro, considerando um função afim F . No seguinte resultado, consideramos o conjunto poliedral

$$K = \{x \in \mathbb{R}^n : Ax \leq b, Cx = d\} \quad (3.7)$$

Para algumas matrizes dadas $A \in \mathbb{R}^{m \times n}$, $C \in \mathbb{R}^{l \times n}$ e vetores $b \in \mathbb{R}^m$ e $d \in \mathbb{R}^l$. O equivalente MiCP resultante da $VI(K, F)$ que se chama o sistema de Karush - Kuhn - Tucker (KKT) da $VI(K, F)$ depende da representação (3.7), para uma demonstração alternativa da proposição 3.2 veja a referência [3].

Proposição 3.2. *Seja $K \subseteq \mathbb{R}^n$ dado por (3.7). Um vetor x resolve a $VI(K, F)$ se e somente se existem vetores $\lambda \in \mathbb{R}^m$ e $\mu \in \mathbb{R}^l$ tal que:*

$$\begin{aligned}
 0 &= F(x) + C^T \mu + A^T \lambda \\
 0 &= d - Cx \\
 0 &\leq \lambda \perp b - Ax \geq 0
 \end{aligned} \quad (3.8)$$

Demonstração. Começamos fazendo uma observação sobre a $VI(K, F)$; a desigualdade (3.1) é equivalente a:

$$\langle F(x), y \rangle \geq \langle F(x), x \rangle \quad \forall y \in K.$$

Assim, um vetor x é solução da $VI(K, F)$ se e somente se x é uma solução do problema de otimização na variável y (com x considerado fixo):

$$\begin{aligned}
 &\text{Minimizar} \quad \langle F(x), y \rangle \\
 &\text{sujeita a} \quad y \in K.
 \end{aligned} \quad (3.9)$$

Assim, se x é uma solução da $VI(K, F)$. Então, pela observação anterior tem-se que $x \in K$ é um minimizador do problema (3.9) com restrições lineares de igualdade e desigualdade.

Logo, pelo teorema 2.17 (ou também ver (2.35)-(2.39)) existem $(\mu, \lambda) \in \mathbb{R}^l \times \mathbb{R}^m$ tais que:

$$\begin{aligned} \nabla f(x) + \sum_{i=1}^l \mu_i \nabla h_i(x) + \sum_{i=1}^m \lambda_i \nabla g_i(x) &= 0 \\ h(x) &= 0 \\ \lambda_i g_i(x) &= 0 \quad \forall i = 1, \dots, m \\ \lambda_i &\geq 0 \quad \forall i = 1, \dots, m \\ g_i(x) &\leq 0 \quad \forall i = 1, \dots, m \end{aligned}$$

Onde $f(y) = \langle F(x), y \rangle$, $h(y) = Cy - d$ e $g(y) = Ay - b$. Assim, temos que $h_i(y) = c_i^T y - d_i$, $g_i(y) = a_i^T y - b_i$, onde $c_i, a_i \in \mathbb{R}^n$ são os vetores filas das matrizes $C \in \mathbb{R}^{l \times n}$ e $A \in \mathbb{R}^{m \times n}$ respectivamente. Então, dado qualquer $y \in \mathbb{R}^n$ tem-se $\nabla f(y) = F(x)$, $\nabla h_i(y) = c_i^T$ e $\nabla g_i(y) = a_i^T$. Consequentemente, temos que:

$$\begin{aligned} 0 &= F(x) + C^T \mu + A^T \lambda \\ 0 &= d - Cx \\ 0 &\leq \lambda \perp b - Ax \geq 0 \end{aligned}$$

Reciprocamente, sendo K um conjunto poliedral, pelo corolário 2.12, temos que K é convexo. Assim, o problema (3.9) é um problema de programação linear convexo na variável y , e pela hipóteses x é um ponto KKT do problema (3.9) (ou ponto estacionário do problema (3.9)), isto é, verifica-se o sistema (2.35)-(2.39). Então, pelo teorema 2.18 e o teorema 2.9, x é um minimizador global do problema (3.9). Logo, x resolve a $VI(K, F)$. \square

Exemplo 3.1. Para ilustrar a proposição 3.2, consideremos a $VI(K, F)$ caixa restrita, onde $K = \{x \in \mathbb{R}^n : a_i \leq x_i \leq b_i \quad \forall i = 1, \dots, n\}$; desde que K é um conjunto poliedral, a $VI(K, F)$ é equivalente a suas condições KKT, que podem ser escritas como:

$$\begin{aligned} 0 &\leq x - a \perp F(x) + v \geq 0 \\ 0 &\leq b - x \perp v \geq 0. \end{aligned}$$

De fato: $K = \{x \in \mathbb{R}^n : a \leq x \leq b\}$ pode-se escrever como: $K = \{x \in \mathbb{R}^n : \bar{A}x \leq \bar{b}\}$ onde: $\bar{A} = \begin{bmatrix} I \\ -I \end{bmatrix}$ e $\bar{b} = \begin{bmatrix} b \\ -a \end{bmatrix}$, sendo I a matriz identidade $n \times n$. Assim, K é um conjunto poliedral só com restrições de desigualdade. Logo, se $x \in K$ é uma solução da $VI(K, F)$, da proposição 3.2, existe $\lambda = \begin{bmatrix} v \\ w \end{bmatrix} \in \mathbb{R}^{2n}$ com $v, w \in \mathbb{R}^n$ e $v, w \geq 0$ tal que:

$$0 = F(x) + \bar{A}^T \lambda \tag{3.10}$$

$$0 \leq \lambda \perp \bar{b} - \bar{A}x \geq 0 \tag{3.11}$$

De (3.10), temos que $0 = F(x) + [I, -I] \begin{bmatrix} v \\ w \end{bmatrix}$; isto é, $0 = F(x) + Iv - Iw$. Logo, $w = F(x) + v$; por outro lado:

$$\bar{b} - \bar{A}x = \begin{bmatrix} b \\ -a \end{bmatrix} - \begin{bmatrix} I \\ -I \end{bmatrix} x = \begin{bmatrix} b - x \\ -a + x \end{bmatrix}.$$

Assim:

$$\lambda \perp \bar{b} - \bar{A}x, \text{ se e somente se } \begin{bmatrix} v \\ w \end{bmatrix} \perp \begin{bmatrix} b - x \\ -a + x \end{bmatrix}. \text{ Então, } v(b - x) + w(x - a) = 0,$$

onde $b - x \geq 0$, $x - a \geq 0$, $v \geq 0$ e $w \geq 0$. Então, $v(b - x) = 0$ e $w(x - a) = 0$; isto é, $b - x \perp v$ e $x - a \perp w$. Assim, usando (3.11) temos que $0 \leq b - x \perp v \geq 0$ e $0 \leq x - a \perp w \geq 0$ e como $w = F(x) + v$, tem-se:

$$\begin{aligned} 0 \leq x - a \perp F(x) + v \geq 0 \\ 0 \leq b - x \perp v \geq 0. \end{aligned}$$

Em geral um ponto $x \in K$ é uma solução de uma caixa restrita $VI(K, F)$ (ver [3]), se e somente se, para cada $i = 1, \dots, n$ temos:

$$\begin{aligned} x_i = a_i &\implies F_i(x) \geq 0, \\ a_i < x_i < b_i &\implies F_i(x) = 0, \\ x_i = b_i &\implies F_i(x) \leq 0. \end{aligned}$$

Associada com uma $VI(K, F)$ com restrições lineares, se define a função (não linear) aumentada \tilde{F} da seguinte maneira. Para todo x no domínio de F e $(\mu, \lambda) \in \mathbb{R}^{l+m}$

$$\tilde{F}(x, \mu, \lambda) = \begin{pmatrix} F(x) + C^T \mu + A^T \lambda \\ d - Cx \\ b - Ax \end{pmatrix} = \begin{pmatrix} G(x, \mu, \lambda) \\ H(x, \mu, \lambda) \end{pmatrix} \quad (3.12)$$

Onde $G : \mathbb{R}^{n+l} \times \mathbb{R}^m \rightarrow \mathbb{R}^{n+l}$ e $H : \mathbb{R}^{n+l} \times \mathbb{R}^m \rightarrow \mathbb{R}^m$ estão definidas como:

$$G(x, \mu, \lambda) = \begin{pmatrix} F(x) + C^T \mu + A^T \lambda \\ d - Cx \end{pmatrix} \text{ e } H(x, \mu, \lambda) = b - Ax$$

Da proposição 3.2, concluímos que $x \in SOL(K, F)$ se e somente se existem um par de vetores (μ, λ) que são chamados de multiplicadores, tais que (x, μ, λ) resolve o $MiCP(G, H) = MiCP(\tilde{F})$. Esta é precisamente a afirmação da equivalência entre uma $VI(K, F)$ com restrições lineares e um MiCP. Consideremos o caso afim no que F é dada por $F(x) = q + Mx$ a função não linear:

$$\tilde{F}(x, \mu, \lambda) = \begin{pmatrix} q + Mx + C^T \mu + A^T \lambda \\ d - Cx \\ b - Ax \end{pmatrix} = \begin{pmatrix} q \\ d \\ b \end{pmatrix} + \begin{pmatrix} Mx + C^T \mu + A^T \lambda \\ -Cx \\ -Ax \end{pmatrix}$$

Logo:

$$\tilde{F}(x, \mu, \lambda) = \begin{pmatrix} q \\ d \\ b \end{pmatrix} + Q \begin{pmatrix} x \\ \mu \\ \lambda \end{pmatrix} \in \mathbb{R}^{n+l+m}$$

onde:

$$Q = \begin{bmatrix} M & C^T & A^T \\ -C & 0 & 0 \\ -A & 0 & 0 \end{bmatrix}$$

Então, segue-se que a $AVI(K, q, M)$ é equivalente ao MLCP definido por $(\tilde{F}(0), Q)$, portanto, nós estabelecemos a equivalencia entre uma AVI e um determinado aumento de MLCP. Verificar a viabilidade de este ultimo MLCP equivale a verificar a consistência do seguinte sistema de desigualdade linear:

$$\begin{aligned} 0 &= q + Mx + C^T\mu + A^T\lambda \\ 0 &= d - Cx \\ 0 &\leq b - Ax \\ 0 &\leq \lambda \end{aligned}$$

Como tal, isto é seguramente mais fácil que resolver o $MLCP(\tilde{F}(0), Q)$ que contem a condição de complementariedade adicional entre λ e $b - Ax$.

Agora, consideramos a conversão da condição KKT da $AVI(K, q, M)$ num LCP. Esta conversão tem tanto razões historicas como implicações algorítmicas. Especificamente, seja $F(x) = q + Mx$ onde $q \in \mathbb{R}^n$, $M \in \mathbb{R}^{n \times n}$, da proposição 3.2 temos que:

$$\begin{aligned} 0 &= q + Mx + C^T\mu + A^T\lambda \\ 0 &= d - Cx \\ 0 &\leq \lambda \perp b - Ax \geq 0. \end{aligned} \tag{3.13}$$

Suponha que a matriz

$$\begin{bmatrix} M & C^T \\ -C & 0 \end{bmatrix} \tag{3.14}$$

é não singular, podemos então usar as duas primeiras igualdades em (3.13) para resolver as variables (x, μ) em termos de λ , como segue:

$$\begin{aligned} Mx + C^T\mu &= -q - A^T\lambda \\ -Cx + 0\mu &= -d \end{aligned}$$

Então,

$$\begin{bmatrix} M & C^T \\ -C & 0 \end{bmatrix} \begin{bmatrix} x \\ \mu \end{bmatrix} = \begin{bmatrix} -q - A^T\lambda \\ -d \end{bmatrix}.$$

Logo, se $\begin{bmatrix} U & V \\ W & Z \end{bmatrix} = \begin{bmatrix} M & C^T \\ -C & 0 \end{bmatrix}^{-1}$ temos que:

$$\begin{bmatrix} x \\ \mu \end{bmatrix} = \begin{bmatrix} U & V \\ W & Z \end{bmatrix} \begin{bmatrix} -q - A^T \lambda \\ -d \end{bmatrix} = \begin{bmatrix} U(-q - A^T \lambda) - Vd \\ W(-q - A^T \lambda) - Zd \end{bmatrix}$$

Assim, $x = -U(q + A^T \lambda) - Vd$ e $\mu = -W(q + A^T \lambda) - Zd$, eliminando assim (x, μ) do sistema (3.13). Além disso:

$$b - Ax = b + A[U(q + A^T \lambda) + Vd] = b + AUq + AVd + AUA^T \lambda$$

ou equivalentemente,

$$b - Ax = b + [AU \quad AV] \begin{bmatrix} q \\ d \end{bmatrix} + [AU \quad AV] \begin{bmatrix} A^T \\ 0 \end{bmatrix} \lambda$$

Mais ainda,

$$b - Ax = b + [A \quad 0] \begin{bmatrix} U & V \\ W & Z \end{bmatrix} \begin{bmatrix} q \\ d \end{bmatrix} + [A \quad 0] \begin{bmatrix} U & V \\ W & Z \end{bmatrix} \begin{bmatrix} A^T \\ 0 \end{bmatrix} \lambda$$

Isto é:

$$b - Ax = b + [A \quad 0] \begin{bmatrix} M & C^T \\ -C & 0 \end{bmatrix}^{-1} \begin{bmatrix} q \\ d \end{bmatrix} + [A \quad 0] \begin{bmatrix} M & C^T \\ -C & 0 \end{bmatrix}^{-1} \begin{bmatrix} A^T \\ 0 \end{bmatrix} \lambda$$

Logo, definindo:

$$\bar{q} = b + [A \quad 0] \begin{bmatrix} M & C^T \\ -C & 0 \end{bmatrix}^{-1} \begin{bmatrix} q \\ d \end{bmatrix} \quad \text{e} \quad \bar{M} = [A \quad 0] \begin{bmatrix} M & C^T \\ -C & 0 \end{bmatrix}^{-1} \begin{bmatrix} A^T \\ 0 \end{bmatrix}$$

Obtemos o $LCP(\bar{q}, \bar{M})$ só na variable λ

$$0 \leq \lambda \perp \bar{F}(\lambda) \geq 0,$$

onde $\bar{F}(\lambda) = \bar{q} + \bar{M}\lambda$.

Tradicionalmente, o objetivo de transformar um MLCP num LCP é facilitar a implementação dos métodos pivote LCP para resolver o problema acima. Além de exigir a condição de não singularidade, como vimos anteriormente, esta transformação envolve uma carga computacional quando se trata de resolver problemas grandes, com os avanços na metodologia da AVI, a necessidade de transformar um MLCP num LCP para fins computacionais tem diminuíu; porém, tal conversão é as vezes benéfica porque um algoritmo pivote pode processar MLCPs, que são solúveis de outra forma por algoritmos iterativos.

3.3 Integrabilidade e o sistema KKT

As inequações variacionais e os problemas de complementaridade surgem de uma variedade de fontes interessantes. A maioria dessas fontes são problemas de otimização restrita diferenciável. No entanto, nem todos os VIs ou CPs são naturalmente derivados de problemas de otimização. Parte do objetivo desta seção é mostrar que uma VI é uma extensão não trivial de um programa não-linear padrão (PNL). Sabe-se que as condições de Karush-Kuhn-Tucker (KKT) desempenharam um papel fundamental em todos os aspectos da programação não-linear. Acontece que essas condições podem ser facilmente estendidas para uma VI. Para maiores informações veja [3].

3.3.1 Problemas de otimização restrita

Considere o problema de otimização restrito:

$$\begin{aligned} &\text{Minimizar } \theta(x) && (3.15) \\ &\text{sujeito a: } x \in K \end{aligned}$$

Onde a função objetivo θ é continuamente diferenciável. Suponhamos que o conjunto K é convexo, então pela proposição 2.11, qualquer minimizador local x de (3.15) deve satisfazer:

$$\langle \nabla\theta(x), y - x \rangle \geq 0 \quad \forall y \in K.$$

Esta última desigualdade pode ser vista facilmente como a $VI(K, \nabla\theta)$ que é chamada o problema de ponto estacionário associado ao problema de otimização (3.15). Uma solução da $VI(K, \nabla\theta)$ é chamada ponto estacionário do problema (3.15). Além disso, sabemos que se θ é uma função convexa, então cada ponto estacionário de (3.15) é um mínimo global deste problema de otimização (proposição 2.12). Portanto, para um problema de programação convexa, isto é, com θ uma função convexa e K um conjunto convexo, a $VI(K, \nabla\theta)$ é equivalente ao problema de otimização (3.15).

Definição 3.5. Seja $U \subseteq \mathbb{R}^n$ um conjunto aberto, uma função vectorial $F : U \rightarrow \mathbb{R}^n$ se chama mapa de gradiente, se existe uma função escalar θ tal que $F(x) = \nabla\theta(x)$ para todo x no domínio de F .

Definição 3.6. Seja $U \subseteq \mathbb{R}^n$ um conjunto aberto e conexo. A função $F : U \rightarrow \mathbb{R}^n$ é integrável, se para quaisquer dois vetores $x, y \in U$, a integral de linha de F de x a y é independente de qualquer caminho liso por partes em U que conecte x e y .

A discussão acima levanta uma questão. Dada uma $VI(K, F)$ com uma função vectorial arbitrária F e um conjunto convexo K , esta VI é sempre o problema de ponto estacionário de algum problema de otimização (3.15) com K como o conjunto viável?. A resposta a esta questão é negativa. Uma questão relacionada pode ser feita em termos

da função F : é dizer, quando uma função vetorial F é um mapa de gradiente?. Podemos obter uma resposta completa a esta última pergunta. O seguinte teorema é clássico, e essencialmente, mostra que os três conceitos mapa de gradiente, função integrável e simetria, são todos equivalentes. Para maiores informações ver [3].

Teorema 3.1. *Seja $F : U \rightarrow \mathbb{R}^n$ uma função de classe C^1 no conjunto aberto e convexo $U \subseteq \mathbb{R}^n$. As seguintes afirmações são equivalentes:*

- a) *Existe uma função real θ tal que: $F(x) = \nabla\theta(x)$ para todo $x \in U$.*
- b) *A matriz jacobiana $JF(x)$ é simétrica para todo $x \in U$.*
- c) *F é integrável em U .*

Se qualquer uma dessas afirmações é verdadeira, então a função escalar desejada θ é dada por:

$$\theta(x) = \int_0^1 \langle F(x_0 + t(x - x_0)), x - x_0 \rangle dt \text{ para um vetor arbitrário } x_0 \in U.$$

Uma classe importante de funções integráveis são as funções separáveis; estas são funções contínuas $F(x)$ tais que cada função componente $F_i(x)$ depende da única variável x_i , isto é, $F(x) = (F_i(x_i) : i = 1, \dots, n)$. Se $F(x)$ é separável e diferenciável, então a matriz jacobiana $JF(x)$ é uma matriz diagonal para todo x no conjunto onde F é diferenciável.

O resultado do teorema 3.1 é que, se F é um mapa de gradiente definida num conjunto convexo K . Então, a $VI(K, F)$ é o problema ponto estacionário do problema de otimização (3.15), onde $F(x) = \nabla\theta(x)$. Para ilustrar esta conclusão, considere o caso afim, onde F é dada por $F(x) = q + Mx$, $q \in \mathbb{R}^n$, $M \in \mathbb{R}^{n \times n}$. Como $JF(x) = M$, para todo $x \in \mathbb{R}^n$, segue-se que F é um mapa de gradiente em \mathbb{R}^n se e somente se M é uma matriz simétrica, neste caso, a $VI(K, q, M)$ é o problema de ponto estacionário do seguinte problema de otimização:

$$\begin{array}{ll} \text{Minimizar} & q^T x + \frac{1}{2} x^T M x \\ \text{sujeito a} & x \in K, \end{array}$$

o qual é um problema quadrático (QP) quando K é um conjunto poliedral.

Voltando ao problema de otimização (3.15); consideremos o caso em que K não é convexo. Neste caso não convexo, pode-se ainda obter uma condição necessária de primeira ordem para um minimizador local de (3.15), mas isso não corresponde mais a uma VI. No entanto, sob algumas suposições adequadas, pode-se obter uma condição primal-dual necessária para a otimização que corresponde a um MiCP.

Como vimos na definição 2.7, o cone tangente de Bouligand $\mathcal{B}(x, K)$, consiste de todos os vetores $d \in \mathbb{R}^n$, chamados vetores tangentes a K em x , para o qual existe uma sequência

de vetores $(x_k) \subseteq K$ e uma sequência de escalares positivos $(t_k) \subseteq \mathbb{R}_+$ tais que:

$$\lim_{k \rightarrow \infty} x_k = x, \quad \lim_{k \rightarrow \infty} t_k = 0, \quad \text{e} \quad \lim_{k \rightarrow \infty} \frac{x_k - x}{t_k} = d.$$

Para mais detalhes ver [8], [21]. Pelo teorema 2.1, se x é um minimizador local de (3.15),

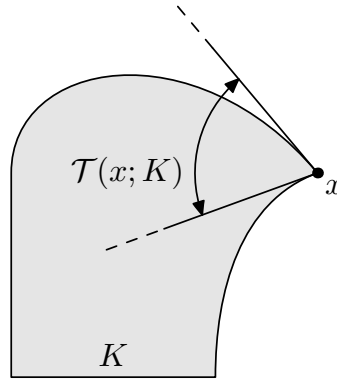


Figura 12 – O Cone Tangente.

então,

$$\langle \nabla \theta(x), y - x \rangle \geq 0 \quad \forall y \in x + \mathcal{B}(x, K). \quad (3.16)$$

Em geral, um vetor $x \in K$ que satisfaz a desigualdade (3.16), se denomina ponto estacionário do problema (não convexo) (3.15).

A desigualdade (3.16) é um exemplo de uma desigualdade quase variacional, abreviada como QVI. Formalmente uma QVI se define como segue.

Definição 3.7. Seja $K : \mathbb{R}^n \rightarrow \mathcal{P}(\mathbb{R}^n)$ uma aplicação de \mathbb{R}^n em subconjuntos de \mathbb{R}^n , isto é, para todo $x \in \mathbb{R}^n$; $K(x)$ é um subconjunto (possivelmente vazio) de \mathbb{R}^n ; seja $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ uma aplicação de \mathbb{R}^n em si mesmo. A QVI definida pelo par (K, F) é encontrar um vetor $x \in K(x)$ tal que:

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

Note que se $F(x)$ é identicamente nula, a QVI é reduzida ao problema de encontrar um vetor x que satisfaz $x \in K(x)$, esse vetor é chamado ponto fixo da aplicação ponto a conjunto K . Assim, além de ser uma extensão da VI, a QVI também inclui o problema clássico de encontrar um ponto fixo de uma aplicação ponto a conjunto como um caso especial.

O cone tangente (de Bouligand) é outro objeto que desempenha um papel importante na análise convexa e programação não linear. No caso de um conjunto convexo K , o teorema 2.7 diz que para qualquer $x \in K$ cumpre-se:

$$(\mathcal{B}(x, K))^* = \mathcal{N}(x, K). \quad (3.17)$$

Sem uma descrição adicional do conjunto não convexo K , em geral não é fácil de tratar com o cone tangente $\mathcal{B}(x, K)$. No caso importante de um problema não linear NLP padrão, isto é, o problema (3.15) onde K é representado por uma quantidade finita de igualdades e desigualdades diferenciáveis:

$$K = \{x \in \mathbb{R}^n : h(x) = 0, g(x) \leq 0\} \quad (3.18)$$

Com $h : \mathbb{R}^n \rightarrow \mathbb{R}^l$ e $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$ funções continuamente diferenciáveis, há muitas condições bem conhecidas sobre as funções h e g sob as quais o cone tangente $\mathcal{B}(x, K)$ torna-se um cone poliédrico. Mais importante ainda, a condição de estacionariedade (3.16) tem então uma descrição equivalente primal - dual. Essas condições sobre as restrições são conhecidas como condições de qualificação (CQs); uma das mais gerais é a CQ de Abadie, (ver o teorema 2.15) que disse que o cone tangente $\mathcal{B}(x, K)$ é igual ao cone de linearização $\mathcal{L}(x, K)$ de K em $x \in K$, definido em (2.26), onde K é definido como em (3.18).

$$\mathcal{L}(x, K) = \{d \in \mathbb{R}^n : \langle d, \nabla h_j(x) \rangle = 0 \ \forall j = 1, \dots, l; \ \langle d, \nabla g_i(x) \rangle \leq 0 \ \forall i \in I(x)\}$$

Em geral, pela proposição 2.9, para qualquer conjunto K definido como em (3.18) o cone tangente $\mathcal{B}(x, K)$ esta sempre contido em $\mathcal{L}(x, K)$ para qualquer $x \in K$. Assim, a CQ de Abadie é equivalente à premissa de que cada vetor no cone de linearização $\mathcal{L}(x, K)$ é um vetor tangente. Sob algumas condições de regularidade o cone tangente $\mathcal{B}(x, K)$ é igual ao cone de linearização das restrições $\mathcal{L}(x, K)$ (ver teorema 2.15).

A igualdade entre $\mathcal{B}(x, K)$ e $\mathcal{L}(x, K)$ implica que:

$$\begin{aligned} \mathcal{B}(x, K)^* &= \mathcal{L}(x, K)^* \\ &= \left\{ v \in \mathbb{R}^n : \exists(\mu, \lambda_{I(x)}), \lambda_{I(x)} \geq 0 \text{ tal que } 0 = v + \sum_{j=1}^l \mu_j \nabla h_j(x) + \sum_{i \in I(x)} \lambda_i \nabla g_i(x) \right\} \end{aligned}$$

Se K também é convexo, de 3.17 temos uma representação poliédrica de $\mathcal{N}(x, K)$. Assumindo que a CQ de Abadie é verificada para o vetor $x \in K$, podemos derivar a formulação primal - dual equivalente de (3.16). Sob esta CQ, a condição (3.16) diz que $-\nabla\theta(x)$ pertence ao cone dual de $\mathcal{L}(x, K)$ (ver (2.31)). Assim, pelo teorema 2.17 temos que (3.16) é válido se e somente se existem multiplicadores de restrições (variáveis dual) $\mu \in \mathbb{R}^l$ e $\lambda \in \mathbb{R}^m$ tais que o sistema de Karush - Kuhn - Tucker (KKT) abaixo é válido:

$$\begin{aligned} 0 &= \nabla\theta(x) + \sum_{j=1}^l \mu_j \nabla h_j(x) + \sum_{i=1}^m \lambda_i \nabla g_i(x) \\ 0 &= h(x) \\ 0 &\leq \lambda \perp g(x) \leq 0 \end{aligned}$$

Reconhecemos facilmente o último sistema como um MiCP nas variáveis (x, μ, λ) . Este sistema KKT é a descrição primal - dual de um ponto estacionário de (3.15), primal - dual no sentido de que o sistema contém tanto a variável primal x e as variáveis duais (ou multiplicadores) (μ, λ) .

3.3.2 O Sistema Karush - Kuhn - Tucker

Nesta subseção, estendemos as condições KKT para a VI. Igual à VI linearmente restrita, podemos facilmente estabelecer a seguinte proposição, parte de cuja prova consiste simplesmente em substituir $\nabla\theta(x)$ na anterior derivação, por uma função vetorial arbitrária $F(x)$.

Proposição 3.3. *Seja o conjunto $K = \{x \in \mathbb{R}^n : h(x) = 0, g(x) \leq 0\}$, onde as funções h_j e g_i são continuamente diferenciáveis, seja $F : K \rightarrow \mathbb{R}^n$ uma aplicação de K em \mathbb{R}^n . As duas afirmações seguintes são válidas:*

a) *Seja $x \in SOL(K, F)$. Se a CQ de Abadie é válida em x , então existem vetores $\mu \in \mathbb{R}^l$ e $\lambda \in \mathbb{R}^m$ tais que:*

$$\begin{aligned} 0 &= F(x) + \sum_{j=1}^l \mu_j \nabla h_j(x) + \sum_{i=1}^m \lambda_i \nabla g_i(x) \\ 0 &= h(x) \\ 0 &\leq \lambda \perp g(x) \leq 0 \end{aligned} \tag{3.19}$$

b) *Por outro lado, se cada função h_j é afim, cada função g_i é convexa e se (x, μ, λ) satisfaz (3.19), então x resolve a VI(K, F).*

Demonstração. Para provar (a) notemos que $x \in SOL(K, F)$, se e somente se

$$\langle F(x), y \rangle \geq \langle F(x), x \rangle \quad \forall y \in K$$

o qual é equivalente a dizer que x resolve o seguinte problema de otimização não linear na variável y :

$$\begin{aligned} &\text{Minimizar} && \langle F(x), y \rangle \\ &\text{Sujeito a} && y \in K \end{aligned} \tag{3.20}$$

Dado que (3.19) é precisamente o sistema KKT de este problema de otimização não linear pelo teorema (2.17) ((2.35)-(2.39)), cumpre-se a).

Reciprocamente, se as funções h_j e g_i são funções afins e convexas respectivamente, pelo corolário 2.12 o conjunto K é convexo, e como a função objetivo do problema (3.20) é convexa, o problema (3.20) é um problema de programação convexo na variável y . Logo, pelos teoremas 2.18 e 2.9 tem-se que cada ponto KKT de este problema é uma solução global (que não necessariamente é única), logo x deve resolver a VI(K, F), isto é $x \in SOL(K, F)$. \square

O MiCP exibido na proposição acima é chamado o sistema KKT da VI(K, F) com um conjunto "finitamente representável" K , isto é, com K definido por (3.18). O sistema

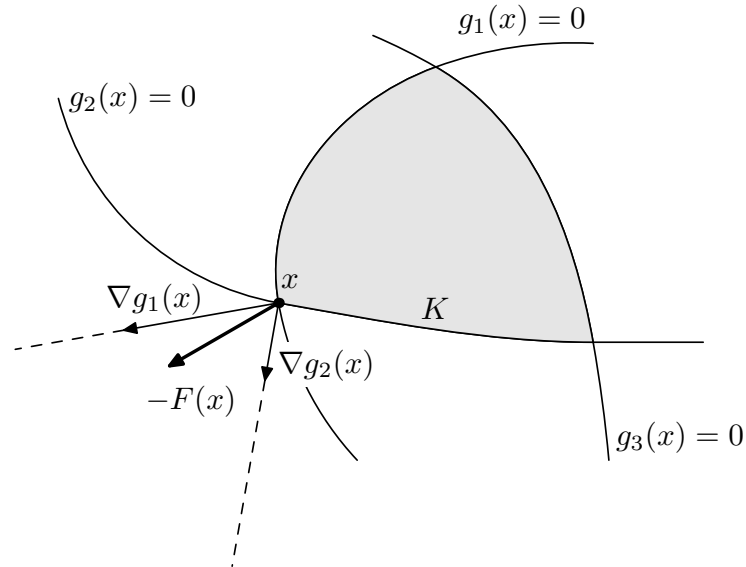


Figura 13 – Condições KKT de uma VI.

KKT representa uma conversão da VI em um MiCP.

É útil introduzir a função vetorial:

$$\mathbf{L}(x, \mu, \lambda) = F(x) + \sum_{j=1}^l \mu_j \nabla h_j(x) + \sum_{i=1}^m \lambda_i \nabla g_i(x) \quad \forall (x, \mu, \lambda) \in \mathbb{R}^{n+l+m}$$

A qual chamamos a função Lagrangiana da $VI(K; F)$. Esta terminologia é emprestada do caso em que F é um mapa de gradiente, digamos $F(x) = \nabla \theta(x)$. Neste último caso, temos:

$$\mathbf{L}(x, \mu, \lambda) = \nabla_x l(x, \mu, \lambda)$$

onde:

$$l(x, \mu, \lambda) = \theta(x) + \langle \mu, h(x) \rangle + \langle \lambda, g(x) \rangle \quad (3.21)$$

é a função escalar lagrangiana familiar do PNL diferenciável:

$$\begin{aligned} &\text{Minimizar} && \theta(x) \\ &\text{sujeito a} && h(x) = 0 \\ &&& g(x) \leq 0 \end{aligned}$$

A função que define o MiCP (3.19) é uma generalização da função dada por (3.12) para um K não poliédrico; em termos da função lagrangiana \mathbf{L} para a $VI(K, F)$, a função de definição para o sistema KKT (3.19) é:

$$\tilde{F}(x, \mu, \lambda) = \begin{pmatrix} \mathbf{L}(x, \mu, \lambda) \\ h(x) \\ -g(x) \end{pmatrix} \quad \forall (x, \mu, \lambda) \in \mathbb{R}^{n+l+m} \quad (3.22)$$

Como um MiCP, o sistema (3.19) está bem definido para todas as funções diferenciáveis h e g . Este sistema fornece um conjunto de condições necessárias que devem ser satisfeitas

por todas as soluções de a VI que obedecem à CQ Abadie; Inversamente, essas condições dão soluções à VI se as funções h_j forem afins e g_i forem convexas (pelo corolário 2.12 o conjunto K é convexo neste caso).

3.4 Problemas de equilíbrio de Nash

Um conceito fundamental de equilíbrio na teoria não-cooperativa de jogos foi introduzido por J. Nash, que recebeu o Prêmio Nobel de Ciências Econômicas em 1994 por essa contribuição. Acontece que o cálculo de um "equilíbrio de Nash" pode ser realizado resolvendo uma desigualdade variacional. No que se segue, nós explicamos a abordagem de VI para o cálculo de equilíbrios de Nash. Para maiores informações veja a referência [3]. Conforme a definição 2.21, em um jogo não-cooperativo geral, existem N jogadores cada um dos quais tem uma certa função de custo e conjunto de estratégias que pode depender das ações dos outros jogadores. Para a discussão inicial, assumimos que o conjunto de estratégias do jogador i é $K_i \subseteq \mathbb{R}^{n_i}$ e é independente das estratégias dos outros jogadores. A função de custo do jogador i que é $\theta_i(x)$ depende de todas as estratégias dos jogadores, que são descritas pelo vetor x que consiste nos subvetores $x^i \in \mathbb{R}^{n_i}$ para $i = 1, \dots, N$. Segundo a definição 2.24 de equilíbrio de Nash, o problema do jogador i é determinar, para cada tupla fixa mas arbitrária $\tilde{x}^i = (x^j : j \neq i)$ das estratégias de outros jogadores, uma estratégia ótima x^i que resolva o problema de minimização de custos na variável y^i :

$$\begin{aligned} &\text{Minimizar } \theta_i(y^i, \tilde{x}^i) \\ &\text{sujeito a: } y^i \in K_i \end{aligned}$$

Denotamos o conjunto de soluções deste problema de otimização por $S_i(\tilde{x}^i)$; observe a dependência desse conjunto na tupla \tilde{x}^i . Um ligeiro abuso de notação ocorre na função objetivo $\theta_i(y^i, \tilde{x}^i)$; Entende-se que isto significa a função θ_i avaliada no vetor cujo j -ésimo subvetor é x^j para $j \neq i$ e cujo i -ésimo subvetor é y^i . Conforme a definição 2.24, temos a seguinte.

Definição 3.8. Um equilíbrio de Nash é uma tupla de estratégias $x = (x^i : i = 1, \dots, N)$ com a propriedade de que para cada $i = 1, \dots, N$, $x^i \in S_i(\tilde{x}^i)$.

Em palavras, um equilíbrio de Nash é uma tupla de estratégias, uma para cada jogador, de modo que nenhum jogador pode diminuir o custo desviando unilateralmente sua ação de sua estratégia designada. O resultado seguinte (ver [3]) fornece um conjunto de condições suficientes sob as quais um equilíbrio de Nash pode ser obtido pela resolução de uma VI.

Proposição 3.4. *Seja cada $K_i \subseteq \mathbb{R}^{n_i}$ um conjunto convexo e fechado. Suponha que para cada tupla fixa \tilde{x}^i , a função $\theta_i(y^i, \tilde{x}^i)$ seja convexa e continuamente diferenciável em*

y^i . Então, uma tupla $x = (x^i : i = 1, \dots, N)$ é um equilíbrio de Nash se e somente se $x \in SOL(K, F)$, onde:

$$K = \prod_{i=1}^N K_i \quad e \quad F(x) = (\nabla_{x^i} \theta_i(x) : i = 1, \dots, N).$$

Demonstração. Sabemos que $x = (x^i : i = 1, \dots, N)$ é um equilíbrio de Nash se e somente se para cada $i = 1, \dots, N$, x^i resolve o problema de minimização convexo na variável y^i :

$$\begin{aligned} & \text{Minimizar } \theta_i(y^i, \tilde{x}^i) \\ & \text{sujeito a: } y^i \in K_i \end{aligned}$$

Logo, como cada $K_i \subseteq \mathbb{R}^{n_i}$ é um conjunto fechado e convexo, pelas proposições 2.11 e 2.12 temos que x é um equilíbrio de Nash se e somente se para cada $i = 1, \dots, N$

$$\langle \nabla_{x^i} \theta_i(x), y^i - x^i \rangle \geq 0 \quad \forall y^i \in K_i. \quad (3.23)$$

Por outro lado, como $F(x) = (\nabla_{x^i} \theta_i(x) : i = 1, \dots, N)$, então cada função componente de F , vem dada por: $F_i(x) = \nabla_{x^i} \theta_i(x)$. Assim, usando (3.23) temos que:

$$\langle F(x), y - x \rangle = \sum_{i=1}^N \langle F_i(x), y^i - x^i \rangle = \sum_{i=1}^N \langle \nabla_{x^i} \theta_i(x), y^i - x^i \rangle \geq 0.$$

Portanto, x resolve a $VI(K, F)$, isto é $x \in SOL(K, F)$.

Reciprocamente, se $x = (x^i : i = 1, \dots, N)$ resolve a $VI(K, F)$ então

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

em particular, para cada $i = 1, \dots, N$, seja y a tupla cujo j -ésimo subvetor é igual a x^j para $j \neq i$ e o i -ésimo subvetor é igual a y^i , onde y^i é um elemento arbitrário do conjunto K_i . Assim, $y = (y^i, \tilde{x}^i) \in K$ e além disso:

$$\begin{aligned} 0 \leq \langle F(x), y - x \rangle &= \sum_{j=1}^N \langle F_j(x), y^j - x^j \rangle = \sum_{\substack{j=1 \\ j \neq i}}^N \langle F_j(x), x^j - x^j \rangle + \langle F_i(x), y^i - x^i \rangle \\ &= \langle \nabla_{x^i} \theta_i(x), y^i - x^i \rangle \end{aligned}$$

Logo, para cada $i = 1, \dots, N$ temos que:

$$0 \leq \langle \nabla_{x^i} \theta_i(x), y^i - x^i \rangle \quad \forall y^i \in K_i$$

Esta última desigualdade então se torna (3.23). Portanto, x é um equilíbrio de Nash. \square

Vale ressaltar que o conjunto de definição K da VI na Proposição 3.4 é o produto cartesiano de conjuntos de dimensões menores.

4 UM ESTUDO DE UM CASO DE EQUILÍBRIO DE NASH

Neste capítulo, apresentamos um estudo de um caso de equilíbrio de Nash do modelo econômico de Nash - Cournot (ver [20]). Em primeiro lugar, conforme a definição 2.24, definimos para este caso um ponto de equilíbrio de Nash, assim como também damos uma orientação geral sobre como tal equilíbrio será determinado. A seguir, apresenta-se uma serie de resultados a certos programas de otimização relacionados aos equilíbrios de Nash. Vamos a seguir as ideias propostas no artigo [20] para caracterizar um ponto de equilíbrio em termos de uma familia de programas de otimização convexos, onde as soluções para esta familia de programas convexos são mostrados como equilíbrios de Nash no sentido formal dos jogos de N pessoas como vimos na seção 2.3 do capítulo 1. Finalmente, apresentamos um exemplo numérico que ilustra as ideias principais; o qual é resolvido como problema de complementaridade (MiCP e NCP), usando-se a metodologia proposta em [20].

4.1 Definição do problema

O problema a ser resolvido é definido como segue: considere uma estrutura de mercado oligopolística na qual N empresas fornecem um produto homogêneo de uma forma não cooperativa. Seja $p(Q)$, $Q \geq 0$, a curva de demanda inversa. Ou seja, $p(Q)$ é o preço a que os consumidores irão exigir (e comprar) uma quantidade Q . Em conformidade com o comportamento econômico geralmente aceito, assumimos que $p(Q)$ é estritamente decrescente e que a curva de ingressos $I(Q) = Qp(Q)$ é uma função côncava de Q para $Q \geq 0$. Além disso, assumimos que $p(Q)$ é continuamente diferenciável.

Agora, suponha que a curva de oferta individual para a i -ésima empresa seja representada pela função de custo total $f_i(q_i)$, $q_i \geq 0$ para $i = 1 \dots N$. Isto é, seja $f_i(q_i)$ o custo total da i -ésima empresa de fornecer q_i unidades do produto. Assumimos no seguinte desenvolvimento que $f_i(q_i)$ é uma função convexa (geralmente não decrescente) continuamente diferenciável de q_i para $i = 1 \dots N$.

O objetivo principal deste trabalho é determinar um conjunto de níveis de produção $q_1^*, q_2^* \dots q_N^*$ para as N empresas (um ponto de equilíbrio de Nash) conforme à definição 2.24. Assim, temos a seguinte definição.

Definição 4.1. Um conjunto de níveis de saída não negativos $q_1^*, q_2^* \dots q_N^*$ para as empresas $1, 2 \dots N$, respectivamente, constituem uma solução de *equilíbrio de Nash* para o mercado, se e somente se q_i^* maximiza o lucro da i -ésima empresa, dado que as outras empresas produzem quantidades q_j^* , $j \neq i$, para cada $i = 1, \dots, N$.

Ou seja, para $q_1^*, q_2^* \dots q_N^*$ ser um equilíbrio de Nash, ou simplesmente uma solução

de equilíbrio, q_i^* , deve ser uma solução ótima para o problema

$$\max_{q_i \geq 0} \{q_i p(q_i + Q_i^*) - f_i(q_i)\} \quad \text{onde} \quad Q_i^* = \sum_{\substack{j=1 \\ j \neq i}}^N q_j^* \quad \text{para cada} \quad i = 1, \dots, N. \quad (4.1)$$

No problema (4.1) acima, a função objetivo é uma expressão para o lucro da i -ésima empresa, ou seja, ingressos $q_i p(q_i + Q_i^*)$ menos custo $f_i(q_i)$ a um nível de produção q_i , dado um volume de saída de q_j^* para cada uma das outras empresas $j \in \{1 \dots N\}$, $j \neq i$.

A estratégia geral desenvolvida neste trabalho para estabelecer a existência de uma solução de equilíbrio, bem como determinar essa solução envolve a construção de uma determinada família de programas convexos. A forma geral desta família é:

$$\begin{aligned} v(Q) &= \max h(q_1, q_2, \dots, q_N; Q) \\ \sum_{i=1}^N q_i &= Q \quad (\lambda) \\ q_i &\geq 0 \end{aligned} \quad (4.2)$$

Onde Q é um parâmetro escalar e $h(q_1, q_2, \dots, q_N; Q)$ é côncava para cada Q fixo. A função h será construída de modo que se a escolha de $Q = Q^*$ resulta em uma solução ótima $(q_1^*, q_2^*, \dots, q_N^*)$ com a variável de Kuhn-Tucker $\lambda^* = 0$, então $(q_1^*, q_2^*, \dots, q_N^*)$ será um equilíbrio de Nash. A natureza da variável de Kuhn-Tucker λ , em função de Q , será de importância central tanto para estabelecer a existência de soluções de equilíbrio bem como para determinar procedimentos computacionais para encontrar um equilíbrio. Por exemplo, será mostrado que λ , sob suposições apropriadas, é uma função de Q , não-crescente e contínua. Portanto, se $Q_1 < Q_2$ e soluções ótimas para $v(Q_1)$ e $v(Q_2)$ resultam em $\lambda(Q_1) \geq 0$ e $\lambda(Q_2) \leq 0$, então a continuidade de λ estabelece a existência de um Q^* para o qual $\lambda(Q^*) = 0$ e a monotonicidade de λ fornece um quadro bem definido para determinar iterativamente Q^* .

4.2 Desenvolvimento matemático

O Teorema 4.1 abaixo reitera um conceito econômico fundamental associado a uma solução de equilíbrio. Para provar este resultado, precisamos do seguinte lema.

Lema 4.1. *Para $Q \geq 0$, seja $I(Q) = Qp(Q)$ uma função côncava de Q e suponha que $p(Q)$ é uma função não-crescente ou convexa de Q . Então, para cada $K > 0$, $\varphi(Q) = Qp(Q+K)$ é uma função côncava de Q para $Q \geq 0$. Além disso, se $p(Q)$ é estritamente decrescente ou é estritamente convexa, então $\varphi(Q) = Qp(Q+K)$ é uma função estritamente côncava na linha real não-negativa.*

Demonstração. Estabeleceremos este resultado sem assumir que $p(Q)$ é duas vezes diferenciável. No entanto, se $p(Q)$ é duas vezes diferenciável, um argumento mais simples está

disponível e é apresentado posteriormente. Seja $\lambda \in (0, 1)$, $Q_1 \geq 0$, $Q_2 \geq 0$ e defina

$$X = \varphi(\lambda Q_1 + (1 - \lambda)Q_2) = [\lambda Q_1 + (1 - \lambda)Q_2] p(\lambda Q_1 + (1 - \lambda)Q_2 + K)$$

$$Y = \lambda \varphi(Q_1) + (1 - \lambda)\varphi(Q_2) = \lambda Q_1 p(Q_1 + K) + (1 - \lambda)Q_2 p(Q_2 + K)$$

Vamos mostrar primeiro que se $p(Q)$ é não-crescente ou convexa, então $X \geq Y$. Suponha ao contrário que $X < Y$. Note-se que

$$\begin{aligned} X &= [\lambda Q_1 + (1 - \lambda)Q_2] p(\lambda Q_1 + (1 - \lambda)Q_2 + K) \\ &= [\lambda Q_1 + (1 - \lambda)Q_2 + K] p(\lambda Q_1 + (1 - \lambda)Q_2 + K) - K p(\lambda Q_1 + (1 - \lambda)Q_2 + K) \\ &= [\lambda Q_1 + (1 - \lambda)Q_2 + \lambda K + (1 - \lambda)K] p(\lambda Q_1 + (1 - \lambda)Q_2 + \lambda K + (1 - \lambda)K) \\ &\quad - K p(\lambda Q_1 + (1 - \lambda)Q_2 + K) \\ &= [\lambda(Q_1 + K) + (1 - \lambda)(Q_2 + K)] p(\lambda(Q_1 + K) + (1 - \lambda)(Q_2 + K)) \\ &\quad - K p(\lambda Q_1 + (1 - \lambda)Q_2 + K) \end{aligned}$$

Então pelo anterior, temos:

$$\begin{aligned} 0 > X - Y &= [\lambda(Q_1 + K) + (1 - \lambda)(Q_2 + K)] p(\lambda(Q_1 + K) + (1 - \lambda)(Q_2 + K)) \\ &\quad - K p(\lambda Q_1 + (1 - \lambda)Q_2 + K) - Y \\ &= I(\lambda(Q_1 + K) + (1 - \lambda)(Q_2 + K)) - K p(\lambda Q_1 + (1 - \lambda)Q_2 + K) - Y \end{aligned}$$

Agora, como a curva de ingresos da indústria $I(Q) = Qp(Q)$ é uma função côncava, para todo $Q \geq 0$ e $(Q_1 + K) \geq 0$, $(Q_2 + K) \geq 0$ então

$$I(\lambda(Q_1 + K) + (1 - \lambda)(Q_2 + K)) \geq \lambda I(Q_1 + K) + (1 - \lambda)I(Q_2 + K)$$

Logo

$$\begin{aligned} 0 > X - Y &\geq \lambda(Q_1 + K)p(Q_1 + K) + (1 - \lambda)(Q_2 + K)p(Q_2 + K) \\ &\quad - K p(\lambda Q_1 + (1 - \lambda)Q_2 + K) - Y \\ &= \lambda Q_1 p(Q_1 + K) + \lambda K p(Q_1 + K) + (1 - \lambda)Q_2 p(Q_2 + K) + (1 - \lambda)K p(Q_2 + K) \\ &\quad - K p(\lambda Q_1 + (1 - \lambda)Q_2 + K) - \lambda Q_1 p(Q_1 + K) - (1 - \lambda)Q_2 p(Q_2 + K) \\ &= \lambda K p(Q_1 + K) + (1 - \lambda)K p(Q_2 + K) - K p(\lambda Q_1 + (1 - \lambda)Q_2 + K) \\ &= K [\lambda p(Q_1 + K) + (1 - \lambda)p(Q_2 + K) - p(\lambda Q_1 + (1 - \lambda)Q_2 + K)] \\ &= K [Z] \end{aligned}$$

onde $Z = \lambda p(Q_1 + K) + (1 - \lambda)p(Q_2 + K) - p(\lambda Q_1 + (1 - \lambda)Q_2 + K)$.

Assim, temos mostrado que $X < Y$ implica que $Z < 0$.

Suponhamos que $p(Q)$ é uma função convexa, para todo $Q \geq 0$, então

$$\begin{aligned} \lambda p(Q_1 + K) + (1 - \lambda)p(Q_2 + K) &\geq p(\lambda(Q_1 + K) + (1 - \lambda)(Q_2 + K)) \\ &= p(\lambda Q_1 + (1 - \lambda)Q_2 + K) \end{aligned}$$

isto implica que $Z \geq 0$, esta é uma contradição. Suponha que $p(Q)$ é simplesmente não-crescente. Então acima, desde que $Z < 0$,

$$\lambda p(Q_1 + K) + (1 - \lambda)p(Q_2 + K) < p(\lambda Q_1 + (1 - \lambda)Q_2 + K)$$

Assim, temos:

$$\begin{aligned} X - Y &= [\lambda Q_1 + (1 - \lambda)Q_2] p(\lambda Q_1 + (1 - \lambda)Q_2 + K) - Y \\ &> [\lambda Q_1 + (1 - \lambda)Q_2] [\lambda p(Q_1 + K) + (1 - \lambda)p(Q_2 + K)] - Y \\ &= \lambda(1 - \lambda)Q_2 p(Q_1 + K) + \lambda(1 - \lambda)Q_1 p(Q_2 + K) + \lambda^2 Q_1 p(Q_1 + K) \\ &\quad + (1 - \lambda)^2 Q_2 p(Q_2 + K) - \lambda Q_1 p(Q_1 + K) - (1 - \lambda)Q_2 p(Q_2 + K) \\ &= \lambda(1 - \lambda) [Q_2 p(Q_1 + K) + Q_1 p(Q_2 + K)] \\ &\quad + \lambda Q_1 p(Q_1 + K) [\lambda - 1] + (1 - \lambda)Q_2 p(Q_2 + K) [(1 - \lambda) - 1] \\ &= \lambda(1 - \lambda) [Q_2 p(Q_1 + K) + Q_1 p(Q_2 + K)] - \lambda(1 - \lambda)Q_1 p(Q_1 + K) \\ &\quad - \lambda(1 - \lambda)Q_2 p(Q_2 + K) \\ &= \lambda(1 - \lambda) [Q_2 p(Q_1 + K) + Q_1 p(Q_2 + K) - Q_1 p(Q_1 + K) - Q_2 p(Q_2 + K)] \\ &= \lambda(1 - \lambda)(Q_1 - Q_2) [p(Q_2 + K) - p(Q_1 + K)] \end{aligned}$$

Mas, como $p(Q)$ é não-crescente, isto significa que se $Q_1 < Q_2$, então $Q_1 + K < Q_2 + K$ logo $p(Q_1 + K) \geq p(Q_2 + K)$, assim temos que:

$$(Q_1 - Q_2) [p(Q_2 + K) - p(Q_1 + K)] \geq 0$$

Similarmente, se $Q_2 < Q_1$, então $Q_2 + K < Q_1 + K$ logo $p(Q_2 + K) \geq p(Q_1 + K)$, assim temos que:

$$(Q_1 - Q_2) [p(Q_2 + K) - p(Q_1 + K)] \geq 0$$

logo, podemos concluir que:

$$0 > X - Y > \lambda(1 - \lambda)(Q_1 - Q_2) [p(Q_2 + K) - p(Q_1 + K)] \geq 0$$

que é uma contradição, portanto $\varphi(Q) = Qp(Q + K)$ é uma função côncava, para todo $Q \geq 0$. É fácil verificar que se $p(Q)$ é estritamente decrescente ou é estritamente convexa, então a suposição de que $X \leq Y$ leva a contradições semelhantes. Isso completa a prova. (Caso diferenciável) Se $p(Q)$ for duas vezes diferenciável, então a prova acima simplifica como se segue. Suponha por contradição que $\varphi''(Q) > 0$ para algum Q . Então,

$$\begin{aligned} 0 < \varphi''(Q) &= 2p'(Q + K) + Qp''(Q + K) \\ &= [2p'(Q + K) + (Q + K)p''(Q + K)] - Kp''(Q + K) \end{aligned} \quad (4.3)$$

como a função $I(Q) = Qp(Q)$ é côncava, então $I''(Q) \leq 0$ para todo $Q \geq 0$, logo para $Q + K > 0$ tem-se

$$I''(Q + K) = 2p'(Q + K) + (Q + K)p''(Q + K) \leq 0$$

então de (4.3) temos que:

$$0 < \varphi''(Q) \leq -Kp''(Q + K)$$

e desde que $K > 0$ então $p''(Q + K) < 0$. Se p é convexa, então $p''(Q + K) \geq 0$, claramente esta é uma contradição. Por outro lado, suponha que $p(Q)$ é simplesmente não-crescente; então $p'(Q + K) \leq 0$. Mas, acima temos que $p''(Q + K) < 0$, o que implica que $Qp''(Q + K) \leq 0$ para $Q \geq 0$. Então

$$0 < \varphi''(Q) = 2p'(Q + K) + Qp''(Q + K) \leq 0$$

Uma contradição. Se $p(Q)$ for estritamente decrescente ou for estritamente convexa, então contradições semelhantes são obtidas supondo que $\varphi''(Q) > 0$. Isto completa a prova. \square

É interessante notar que alguns autores consideraram necessário, para estabelecer matematicamente a existência de uma solução de equilíbrio, assumir que $\varphi(Q) = Qp(Q+K)$, com $Q \geq 0$, é côncava para cada $K > 0$. Contudo, o Lema 4.1 atenua essa dificuldade mostrando que essa suposição está implícita nas condições usuais de não-crescente de $p(Q)$ e concavidade de $I(Q) = Qp(Q)$. Consideremos agora o seguinte resultado fundamental.

Teorema 4.1. *Considere um oligopólio com N empresas como descrito acima e suponha que as curvas de custo individuais f_i são convexas e continuamente diferenciáveis para todo $i = 1, \dots, N$ e que a função de demanda inversa p é estritamente decrescente e continuamente diferenciável. Além disso, suponha que a curva de ingressos da indústria $I(Q) = Qp(Q)$ é côncava, $Q \geq 0$. Então, $(q_1^*, q_2^*, \dots, q_N^*)$ é uma solução de equilíbrio (Nash) se e somente se*

$$p(Q^*) + q_i^* p'(Q^*) \leq f'_i(q_i^*) \quad \text{para cada } i = 1, \dots, N \quad (4.4)$$

$$q_i^* [p(Q^*) + q_i^* p'(Q^*) - f'_i(q_i^*)] = 0 \quad \text{para cada } i = 1, \dots, N \quad (4.5)$$

onde

$$Q^* = \sum_{i=1}^N q_i^*, \quad q_i^* \geq 0 \quad (4.6)$$

Além disso, se $(q_1^*, \dots, q_i^*, \dots, q_N^*)$ e $(q_1^*, \dots, \hat{q}_i, \dots, q_N^*)$ são soluções de equilíbrio, então $q_i^* = \hat{q}_i$.

Demonstração. Pela definição 4.1 temos que q_1^*, \dots, q_N^* é uma solução de equilíbrio de Nash, se e somente se, para cada $i = 1, \dots, N$, q_i^* deve ser uma solução ótima para o problema (4.1)

$$\max_{q_i \geq 0} \{q_i p(q_i + Q_i^*) - f_i(q_i)\} \quad \text{onde } Q_i^* = \sum_{\substack{j=1 \\ j \neq i}}^N q_j^*.$$

Agora, pela hipótese, a curva de ingressos da indústria $I(Q) = Qp(Q)$ é côncava para todo $Q \geq 0$ e a função de demanda inversa p é estritamente decrescente. Então, pelo lema

4.1 a função $q_i p(q_i + Q_i^*)$ é estritamente côncava para todo $q_i \geq 0$. Além disso, como as curvas de custo individuais f_i são convexas; então $-f_i$ são funções côncavas para todo $i = 1, \dots, N$. Portanto, a função $q_i p(q_i + Q_i^*) - f_i(q_i)$ é estritamente côncava para todo $i = 1, \dots, N$. Assim, a função $\theta_i(q_1, \dots, q_i, \dots, q_N) = f_i(q_i) - q_i p(q_i + Q_i^*)$ é estritamente convexa. Então, resolver o problema (4.1) é equivalente a resolver, para cada $i = 1, \dots, N$, o seguinte problema de programação convexa.

$$\min_{q_i \geq 0} \{f_i(q_i) - q_i p(q_i + Q_i^*)\} \quad \text{onde} \quad Q_i^* = \sum_{\substack{j=1 \\ j \neq i}}^N q_j^* \quad (4.7)$$

Note-se que: para cada $i = 1, \dots, N$, a restrição de desigualdade é $g(q) \leq 0$, onde $g_i(q) = -q_i$. Além disso,

$$\frac{\partial \theta_i}{\partial q_i}(q) = f_i'(q_i) - p(q_i + Q_i^*) - q_i p'(q_i + Q_i^*) \quad \text{e} \quad \frac{\partial g}{\partial q_i}(q) = -1.$$

Logo, aplicando as condições KKT (2.35)-(2.39) (que pelos teoremas 2.17 e 2.18 são condições necessárias e suficientes) ao problema de programação convexa 4.7 temos que, para cada $i = 1, \dots, N$, q_i^* é um minimizador de (4.7), se e somente se, existe $\mu_i \in \mathbb{R}^+$ tal que:

$$\begin{aligned} f_i'(q_i^*) - p(q_i^* + Q_i^*) - q_i^* p'(q_i^* + Q_i^*) + \mu_i(-1) &= 0 \\ \mu_i(-q_i^*) &= 0 \\ \mu_i \geq 0, \quad q_i^* &\geq 0 \end{aligned}$$

isto é equivalente a:

$$\begin{aligned} 0 \leq \mu_i = f_i'(q_i^*) - p(Q^*) - q_i^* p'(Q^*) \\ -\mu_i q_i^* = 0 \\ q_i^* \geq 0 \quad \text{onde} \quad Q^* = \sum_{j=1}^N q_j^* \end{aligned}$$

e isto é equivalente a dizer

$$\begin{aligned} p(Q^*) + q_i^* p'(Q^*) \leq f_i'(q_i^*) \quad \text{para cada} \quad i = 1, \dots, N \\ q_i^* [p(Q^*) + q_i^* p'(Q^*) - f_i'(q_i^*)] = 0 \quad \text{para cada} \quad i = 1, \dots, N \end{aligned}$$

$$\text{onde:} \quad Q^* = \sum_{j=1}^N q_j^*, \quad q_i^* \geq 0$$

Além disso, note que, pelo teorema 2.9, se o problema 4.7 tem um minimizador, então este é único. \square

Note que, considerando as aplicações $G : \mathbb{R} \times \mathbb{R}_+^N \longrightarrow \mathbb{R}$ e $H : \mathbb{R} \times \mathbb{R}_+^N \longrightarrow \mathbb{R}^N$ definidas como

$$G(Q, q) = \sum_{i=1}^N q_i - Q,$$

$$H_i(Q, q) = f'_i(q_i) - p(Q) - q_i p'(Q) \quad \text{para cada } i = 1, \dots, N, \quad (4.8)$$

tem-se que, conforme a definição 3.4 de MiCP e o teorema 4.1, o ponto $(Q^*, q_1^*, \dots, q_N^*)$ satisfaz o seguinte MiCP:

$$G(Q, q) = 0, \quad (4.9)$$

$$0 \leq q \perp H(Q, q) \geq 0, \quad (4.10)$$

se e somente se, $(q_1^*, q_2^*, \dots, q_N^*)$ é uma solução de equilíbrio de Nash no sentido da definição 4.1. Assim, podemos calcular soluções de equilíbrio de Nash resolvendo o MiCP (4.9)-(4.10).

Por outro lado, usando a equação (4.9), podemos isolar Q em função das q_i , a saber $Q = \sum_{i=1}^N q_i$, e substituir na equação (4.8) para obter o seguinte:

$$H_i(q) = f'_i(q_i) - p\left(\sum_{i=1}^N q_i\right) - q_i p'\left(\sum_{i=1}^N q_i\right) \quad \text{para cada } i = 1, \dots, N.$$

Agora, usando a equação (4.10) temos o seguinte problema de complementaridade não linear (NCP):

$$0 \leq q \perp H(q) \geq 0 \quad (4.11)$$

Logo, pelo teorema 4.1 tem-se que o ponto $(q_1^*, q_2^*, \dots, q_N^*)$ é uma solução de equilíbrio de Nash, se e somente se, este ponto é solução do NCP (4.11), isto é, também podemos obter soluções de equilíbrio, resolvendo o NCP (4.11).

Observe que as relações (4.4), (4.5) e (4.6) afirmam essencialmente que, em um equilíbrio de Nash, o ingresso marginal deve ser igual ao custo marginal para cada empresa que está produzindo uma quantidade positiva e que o ingresso marginal não deve exceder o custo marginal para as empresas que preferem não estar no negócio. Esta afirmação intuitiva tem uma base econômica bem aceita.

Matematicamente, o Teorema 4.1 indica que simplesmente precisamos encontrar uma solução para o $MiCP(G, H)$ definido por o sistema (4.9)-(4.10) ou para o $NCP(H)$ dado por (4.11) para obter um conjunto de saídas em condições de equilíbrio e garante a disponibilidade de uma solução para este sistema no caso de que um equilíbrio existe.

No entanto, devido que na década de 1980 os algoritmos para resolver MiCPs e/ou NCPs eram incipientes, o autor [20] teve que usar um método alternativo. Para realizar a tarefa

de encontrar esse conjunto de níveis de saída, considere o seguinte problema de equilíbrio $EP(Q)$.

$$\begin{aligned} \text{Maximizar} \quad & P(Q) \sum_{i=1}^N q_i + \frac{1}{2} p'(Q) \sum_{i=1}^N q_i^2 - \sum_{i=1}^N f_i(q_i) \\ \text{Sujeito a} \quad & \sum_{i=1}^N q_i = Q \end{aligned} \quad (4.12)$$

$$q_i \geq 0, \quad i = 1, \dots, N \quad (4.13)$$

Observe que, para um determinado $Q \geq 0$, o problema $EP(Q)$ busca o máximo de uma função objetivo

$$F(q) = P(Q) \sum_{i=1}^N q_i + \frac{1}{2} p'(Q) \sum_{i=1}^N q_i^2 - \sum_{i=1}^N f_i(q_i) \quad \text{onde } q = (q_1, q_2, \dots, q_N)$$

Logo, para cada $i = 1, \dots, N$, tem-se:

$$\frac{\partial F}{\partial q_i}(q) = p(Q) + p'(Q)q_i - f'_i(q_i)$$

Assim, como $p'(Q) < 0$ e para cada $i = 1, \dots, N$ as funções f_i são convexas, obtemos:

$$\frac{\partial^2 F}{\partial q_i^2}(q) = p'(Q) - f''_i(q_i) < 0$$

Assim, a função F é estritamente côncava e está definida sobre um conjunto poliedral $K = \{q \in \mathbb{R}^N : h(q) = 0, g(q) \leq 0\}$ limitado, não vazio, e pelo corolário 2.10 é convexo, onde as funções $h : \mathbb{R}^N \rightarrow \mathbb{R}$ e $g : \mathbb{R}^N \rightarrow \mathbb{R}^N$ são definidas como segue:

$$h(q) = \sum_{i=1}^N q_i - Q \quad \text{e} \quad g_i(q) = -q_i, \quad i = 1, \dots, N.$$

Portanto, pelo teorema 2.9, não pode haver mais de uma solução ótima do problema $EP(Q)$, e pelos teoremas 2.17 e 2.18, as condições de KKT (2.35) - (2.39) para este problema são necessárias e suficientes para a otimização. Estas condições são apresentadas abaixo com λ e μ_i , $i = 1, \dots, N$, sendo os multiplicadores duais associados às restrições (4.12) e (4.13), respectivamente.

$$\begin{aligned} -\nabla F(q) + \lambda \nabla h(q) + \sum_{i=1}^N \mu_i \nabla g_i(x) &= 0 \\ h(q) &= 0 \\ \mu_i g_i(q) &= 0, \quad i = 1, \dots, N \\ \mu_i &\geq 0, \quad i = 1, \dots, N \\ g_i(q) &\leq 0, \quad i = 1, \dots, N \end{aligned}$$

e desde que $\frac{\partial h}{\partial q_i}(q) = 1$ e $\frac{\partial g_i}{\partial q_i}(q) = -1$ para cada $i = 1, \dots, N$, temos que:

$$\begin{aligned} -p(Q) - p'(Q)q_i + f'_i(q_i) + \lambda + \mu_i(-1) &= 0, \quad i = 1, \dots, N \\ \sum_{i=1}^N q_i - Q &= 0 \\ \mu_i(-q_i) &= 0, \quad i = 1, \dots, N \\ \mu_i &\geq 0, \quad i = 1, \dots, N \\ -q_i &\leq 0, \quad i = 1, \dots, N \end{aligned}$$

isto é

$$p(Q) + p'(Q)q_i - f'_i(q_i) - \lambda + \mu_i = 0, \quad i = 1, \dots, N \quad (4.14)$$

$$\sum_{i=1}^N q_i = Q \quad (4.15)$$

$$\mu_i q_i = 0, \quad i = 1, \dots, N \quad (4.16)$$

$$\mu_i \geq 0, \quad q_i \geq 0 \quad i = 1, \dots, N \quad (4.17)$$

O papel desempenhado pelo problema $EP(Q)$ é evidente a partir da seguinte afirmação.

Teorema 4.2. *Seja $\hat{Q} \geq 0$ tal que a única solução ótima $\hat{q}_1, \dots, \hat{q}_N$ para o problema $EP(\hat{Q})$ resulta em um valor de $\lambda = 0$ no sistema (4.14) - (4.17). Então, $\hat{q}_1, \dots, \hat{q}_N$ é uma solução de equilíbrio. Reciprocamente, seja q_1^*, \dots, q_N^* uma solução de equilíbrio. Então q_1^*, \dots, q_N^* resolve $EP(Q^*)$ onde $Q^* = \sum_{i=1}^N q_i^*$. Além disso, se $Q^* > 0$, então λ é necessariamente zero.*

Demonstração. Dado $\hat{Q} \geq 0$, suponha que $\hat{q}_1, \dots, \hat{q}_N$ é a única solução do problema $EP(\hat{Q})$ e que $\lambda = 0$ no sistema (4.14) - (4.17). Então, de (4.14) e (4.17) temos que:

$$0 \leq \mu_i = f'_i(\hat{q}_i) - p(\hat{Q}) - \hat{q}_i p'(\hat{Q}) \quad \text{para cada } i = 1, \dots, N$$

portanto

$$p(\hat{Q}) + \hat{q}_i p'(\hat{Q}) \leq f'_i(\hat{q}_i) \quad \text{para cada } i = 1, \dots, N$$

de (4.16) tem-se:

$$\hat{q}_i [p(\hat{Q}) + \hat{q}_i p'(\hat{Q}) - f'_i(\hat{q}_i)] = 0 \quad \text{para cada } i = 1, \dots, N$$

finalmente, de (4.15) e (4.17) temos:

$$\sum_{i=1}^N \hat{q}_i = \hat{Q}, \quad \hat{q}_i \geq 0 \quad i = 1, \dots, N$$

Assim, $\hat{q}_1, \dots, \hat{q}_N$ cumpre (4.4), (4.5) e (4.6); logo, do teorema 4.1 concluímos que $\hat{q}_1, \dots, \hat{q}_N$ é um equilíbrio de Nash. Por outro lado, suponha que q_1^*, \dots, q_N^* é uma solução de equilíbrio

de Nash. Então q_1^*, \dots, q_N^* cumpre as relações (4.4), (4.5) e (4.6) do teorema 4.1. Logo, para cada $i = 1, \dots, N$ considerando

$$q_i = q_i^*, \quad Q^* = \sum_{i=1}^N q_i^*, \quad \lambda = 0, \quad \mu_i = f'_i(q_i^*) - p(Q^*) - q_i^* p'(Q^*)$$

q_1^*, \dots, q_N^* cumpre as relações (4.14), (4.15), (4.16) e (4.17) e portanto é uma solução ótima do problema $EP(Q^*)$. Além disso, se $Q^* > 0$, então $q_i > 0$ para algum $i \in \{1, \dots, N\}$; logo, desde que q_1^*, \dots, q_N^* resolve o $EP(Q^*)$, por (4.16) temos que $\mu_i = 0$ e substituindo em (4.14) concluímos que $\lambda = 0$, isto completa a demonstração do teorema. \square

Corolário 4.3. *Se (q_1^*, \dots, q_N^*) e $(\hat{q}_1, \dots, \hat{q}_N)$ são soluções de equilíbrio distintas, então*

$$\sum_{i=1}^N q_i^* \neq \sum_{i=1}^N \hat{q}_i$$

Demonstração. Segue diretamente da unicidade de uma solução ótima para o problema $EP(Q)$ para um $Q \geq 0$ dado. \square

Até agora, reduzimos a tarefa de encontrar uma solução de equilíbrio à de encontrar uma quantidade Q^* para a qual o problema $EP(Q^*)$ possui um multiplicador dual ótimo associado à restrição (4.12) com valor zero. Para então, o teorema 4.2 afirma que a solução ótima (q_1^*, \dots, q_N^*) para o problema $EP(Q^*)$ é uma solução de equilíbrio. Além disso, o Teorema 4.2 mostra que, se existir uma solução de equilíbrio, então, um deve ser capaz de realizar esta tarefa. Para este fim, demonstramos a seguir uma propriedade crucial da variável λ quando é vista como uma função de Q para $Q \geq 0$. Ou seja, mostra-se que, $\lambda(Q)$ é uma função contínua não crescente de Q . A importância desta propriedade de um ponto de vista algorítmico é claramente evidente. Para estabelecer esse resultado, precisaremos usar as seguintes lemas.

Lema 4.2. *Considere o problema $EP(Q)$ e seja $q_i(Q)$, $\lambda(Q)$, $\mu_i(Q)$, $i = 1, \dots, N$, as soluções ótimas de KKT fornecidas por (4.14) - (4.17). Então, q_i , λ e μ_i são funções contínuas de Q em $(0, +\infty)$ para $i = 1, \dots, N$.*

Demonstração. Denote a função objetivo de $EP(Q)$ por $F(q, Q)$ e seja $v(Q)$ o valor da função objetivo ótimo. Isto é,

$$\begin{aligned} F(q, Q) &= P(Q) \sum_{i=1}^N q_i + \frac{1}{2} p'(Q) \sum_{i=1}^N q_i^2 - \sum_{i=1}^N f_i(q_i) \\ v(Q) &= F(q(Q), Q) \end{aligned} \quad (4.18)$$

onde $q(Q) = (q_1(Q), \dots, q_N(Q))$. Pode ser verificado em [20] que $v(Q)$ é contínua em $(0, +\infty)$. Assim, dada uma seqüência não negativa $Q^n \rightarrow Q$, $Q > 0$, segue-se que

$$v(Q^n) = F(q(Q^n), Q^n) \rightarrow v(Q) = F(q(Q), Q) \quad (4.19)$$

Além disso, como (Q^n) é convergente, ela também é limitada. Então, considerando a norma da soma temos que existe $c > 0$ tal que:

$$\|q(Q^n)\| = \sum_{i=1}^N q_i(Q^n) = Q^n \leq c$$

Logo, a seqüência $(q(Q^n))$ é limitada e portanto possui uma subseqüência $(q(Q^{n_k}))$ convergente. Isto é, existe $\bar{q} = (\bar{q}_1, \dots, \bar{q}_N)$ tal que:

$$\lim_{k \rightarrow +\infty} q(Q^{n_k}) = \bar{q} \quad (4.20)$$

Suponha que $\bar{q} \neq q(Q)$. Então, desde que a função $\psi(x) = \sum_{i=1}^N x_i$ é contínua, para todo $x = (x_1, \dots, x_N) \in \mathbb{R}^N$ e como

$$\psi(q(Q^{n_k})) = \sum_{i=1}^N q_i(Q^{n_k}) = Q^{n_k}, \quad \lim_{k \rightarrow +\infty} Q^{n_k} = Q, \quad q(Q^{n_k}) \geq 0 \quad \text{para cada } k \in \mathbb{N} \quad (4.21)$$

De (4.21) e (4.20) segue-se que:

$$\lim_{k \rightarrow +\infty} \psi(q(Q^{n_k})) = \lim_{k \rightarrow +\infty} Q^{n_k} = Q = \psi(\bar{q})$$

Isto é:

$$\sum_{i=1}^N \bar{q}_i = Q \quad \text{e} \quad \bar{q} \geq 0 \quad (4.22)$$

Isso significa que \bar{q} é viável para $EP(Q)$. Mas como F é uma função contínua, temos que:

$$\lim_{k \rightarrow +\infty} F(q(Q^{n_k}), Q^{n_k}) = F(\bar{q}, Q)$$

Logo, a unicidade do limite e (4.19) implicam que:

$$F(\bar{q}, Q) = F(q(Q), Q) \quad (4.23)$$

Então, (4.23) implica que o problema $EP(Q)$ tem dois maximizadores diferentes; o que contradiz a unicidade de $q(Q)$ para o $Q > 0$ dado. Assim, q é contínua em $(0, +\infty)$.

A continuidade das variáveis de KKT λ e $\mu = (\mu_1, \dots, \mu_N)$ agora segue-se diretamente de (4.14) - (4.17). De fato; seja $Q > 0$, de (4.15) tem-se que $q_i > 0$ para algum $i \in \{1, \dots, N\}$, então graças a (4.16) temos que $\mu_i = 0$, agora substituindo isto em (4.14) obtemos:

$$\lambda(Q) = p(Q) + p'(Q)q_i(Q) - f'_i(q_i(Q)) \quad (4.24)$$

E como p , p' , f'_i e q_i são todas funções contínuas, Segue-se que λ é contínua. A equação (4.14) então mostra que μ é contínua. \square

A seguinte proposição é importante para provar o Lema 4.3.

Proposição 4.1. *Seja $X \subseteq \mathbb{R}$ e $f : X \rightarrow \mathbb{R}$ derivável à direita no ponto $a \in X \cap X'_+$. Se $f'_+(a) > 0$, então existe $\delta > 0$ tal que $x \in X$, $a < x < a + \delta$ implica que $f(a) < f(x)$.*

Demonstração. A demonstração pode ser encontrada na referência [14]. \square

Lema 4.3. *Considere o problema $EP(Q)$ e seja q, λ, μ representam a solução para o sistema de (4.14) - (4.17). Definir mais o conjunto*

$$I = \{i \in \{1, \dots, N\} : D_Q^+(q_i) > 0\} \quad (4.25)$$

Onde $D_x^+(\cdot)$ denota a derivada pela direita do argumento em relação a x . Então

$$\lambda = p(Q) + p'(Q)q_i - f'_i(q_i) \quad \text{para } i \in I \quad (4.26)$$

Além disso, (4.26) é válido para cada $i \in I$, e Q que varia de Q para $Q + \epsilon$ para algum $\epsilon > 0$.

Demonstração. O resultado é claro para qualquer $i \in I$ satisfazendo $q_i > 0$, pois neste caso, graças à equação (4.16) temos que $\mu_i = 0$; assim, cumpre-se (4.26). Portanto, dado qualquer $Q > 0$ e $i \in I$, suponha que $q_i(Q) = 0$, vamos mostrar que $\mu_i(Q) = 0$. Por contradição, suponha que $\mu_i(Q) > 0$ para algum $i \in I$. Logo, como $i \in I$, temos que $D_Q^+(q_i) > 0$. Então, pela proposição 4.1 existe $\delta > 0$ tal que:

$$0 = q_i(Q) < q_i(\bar{Q}) \quad \text{para todo } \bar{Q} > 0 \text{ e } Q < \bar{Q} < Q + \delta$$

Agora considere $0 < \epsilon < \delta$ e $\bar{Q} = Q + \epsilon$, então $0 < q_i(Q + \epsilon)$. Logo, da equação (4.16) tem-se que $\mu_i(Q + \epsilon) = 0$. Mas, por outro lado, sabemos que μ_i é contínua em Q e pela hipótese auxiliar $\mu_i(Q) > 0$, então $\mu_i(Q + \epsilon) > 0$ para ϵ suficientemente pequeno, isto é uma contradição. Logo, temos que $\mu_i(Q) = 0$ para todo $i \in I$, então de (4.14) obtemos:

$$\lambda = p(Q) + p'(Q)q_i - f'_i(q_i) \quad \text{para todo } i \in I$$

Assim, verifica-se (4.26). Além disso, da demonstração, observe que (4.26) também é válida no intervalo $[Q, Q + \epsilon]$ para algum $\epsilon > 0$. \square

Lema 4.4. *Considere o problema $EP(Q)$ e seja $\lambda(Q)$ o multiplicador dual associado à restrição (4.12), suponha que f_i são funções convexas para $i = 1, \dots, N$, a função de demanda inversa p é estritamente decrescente e $Qp(Q)$ é uma função côncava. Então, λ é uma função não crescente de Q , $Q \geq 0$.*

Demonstração. Seja $Q \geq 0$, vamos mostrar que $D_Q^+(\lambda) \leq 0$. Definindo I como em (4.25), o Lema 4.3 afirma que (4.26) também é válida no intervalo $[Q, Q + \epsilon]$, para algum $\epsilon > 0$; então, substituindo (4.26) em (4.14) tem-se:

$$p(Q) + p'(Q)q_i - f'_i(q_i) - (p(Q) + p'(Q)q_i - f'_i(q_i)) + \mu_i = 0, \quad \text{para todo } i \in I$$

Então $\mu_i = 0$ em $[Q, Q + \epsilon]$ para todo $i \in I$. Logo, segue-se que $D_Q^+(\mu_i) = 0$ para cada $i \in I$. Consequentemente, a partir de (4.26) e regra da cadeia, obtemos para cada $i \in I$,

$$D_Q^+(\lambda) = p'(Q) + q_i D_Q^+(p'(Q)) + p'(Q) D_Q^+(q_i) - D_{q_i}^+(f'_i(q_i)) D_Q^+(q_i) \quad (4.27)$$

Adicionando (4.27) sobre $i \in I$, obtemos:

$$\sum_{i \in I} D_Q^+(\lambda) = \sum_{i \in I} p'(Q) + \sum_{i \in I} q_i D_Q^+(p'(Q)) + \sum_{i \in I} p'(Q) D_Q^+(q_i) - \sum_{i \in I} D_{q_i}^+(f'_i(q_i)) D_Q^+(q_i)$$

Agora, fazendo $\sum_{i \in I} 1 = |I| = r \geq 1$, tem-se o seguinte:

$$r D_Q^+(\lambda) = r p'(Q) + D_Q^+(p'(Q)) \sum_{i \in I} q_i + p'(Q) \sum_{i \in I} D_Q^+(q_i) - \sum_{i \in I} D_{q_i}^+(f'_i(q_i)) D_Q^+(q_i)$$

Assim, temos:

$$r D_Q^+(\lambda) = p'(Q) \left[r + \sum_{i \in I} D_Q^+(q_i) \right] + D_Q^+(p'(Q)) \sum_{i \in I} q_i - \sum_{i \in I} D_{q_i}^+(f'_i(q_i)) D_Q^+(q_i)$$

Como f_i é convexa e continuamente diferenciável, então $D_{q_i}^+(f'_i(q_i)) \geq 0$ para cada $i = 1, \dots, N$. Além disso, como $D_Q^+(q_i) > 0$ para todo $i \in I$, nós obtemos:

$$r D_Q^+(\lambda) \leq p'(Q) \left[r + \sum_{i \in I} D_Q^+(q_i) \right] + D_Q^+(p'(Q)) \sum_{i \in I} q_i \quad (4.28)$$

Se $\sum_{i \in I} q_i = 0$ em qualquer Q , então

$$r D_Q^+(\lambda) \leq p'(Q) \left[r + \sum_{i \in I} D_Q^+(q_i) \right].$$

Logo, como $p'(Q) < 0$ e $\left[r + \sum_{i \in I} D_Q^+(q_i) \right] > 0$, então $D_Q^+(\lambda) < 0$. Portanto, suponha que

$\sum_{i \in I} q_i > 0$. Assim, multiplicando (4.28) por $\frac{Q}{\sum_{i \in I} q_i} \geq 0$ obtemos o seguinte:

$$\frac{Q}{\sum_{i \in I} q_i} r D_Q^+(\lambda) \leq \frac{Q}{\sum_{i \in I} q_i} \left[r + \sum_{i \in I} D_Q^+(q_i) \right] p'(Q) + Q D_Q^+(p'(Q)) \quad (4.29)$$

Por outro lado, como $\sum_{i \in I} q_i \leq \sum_{i=1}^N q_i = Q$, temos que $1 \leq \frac{Q}{\sum_{i \in I} q_i}$ e como $r \geq 1$, isto implica que

$$1 \leq \frac{Q}{\sum_{i \in I} q_i} r \quad (4.30)$$

Além disso, $\sum_{i=1}^N q_i(Q) = Q$, então $\sum_{i=1}^N D_Q^+(q_i) = 1$. Assim,

$$\sum_{i=1}^N D_Q^+(q_i) = \sum_{i \in I} D_Q^+(q_i) + \sum_{i \notin I} D_Q^+(q_i) = 1.$$

Mas $D_Q^+(q_i) \leq 0 \quad \forall i \notin I$; logo,

$$\sum_{i \in I} D_Q^+(q_i) = 1 - \sum_{i \notin I} D_Q^+(q_i) \geq 1,$$

portanto, temos que

$$1 \leq \frac{Q}{\sum_{i \in I} q_i} \sum_{i \in I} D_Q^+(q_i) \quad (4.31)$$

Agora, adicionando (4.30) e (4.31), obtemos:

$$2 \leq \frac{Q}{\sum_{i \in I} q_i} \left[r + \sum_{i \in I} D_Q^+(q_i) \right]$$

Em seguida, multiplicando por $p'(Q) < 0$ em ambos lados da desigualdade, obtemos:

$$\frac{Q}{\sum_{i \in I} q_i} \left[r + \sum_{i \in I} D_Q^+(q_i) \right] p'(Q) \leq 2p'(Q) \quad (4.32)$$

Logo, a partir de (4.29) e (4.32) segue-se

$$\frac{Q}{\sum_{i \in I} q_i} r D_Q^+(\lambda) \leq 2p'(Q) + Q D_Q^+(p'(Q)) = D_Q^+ \left\{ \frac{d}{dQ} [Qp(Q)] \right\} \quad (4.33)$$

Mas, pela hipótese geral, a função $Qp(Q)$ é concava, então $D_Q^+ \left\{ \frac{d}{dQ} [Qp(Q)] \right\} \leq 0$. Logo, de (4.30) e (4.33) segue-se que $D_Q^+(\lambda) \leq 0$ para $Q > 0$. Portanto a função λ é não-crescente. \square

A importância do Lema 4.4 é uma caracterização da família de problemas $\{EP(Q)\}$. Em particular, se para algum Q_1 um resolve $EP(Q_1)$ e encontra $\lambda(Q_1) > 0$, então segue que o aumento de Q para $Q_2 > 0$ resultará em $\lambda(Q_2) \leq \lambda(Q_1)$. Uma vez que se procura um Q^* tal que $\lambda(Q^*) = 0$, a monotonicidade de λ garante que sempre se pode mover na direção correta.

O próximo lema aborda a questão do conjunto de soluções de equilíbrio e sua unicidade em condições de regularidade bastante fracas.

Lema 4.5. *Sob as hipóteses do lema 4.4, definir o conjunto*

$$S = \left\{ Q : \sum_{i=1}^N q_i = Q, \text{ para uma solução de equilíbrio } (q_1, \dots, q_N) \right\}.$$

Então S é um conjunto convexo. Além disso, se uma das seguintes três condições adicionais se mantiver para o mercado oligopolista, e se $S \neq \emptyset$, então S é um conjunto unitario e portanto, existe uma solução de equilíbrio única:

i) f_i são estritamente convexas para $i = 1, \dots, N$.

ii) $Qp(Q)$ é estritamente côncava, $Q > O$.

iii) $p(Q)$ é convexa, $Q > 0$.

Demonstração. Uma vez que λ é uma função não-crescente e contínua de Q (Lemma 4.4 e Lemma 4.2) e como $Q \in S$ se e somente se $EP(Q)$ tiver $\lambda(Q) = 0$ na otimidade (Teorema 4.2), então $S = \lambda^{-1}(0)$. Logo, S é um conjunto convexo. Além disso, a partir da prova do Lema 4.4, segue-se que, se cumpre-se uma das condições estabelecidas, então λ é uma função estritamente decrescente de Q e portanto injetiva. Por isso, se existe uma solução de equilíbrio, então o Teorema 4.2 implica que deve corresponder a $\lambda = 0$, que deve ser para um valor único de Q . Isto é, se $Q_1, Q_2 \in S$, então $\lambda(Q_1) = 0 = \lambda(Q_2)$ implica que $Q_1 = Q_2$. Assim S é um conjunto unitário. Mas o $EP(Q)$ possui uma solução ótima única para cada Q . Assim, existe uma solução de equilíbrio única. Isso completa a prova. \square

O estágio agora está definido para descrever o procedimento algorítmico para determinar um equilíbrio de Nash $(q_1^*, q_2^*, \dots, q_N^*)$ para a estrutura do mercado oligopolista. Primeiro, derivamos limites superiores e inferiores aplicáveis a todos os elementos em S , o conjunto de todas as soluções de equilíbrio. Esses limites superiores e inferiores na produção total correspondem a um equilíbrio assumindo, em primeiro lugar, um mercado em que o custo marginal é igual ao preço e, segundo, uma estrutura de mercado monopolista. Os dois programas (convexos) a serem resolvidos são:

$$\begin{aligned} \text{Maximizar} \quad & \int_0^Q p(x)dx - \sum_{i=1}^N f_i(q_i) \\ & \sum_{i=1}^N q_i - Q = 0 \quad (\lambda) \\ & q_i, Q \geq 0 \end{aligned} \tag{4.34}$$

$$\begin{aligned} \text{Maximizar} \quad & Qp(Q) - \sum_{i=1}^N f_i(q_i) \\ & \sum_{i=1}^N q_i - Q = 0 \quad (\lambda) \\ & q_i, Q \geq 0 \end{aligned} \tag{4.35}$$

Seja Q_u e Q_l (parte de) uma solução ótima para (4.34) e (4.35), respectivamente. Em seguida, resolver $EP(Q_l)$ e $EP(Q_u)$ para obter $\lambda(Q_l)$ e $\lambda(Q_u)$. O próximo lema demonstra a utilidade e a razão para obter Q_l e Q_u .

Lema 4.6. Para $q_i, Q \geq O$, suponha que $Qp(Q)$ seja côncava, $p(Q)$ é estritamente decrescente, e cada $f_i(q_i)$ é convexa. Então, se Q_l, Q_u existem, obtém-se:

$$a) \lambda(Q_l) \geq 0 \geq \lambda(Q_u)$$

b) $Q_l \leq Q \leq Q_u$ para cada $Q \in S$

Demonstração. a) Primeiro, seja (q_u, Q_u) uma solução óptima para (4.34). Então,

$$p(Q_u) - f'_i(q_{ui}) \leq 0 \quad i = 1, \dots, N$$

de modo que:

$$p(Q_u) + q_{ui}p'(Q_u) - f'_i(q_{ui}) \leq 0 \quad i = 1, \dots, N \quad (4.36)$$

Agora, considere $EP(Q_u)$ e seja \hat{q} a solução única. Suponha que $\lambda(Q_u) > 0$. Definir $I = \{i : \hat{q}_i > 0\}$. Assim, temos:

$$\lambda(Q_u) = p(Q_u) + \hat{q}_i p'(Q_u) - f'_i(\hat{q}_i) > 0 \quad i \in I \quad (4.37)$$

De (4.36) e (4.37) segue-se que $\hat{q}_i < q_{ui} \quad i \in I$.

Mas

$$Q_u = \sum_{i=1}^N \hat{q}_i = \sum_{i \in I} \hat{q}_i$$

e desde que:

$$\sum_{i \in I} \hat{q}_i < \sum_{i \in I} q_{ui} \leq Q_u$$

Uma contradição é obtida. Assim, $\lambda(Q_u) \leq 0$.

Em seguida, seja (q_l, Q_l) uma solução óptima para (4.35). Então

$$p(Q_l) + Q_l p'(Q_l) - f'_i(q_{li}) = 0 \quad i \in I \quad (4.38)$$

Onde $I = \{i : q_{li} > 0\}$. Agora, considere o $EP(Q_l)$ e seja \hat{q} a solução única. Suponha que $\lambda(Q_l) < 0$. Então,

$$\lambda(Q_l) = p(Q_l) + \hat{q}_i p'(Q_l) - f'_i(\hat{q}_i) \quad i \in J \quad (4.39)$$

Além disso,

$$0 > \lambda(Q_l) \geq p(Q_l) - f'_i(0) \quad i \in \bar{J} \quad (4.40)$$

Onde $J = \{i : \hat{q}_i > 0\}$ e $\bar{J} = \{1, 2, \dots, N\} - J$.

Agora mostramos que $I = J$. Suponha que existe um $k \in I$ tal que $k \notin J$. Então,

$$\begin{aligned} 0 > p(Q_l) - f'_k(0) &\geq p(Q_l) - f'_k(q_{lk}) \\ &> p(Q_l) + Q_l p'(Q_l) - f'_k(q_{lk}) = 0 \end{aligned}$$

Assim, de (4.38) e (4.39) segue-se que: $\hat{q}_i > q_{li}$ para cada $i \in I$.

Mas por construção:

$$\sum_{i \in I} \hat{q}_i = \sum_{i \in I} q_{li} = Q_l$$

Isto é uma contradição. Por tanto, $\lambda(Q_l) \geq 0$.

A parte (b) segue-se do teorema 4.2 e os lemas 4.4 e 4.5. □

Como resultado final dessa análise matemática, fornecemos uma prova alternativa para a existência de uma solução de equilíbrio.

Teorema 4.4. *Suponha que $p(Q)$ seja estritamente decrescente, $Qp(Q)$ é côncava e que cada $f_i(q_i)$ é convexa, $q \geq 0$, $Q \geq 0$. Além disso, assumir que existem soluções (q_u, Q_u) e (q_l, Q_l) para os mercados competitivos e monopolísticos, ou seja, existem soluções para (4.34) e (4.35) respectivamente. Então, existe uma solução de equilíbrio (q^*, Q^*) para o mercado oligopolista com $Q_l \leq Q^* \leq Q_u$. Além disso, (q^*, Q^*) é único no caso de qualquer uma das três condições valha:*

- a) p é convexa
- b) f_i é estritamente convexa para cada $i = 1, \dots, N$
- c) $Qp(Q)$ é estritamente côncava,

tudo na linha real não negativa.

Demonstração. Pelo Lema 4.6 observa-se que $\lambda(Q_l) \geq 0 \geq \lambda(Q_u)$ e como λ é contínua (Lema 4.2), deve seguir-se que existe $Q^* \in [Q_l, Q_u]$ tal que $\lambda(Q^*) = 0$. Do teorema 4.2 segue-se que (q^*, Q^*) é uma solução de equilíbrio para o mercado oligopolista. A afirmação de unicidade segue do lema 4.5. \square

Corolário 4.5. *Suponha que existe uma solução (q_u, Q_u) para o problema competitivo (4.34). Se $f_i(0) < p(0)$ para cada $i = 1, 2, \dots, N$, então existe um equilíbrio (q^*, Q^*) para o mercado oligopolista com $0 < Q^* \leq Q_u$.*

Demonstração. Uma vez que $p(0) > f_i(0)$ para $i = 1, 2, \dots, N$, segue-se que $\lambda(0) > 0$. \square

Para resumir, pode-se começar com o intervalo $[Q_L, Q_U] = [Q_l, Q_u]$ e realizar uma pesquisa direta de bissecção para determinar uma raiz de λ . Para cada Q selecionado, o problema $EP(Q)$ é resolvido e $\lambda(Q)$ é assim avaliado. Uma solução de equilíbrio corresponde ao valor de Q para o qual $\lambda(Q) = 0$. Ao invés de um método de bissecção simples, descobrimos que se pode acelerar a busca alternando a interpolação linear uma vez que o limite superior inicial Q_U seja revisado. Quando isso ocorre, os valores subsequentes do iterar Q são determinados por:

$$Q = \frac{\lambda_L Q_U - \lambda_U Q_L}{\lambda_L - \lambda_U}$$

onde $\lambda_L = \lambda(Q_L)$ e $\lambda_U = \lambda(Q_U)$.

4.3 Exemplo de equilíbrio do modelo de Nash - Cournot

Como ilustração da metodologia proposta, nesta seção vamos apresentar um modelo de oligopolio com 5 empresas, em que cada uma das quais possui função de custo total da forma:

$$f_i(q_i) = c_i q_i + \frac{\beta_i}{\beta_i + 1} K_i^{-1/\beta_i} q_i^{(\beta_i+1)/\beta_i} \quad (4.41)$$

Observe que

$$f'_i(q_i) = c_i + K_i^{-1/\beta_i} q_i^{1/\beta_i} \quad (4.42)$$

Isto é, a curva de oferta marginal da empresa i é derivada de uma curva de oferta de elasticidade constante

$$q_i = K_i p_i^{\beta_i}$$

traduzido por uma constante c_i . A especificação dos parâmetros para estas cinco empresas é mostrada na Tabela 1. Além disso, para este exemplo, assume-se que a curva de demanda

Empresa	c_i	K_i	β_i
1	10	5	1.2
2	8	5	1.1
3	6	5	1.0
4	4	5	0.9
5	2	5	0.8

Tabela 1 – Especificação de parâmetros para o exemplo numérico

é dada por:

$$Q = 5000p^{-1.1} \quad \text{ou} \quad p(Q) = 5000^{1/1.1} Q^{-1/1.1} \quad (4.43)$$

Como vimos acima pode-se determinar a solução do equilíbrio de três formas:

- (i) Usamos o algoritmo chamado FDIPA (Feasible Direct Interior Point Algorithm) proposto por José Herskovits (ver [4], [5], [6]) para calcular a solução ótima do problema $EP(Q)$ segundo a metodologia proposta em [20].
- (ii) Usamos o algoritmo chamado FDA-MNCP (Feasible Direct Algorithm) (proposto em [19], [18]) para resolver o $MiCP(G, H)$ definido por (4.9)-(4.10).
- (iii) Finalmente, usamos o algoritmo FDA-NCP proposto por Mazorche e Herskovits (ver [18], [7]) para resolver o $NCP(H)$ dado por (4.11).

Assim, a solução de equilíbrio de Nash (estimativa) usando os três algoritmos citados acima, com tolerância de 10^{-6} , é dada pelos valores de q_i^* descritos na tabela 2 para cada empresa $i = 1, \dots, 5$, onde a saída de equilíbrio total é dada por $Q^* = 204.2966$.

Empresa	q_i^*
1	36.9328
2	41.8184
3	43.7068
4	42.6594
5	39.1791

Tabela 2 – Estimativas das soluções de equilíbrio de Nash.

Para inicializar o procedimento algorítmico descrito em (i) são obtidos limites superiores e inferiores sobre a produção total do oligopólio. Esses limites correspondem a soluções ótimas para (4.34) e (4.35), Ou seja, as soluções para estruturas de mercado competitivas e monopolísticas. As soluções ótimas para (4.34) e (4.35) resultam em $Q_u = 235.4$ e $Q_l = 40$, respectivamente. Assim, pelo Lema 4.6, a saída de equilíbrio oligopolístico total está no intervalo $[Q_l, Q_u] \equiv [40.0, 235.4]$.

Logo, usando o algoritmo FDIPA para resolver problemas $EP(Q)$, obtemos 6 iterações para calcular o número de $Q \geq 0$, onde para cada $Q \geq 0$ calculado; o FDIPA resolveu o problema $EP(Q)$ com 5 iterações, dando um total de 30 iterações para determinar a solução de equilíbrio dadas pelos valores de q_i^* na tabela 2.

Usando o algoritmo FDA-MNCP descrito em (ii), obtemos 8 iterações para determinar a solução de equilíbrio dada na tabela 2.

E o algoritmo FDA-NCP descrito em (iii), fez 12 iterações para calcular os valores da solução de equilíbrio q_i^* como mostra a tabela 2.

Foi usado o mesmo ponto inicial para os algoritmos FDIPA e FDA-MNCP, a saber $x_0 = (20, 20, 20, 20, 20)$ e para o algoritmo FDA-NCP foi usado o ponto inicial $x_0 = (100, 100, 100, 100, 100)$. Da tabela 3 observa-se que o algoritmo FDA-MNCP teve o melhor desempenho, pois fez 8 iterações para determinar a solução de equilíbrio.

Algoritmo	Iterações para encontrar q_i^*
FDIPA	30
FDA-MNCP	8
FDA-NCP	12

Tabela 3 – Comparação dos Algoritmos para calcular o Equilíbrio de Nash.

5 CONCLUSÕES E TRABALHOS FUTUROS

Neste trabalho foi muito importante estabelecer condições necessárias e suficientes para a existência de solução de uma inequação variacional (VI) e sua equivalência com vários casos especiais de problemas de complementaridade (CP), assim como também, com programas de otimização convexos; em particular, vimos a interconexão entre uma VI e um problema de complementaridade misto (MiCP). Além disso, mostramos que o cálculo de um equilíbrio de Nash pode ser realizado resolvendo uma VI. Portanto, estas relações, nos permitem determinar soluções numéricas para um caso de um problema de equilíbrio de Nash, que pode ser resolvido como um MiCP, que por sua vez pode ser reescrito como um NCP, ou resolvendo uma série de programas convexos, usando os métodos descritos em [20] e desenvolvidos neste trabalho. Cabe resaltar que só exibimos os resultados numéricos sem aprofundar no estudo dos algoritmos aqui usados, no futuro pretendo continuar nesta parte de algoritmos e análise numérico que ficou pouco explorada, por exemplo outros problemas com mais empresas, colocar outros parâmetros na tabela 1, etc, para realizar um estudo computacional que indique um algoritmo de melhor performance, fica também como trabalho futuro comparar o melhor algoritmo com otimização multiobjetivo para resolver equilíbrios de Nash.

REFERÊNCIAS

- [1] BOTELHO, G.; PELLEGRINO, D.; TEIXEIRA, E. *Fundamentos de Análise Funcional*. Sociedade Brasileira de Matemática, Rio de Janeiro: SBM, 2015.
- [2] CHARALAMBOS, D.; SUBIR, K. *Games and Decision Making*. Departments of Economics & Mathematics Purdue University, New York, 1998.
- [3] FACCHINEI, F.; SHI PANG, J. *Finite-Dimensional Variational Inequalities and Complementarity Problems, Vol. 1*. Springer, New York, 2003.
- [4] HERSKOVITS, J. *A two-stage feasible directions algorithm for non-linear constrained optimization*. Research Report 103, INRIA, Le Chesnay, France, 1982.
- [5] HERSKOVITS, J. *A two-stage feasible directions algorithm including variable metric techniques for nonlinear optimization*. Research Report 118, INRIA, Le Chesnay, France, 1982.
- [6] HERSKOVITS, J. *A two-stage feasible directions algorithm for non-linear constrained optimization*. Mathematical Programming, 36:19–38, 1986.
- [7] HERSKOVITS, J. ; MAZORCHE, S.A. *Feasible Directions Algorithm for Nonlinear Complementarity Problem and Applications in Mechanics*. Struct Multidisc Optim, 2009.
- [8] IZMAILOV, A.; SOLODOV, M. *Otimização, Vol. 1*. Rio de Janeiro: IMPA, 2009.
- [9] KELLY, A. *Decision Making using Game Theory*. Cambridge University Press, New York, 2003.
- [10] KINDERLEHRER, D.; STAMPACCHIA, G. *An Introduction to variational inequalities and their applications*. Academic Press, New York, 1980.
- [11] KREYSZIG, E. *Introductory Functional Analysis with Applications*. John Wiley & Sons, 1989.
- [12] LIMA, E. L. *Espaços Métricos*. Projeto Euclides, Rio de Janeiro: IMPA, 2015.
- [13] LIMA, E. L. *Álgebra Linear*. Coleção Matemática Universitária, Rio de Janeiro: IMPA, 2014.
- [14] LIMA, E. L. *Curso de Análise, vol. 1*. Projeto Euclides, Rio de Janeiro: IMPA, 2014.
- [15] LIONS, J. L.; STAMPACCHIA, G. *Variational Inequalities*. Communications on Pure and Applied Mathematics 20 (1967) 493–519.
- [16] MANCINO, O. G.; STAMPACCHIA, G. *Convex Programming and Variational Inequalities*. Journal of Optimization Theory and Applications 9 (1972) 3–23.
- [17] MARTÍNEZ, J.; SANTOS, S. *Métodos Computacionais de Otimização*. Departamento de Matemática Aplicada, IMECC-UNICAMP, 1995.
- [18] MAZORCHE, S. R. *Algoritmos para Problemas de Complementaridade não Linear*. Tese de Doutorado. COPPE, UFRJ, Rio de Janeiro, 2007.

- [19] MAZORCHE, S. R.; HERSKOVITS, J. *A feasible point technique for mixed nonlinear complementarity problems*. In Joint EUROPT-OMS Conference 2007, 2nd Conference on Optimization Methods Software and 6th EUROPT Workshop on Advances in Continuous Optimization, Prague, Czech Republic, 2007.
- [20] MURPHY, H.; SHERALI, D.; SOYSTER, L. *A Mathematical Programming Approach for Determining Oligopolistic Market Equilibrium*. Mathematical Programming North-Holland Publishing Company, 1982.
- [21] RIBEIRO, A.; WEGNER, E. *Otimização Contínua: Aspectos Teóricos e Computacionais*. Cengage Learning, São Paulo, 2013.
- [22] STAMPACCHIA, G. *Formes bilineares coercives sur les ensembles convexes*. Comptes Rendus Academie Sciences Paris 258 (1964) 4413–4416.
- [23] STAMPACCHIA, G. *Variational inequalities*. In Theory and Applications of Monotone Operators (Proc. NATO Advanced Study Inst., Venice, 1968), Edizioni “Oderisi” (Gubbio 1969) pp. 101–192.