André Luiz Barbosa Arantes dos Santos

CONJUGAÇÃO, CONDIÇÕES DE OTIMALIDADE E TEORIA DA FIRMA

Curitiba
2008

André Luiz Barbosa Arantes dos Santos

CONJUGAÇÃO, CONDIÇÕES DE OTIMALIDADE E TEORIA DA FIRMA

Monografia apresentada à coordenação do curso de Matemática Industrial, da Universidade Federal do Paraná, como parte dos requisitos para obter aprovação na Disciplina de Projeto de Matemática Industrial.
Orientadora:
Prof^{a}. Dr^{a}. Elizabeth Wegner Karas.
Co-orientador:
Prof. Dr. Ademir Alves Ribeiro.

Curitiba
2008

André Luiz Barbosa Arantes dos Santos

**CONJUGAÇÃO, CONDIÇÕES DE OTIMALIDADE
E TEORIA DA FIRMA**

Monografia aprovada como requisito parcial para a obtenção do grau de bacharel no curso de Matemática Industrial, Setor de Exatas da Universidade Federal do Paraná.

Banca Examinadora:

Profª. Drª. Elizabeth Wegner Karas
Departamento de Matemática - UFPR

Prof. Dr. Wilfredo Sosa
IMCA - Peru

Prof. MSc. Rodrigo Garcia Eustáquio
Departamento de Matemática - UTFPR

Curitiba, 2 de Dezembro de 2008.

Sumário

Introdução

Nosso trabalho de conclusão do curso de Bacharelado em Matemática Industrial insere-se no contexto de Otimização que é uma das grandes áreas que estudamos ao longo do curso.

Otimização consiste em determinar os mínimos ou os máximos de uma função de várias variáveis com valores dentro de uma determinada região do espaço multidimensional [19].

Apresentamos algumas ferramentas de análise convexa, como o conceito de cones, cone polar e Lema de Farkas. Dedicamos um capítulo ao conceito de função conjugada no sentido de Fenchel, [3], e discutimos suas propriedades. Discutimos o problema geral de programação não linear diferenciável e as clássicas condições de otimalidade de Karush-Kuhn-Tucker (KKT). Estas condições foram inicialmente provadas por Karush em 1939 em uma tese de mestrado, porém esta tese foi esquecida. Em 1951 Kuhn e Tucker voltaram a provar estas condições [14]. Assim estas condições passaram as ser conhecidas como condições de Karush-Kuhn-Tucker. Em nosso trabalho apresentamos uma prova das condições de KKT através de uma abordagem cônica, em que o Lema de Farkas é de fundamental importância. Consideramos uma hipótese fraca, mas difícil de ser verificada, que é a igualdade, no ponto ótimo, do polar do cone viável linearizado e do polar do cone tangente. Apresentamos, em seguida, algumas condições de qualificação e as relações de implicação entre elas que são estudadas em profundidade em [2, 20].

Finalmente, procuramos aplicar estes conceitos estudados a problemas de Economia. Apresentamos três diferentes modelos da Teoria da Firma [11, 15]: a concorrência perfeita, o monopólio, e o oligopólio. Estes modelos são descritos como problemas de programação não linear. Discutimos, então, as condições de otimalidade de KKT destes modelos.

Cabe ressaltar que em toda a discussão procuramos enfatizar o aspecto geométrico da teoria.

Estrutura dos capítulos. No Capítulo 1 apresentamos alguns resultados de análise, álgebra linear e análise convexa. Vemos, também, uma breve introdução à teoria de cones além de apresentarmos o Lema de Farkas. No Capítulo 2 descrevemos em profundidade a função conjugada e suas propriedades. No Capítulo 3 apresentamos o problema de programação não linear e abordamos alguns cones importantes relacionados ao problema e as relações entre eles. Também discutimos a condição de otimalidade para problemas sem restrições, um resultado de minimização convexa e a demonstração das condições de KKT. As diversas condições de qualificação estudadas e suas implicações serão vistas ao final do capítulo. No Capítulo 4 apresentamos a teoria da firma para três diferentes modelos: concorrência perfeita, monopólio e oligopólio. Abordaremos essas três formas com as condições de

KKT. A seguir discutimos o problema da economia sob ótica da função conjugada. Ao final apresentamos algumas conclusões.

Capítulo 1

Ferramentas Matemáticas

Inicialmente veremos alguns conceitos básicos que são de grande importância para o desenvolvimento deste trabalho. Estes conceitos nos servirão de ferramentas para obter vários resultados nos capítulos subseqüentes. Veremos alguns resultados de análise e de álgebra linear e em seguida, a definição de supremo de um conjunto e certas propriedades. Depois veremos a definição de conjunto convexo e função convexa. Por fim apresentamos a definição de cone e algumas propriedades. As principais referências deste capítulo são [7, 8, 13, 16, 17, 18].

1.1 Conceitos básicos

Primeiro começaremos a tratar sobre algumas definições de Análise.

Definição 1.1 *Uma seqüência em $\mathbb{R}^n$ é uma aplicação $k \in \mathbb{N} \mapsto x^k \in \mathbb{R}^n$, definida no conjunto $\mathbb{N}$ dos números naturais.*

Denotaremos uma seqüência por $(x^k)_{k\in\mathbb{N}}$, ou por (x^k).

Definição 1.2 *Dizemos que o ponto $a \in \mathbb{R}^n$ é o limite da seqüência de pontos $x^k \in \mathbb{R}^n$ quando, para todo $\varepsilon > 0$ dado, é possível obter $k_0 \in \mathbb{N}$ tal que*

$$k \geq k_0 \Rightarrow \|x^k - a\| < \varepsilon$$

e denotamos $\lim\limits_{k\to\infty} x^k = a$.

A próxima definição diz respeito a sequências em $\mathbb{R}$.

Definição 1.3 *Dizemos que o ponto $a \in \mathbb{R}$ é o limite inferior de uma seqüência $x^k \in \mathbb{R}$ quando, para todo $\varepsilon > 0$ dado, é possível obter $k_0 \in \mathbb{N}$ tal que*

$$k \geq k_0 \Rightarrow x^k - a > -\varepsilon$$

e denotamos $\liminf\limits_{k\to\infty} x^k = a$. Analogamente, dizemos que $a \in \mathbb{R}$ é o limite superior de $x^k \in \mathbb{R}$ quando, para todo $\varepsilon > 0$ dado, é possível obter $k_0 \in \mathbb{N}$ tal que

$$k \geq k_0 \Rightarrow x^k - a < \varepsilon$$

e denotamos $\limsup\limits_{k\to\infty} x^k = a$.

Em particular, quando o limite superior e o limite inferior da sequência (x^k) coincidem, a sequência converge para este limite.

Definição 1.4 *Um ponto* $a \in I\!R^n$ *é dito aderente a um conjunto* $A \subset I\!R^n$ *quando é limite de uma seqüência de pontos desse conjunto, isto é, se qualquer vizinhança de* a *contém algum elemento de* A*. O conjunto dos pontos aderentes a* A *é chamado de fecho de* A*.*

O fecho de um conjunto A será denotado por $cl(A)$.

Definição 1.5 *Um conjunto* A *é fechado quando* $cl(A) = A$*.*

A seguir alguns resultados de Álgebra Linear.

Definição 1.6 *Considere* $A \in I\!R^{m\times n}$*. O subespaço de* $I\!R^n$ *gerado pelos vetores-linha de* A *é chamado espaço-linha de* A *e o subespaço de* $I\!R^m$ *gerado pelos vetores-coluna de* A *é chamado espaço-coluna de* A*.*

Denotaremos a matriz transposta de A por A^t.

Definição 1.7 *O posto de uma matriz* $A \in I\!R^{m\times n}$*, denotado por* $posto(A)$*, é a dimensão comum do espaço-linha e do espaço-coluna.*

Mais adiante precisaremos da definição de semicontinuidade inferior.

Definição 1.8 *Dizemos que uma função* $f : D \to I\!R$ *é semi contínua inferiormente no ponto* $x \in D \subset I\!R^n$*, quando para qualquer seqüência* $x^k \in D$ *tal que* $x^k \to x$ *quando* $k \to \infty$*, tem-se que*

$$\lim_{k\to\infty} \inf f(x^k) \geq f(x).$$

Dizemos que f *é semi contínua superiormente quando, nas mesmas condições,*

$$\lim_{k\to\infty} \sup f(x^k) \leq f(x).$$

As propriedades de uma função semi-contínua são vistas mais intensamente em [10, 13].

1.2 Supremo de um conjunto

Para abordarmos a função conjugada, necessitamos da definição de supremo e de algumas propriedades.

Definição 1.9 *Dizemos que* b *é uma cota superior de um conjunto* $X \subset I\!R$ *se* b *for maior do que qualquer elemento de* X*.*

Definição 1.10 *Seja* $X \subset I\!R$ *limitado superiormente e não-vazio. Chamamos* $b \in I\!R$ *de supremo do conjunto* X *quando* b *é a menor cota superior de* X*, em outras palavras,* b *é supremo se cumpre duas condições:*

(i) b é cota superior de X;

(ii) se $c < b$ então existe $x \in X$ com $c < x \leq b$.

Escreveremos $b = \sup X$ para denotar que b é o supremo do conjunto X.

Veremos, a seguir, algumas propriedades do supremo que seguem diretamente da definição.

Proposição 1.11 *Considere $X \subset I\!R$ um conjunto não vazio limitado superiormente e $\beta \in I\!R$. Valem as seguintes afirmações:*

(i) $\sup\{X + \beta\} = \sup X + \beta$

(ii) $\sup\{\beta X\} = \beta \sup X;\ \beta > 0$

Prova. (i) Considere $b = \sup X$, queremos provar que $b+\beta$ é a menor cota superior do conjunto $X + \beta$. Para provar que é uma cota superior note que

$$b \geq x, \forall x \in X.$$

Se somarmos β em ambos os lados da desigualdade temos

$$b + \beta \geq x + \beta,$$

o que mostra que $b + \beta$ é uma cota superior de $X + \beta$. Iremos, agora, provar que $b + \beta$ é a menor cota superior. Consideremos $d \in I\!R$ menor do que $b + \beta$, logo

$$d < b + \beta.$$

Tome

$$c = d - \beta < b.$$

Pela Definição 1.10, como $b = \sup X$, existe $x \in X$ tal que

$$c < x \leq b.$$

Se somarmos β em todos os membros das desigualdades

$$c + \beta < x + \beta \leq b + \beta$$

e conseqüentemente

$$d < x + \beta \leq b + \beta$$

portanto, $b+\beta$ é a menor cota superior de $X+\beta$. Prova do item (ii). Analogamente precisamos demonstrar que $b\beta$ é a menor cota superior do conjunto $X\beta$. Temos que

$$b \geq x,\ \forall x \in X.$$

Mutiplicando por β ambos os membros da desigualdade, obtemos

$$b\beta \geq x\beta,$$

pois $\beta > 0$. Logo $b\beta$ é uma cota superior de $X\beta$. Agora provaremos que $b\beta$ é a menor cota superior. Consideremos $d \in I\!R$ menor do que $b\beta$, ou seja,

$$d < b\beta.$$

Tome

$$c = \frac{d}{\beta} < b.$$

Pela Definição 1.10, como $b = \sup X$, existe $x \in X$ tal que

$$c < x \leq b$$

se multiplicarmos por $\beta > 0$

$$c\beta < x\beta \leq b\beta$$

$$d < x\beta \leq b\beta$$

assim, $b\beta$ é a menor das cotas superiores de $X\beta$. □

Para maiores referências sobre o supremo de um conjunto veja [17].

1.3 Convexidade

A convexidade é uma noção importantíssima para a otimização. Com as hipóteses de convexidade qualquer minimizador local é global e, além disso, podemos desenvolver a teoria de dualidade de maneira mais fácil, ou seja, podemos associar ao problema original outro problema, chamado de dual, que sob certas hipóteses é equivalente ao primal e muitas vezes mais fácil de resolver do que o problema original.

Definição 1.12 *Um subconjunto* $\Omega \subset I\!R^n$ *é dito convexo se e somente se para todo* $x, y \in \Omega$ *e* $\alpha \in [0,1]$ *se verifica que*

$$\alpha x + (1-\alpha)y \in \Omega.$$

Definição 1.13 *Considere* $\Omega \subset I\!R^n$ *um conjunto convexo. Dizemos que a função* $f : \Omega \to I\!R$ *é convexa quando*

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

para todo par de pontos (x, y) *em* Ω *e* $\alpha \in [0,1]$.

A função convexa é largamente abordada em [10, 13]. No caso em que a desigualdade vale de forma estrita, dizemos que f é estritamente convexa.

Na Figura 1.1 ilustramos a Definição 1.13. Podemos ver que o gráfico da função f no ponto $\alpha x + (1-\alpha)y$ está abaixo da reta.

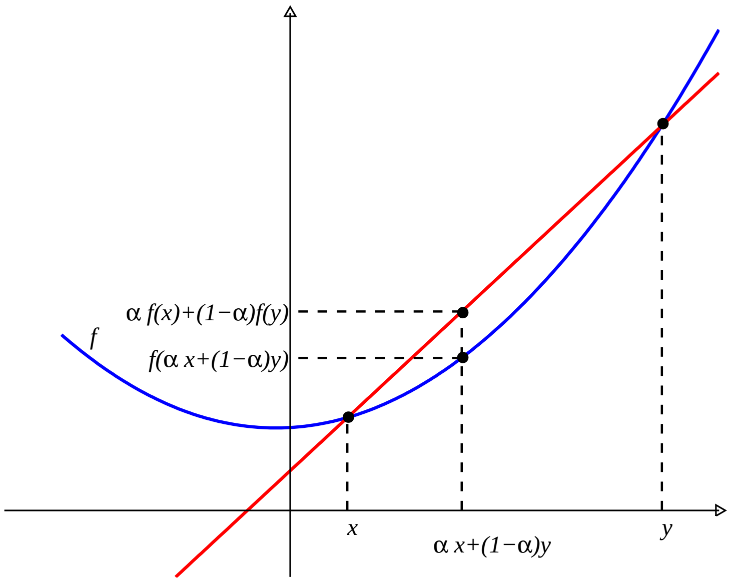

Figura 1.1: Função convexa.

1.4 Cones

Passaremos aqui a analisar a teoria de cones, pois esta é de suma importância para a obtenção das condições de otimalidade de KKT. Para maiores informações vide [2, 13].

Definição 1.14 *Um subconjunto $C \subset \mathbb{R}^n$ é dito cone quando, para todo $t > 0$ e $d \in C$ tem-se $td \in C$.*

Definição 1.15 *Dado um conjunto $S \subset \mathbb{R}^n$, definimos o polar de S por*

$$P(S) = \left\{p \in \mathbb{R}^n \mid p^t x \leq 0, \forall x \in S\right\}.$$

Nas Figuras 1.2 e 1.4 ilustramos um cone e o cone polar de um conjunto.

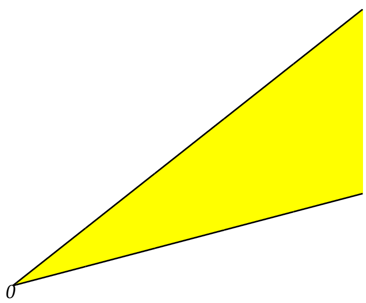

Figura 1.2: Ilustração da Definição 1.14.

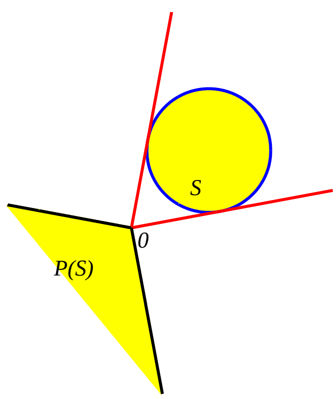

Figura 1.3: Polar de um conjunto S.

Lema 1.16 *Sejam A e B subconjuntos do $I\!R^n$. Se $A \subseteq B$, então $P(B) \subseteq P(A)$.*

Prova. Considere $y \in P(B)$ qualquer. Com isso, temos $y^t x \leq 0$ para todo $x \in B$. Como $A \subseteq B$, segue que $y^t z \leq 0$ para todo $z \in A$. Portanto $y \in P(A)$. □

Proposição 1.17 *Considere $C \subset I\!R^n$ um cone não vazio. O polar de C é convexo e fechado.*

Lema 1.18 *Se $S \subset I\!R^n$, então $S \subset P(P(S))$.*

Prova. Considere $u \in S$ e $A = P(S)$. Queremos mostar que $u \in P(P(S)) = P(A)$. Dado $p \in A$ temos $u^t p \leq 0$. Logo $u \in P(A)$. □

Veremos que adicionado algumas hipóteses sobre o conjunto S, o Lema de Farkas garante a igualdade $S = P(P(S))$. Para isso mostraremos duas versões do Lema de Farkas.

Lema 1.19 (Versão Geométrica do Lema de Farkas) *Seja $C \subset I\!R^n$ um cone convexo e fechado. Então $C = P(P(C))$.*

Lema 1.20 (Versão Algébrica do Lema de Farkas) *Considere $A \in I\!R^{n\times m}$ e $c \in I\!R^n$. Então um dos dois sistemas tem solução:*

$$Ax \leq 0 \quad e \quad c^t x > 0$$

ou

$$A^t y = c \quad e \quad y \geq 0.$$

Seja $C = \{A^t y = c; y \geq 0\}$ com $A \in I\!R^{n\times m}$ um cone finitamente gerado. Então as duas versões do Lema de Farkas, 1.19 e 1.20, são equivalentes.

Capítulo 2

Função Conjugada

2.1 Definição e interpretação geométrica

Neste capítulo introduzimos o conceito de conjugada de uma função e discutimos suas propriedades.

O conceito de conjugação foi introduzido por Fenchel em [3] e é um instrumento importante em otimização. Há vários métodos para se tentar resolver um problema de otimização. Um deles é o Método de Penalidade, que consiste basicamente em minimizar a cada iteração uma função penalidade, que é a soma da função objetivo do problema original com a penalidade das restrições. O Método de Penalidade é explorado em [1] e [4].

Ao problema original, dito problema primal, podemos associar um problema dual que pode ser resolvido pelo Método de Ponto Proximal. Tal método consiste em resolver uma seqüência de problemas irrestritos onde aparece a conjugada da função penalidade. Os Métodos de Penalidade e de Ponto Proximal estão relacionados através da operação de conjugação da função penalidade. Para maiores informações veja [12].

Outra utilização da conjugada é na economia, em que queremos maximizar o lucro. Analisaremos esta utilização na Seção 4.4.

É conveniente considerarmos a função f existente e definida em todo o espaço $\mathbb{R}^n$, permitindo o valor ∞ para $f(x)$. Considere $X \subseteq \mathbb{R}^n$ e $f : X \to \mathbb{R}$. Nós podemos estender f, para além de X, definindo

$$f_e(x) = \begin{cases} f(x), & se \quad \forall\, x \in X, \\ \infty, & se \quad \forall\, x \notin X. \end{cases}$$

A função valor-estendido envia $\mathbb{R}^n$ para $\mathbb{R} \cup \{\infty\}$ e está ilustrada na Figura 2.1. Até o final do capítulo toda função será uma função valor-estendido e subíndice "e" será, por esta razão, omitido. Isto posto podemos analisar a definição conjugada.

Definição 2.1 *Considere $f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ uma função, não identicamente igual a ∞. A conjugada de f, denotada por f^*, é dada por*

$$f^*(s) = \sup_x \{\langle x, s\rangle - f(x) \mid x \in \mathbb{R}^n\}. \tag{2.1}$$

Em particular, considere $f : \mathbb{R} \to \mathbb{R} \cup \{\infty\}$ uma função diferenciável. Note que para um dado $s \in \mathbb{R}$ o supremo em (2.1) ser atingido em um certo $x \in \mathbb{R}$

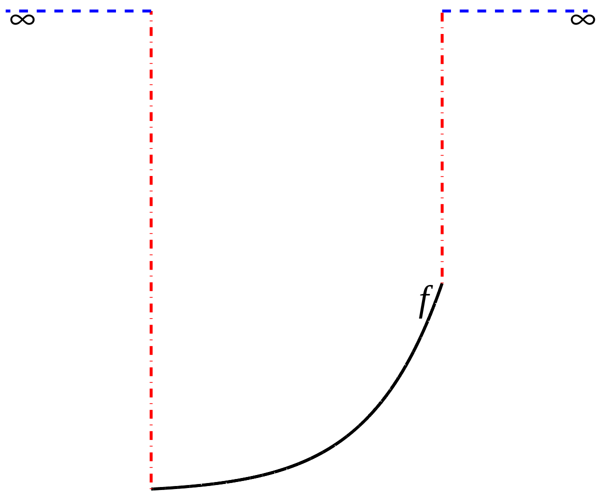

Figura 2.1: Função de $I\!R$ em $I\!R \cup \{\infty\}$

equivale a dizer que $f^*(s) = sx - f(x)$. Ou ainda, temos

$$\frac{d}{dx}\left(sx - f(x)\right) = s - f'(x) = 0.$$

Assim, se $f : I\!R \to I\!R \cup \{\infty\}$ é uma função diferenciável, temos que

$$f^*(s) = sx - f(x) \Leftrightarrow f'(x) = s. \tag{2.2}$$

Tendo em vista (2.2) podemos obter a conjugada sem a utilização do supremo.

Exemplo 1 *Calculemos a conjugada da função exponencial,* $f(x) = e^x$.

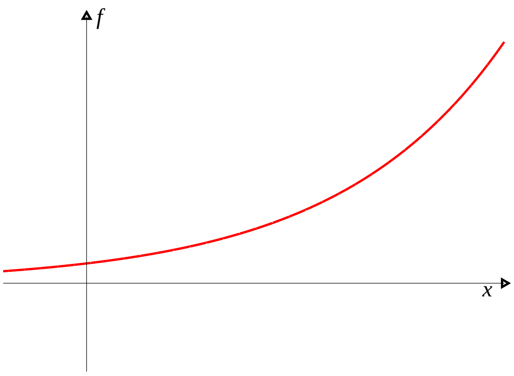

Figura 2.2: Função $f(x) = e^x$.

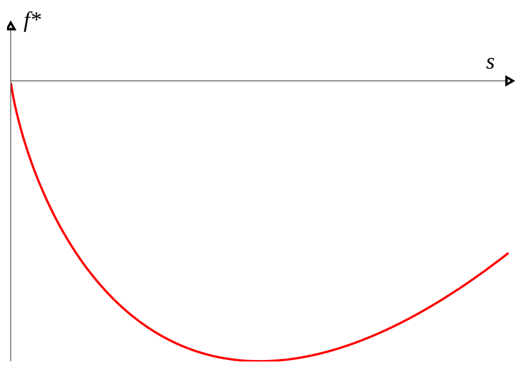

Figura 2.3: Conjugada da $f(x) = e^x$.

Pela Definição 2.1

$$\begin{aligned} f^*(s) &= \sup_x\{\langle x, s\rangle - f(x)\} \\ &= \sup_x\{\langle x, s\rangle - e^x\}. \end{aligned}$$

Derivando o argumento do supremo em relação a x e igualando a 0, obtemos

$$s - e^x = 0 \Leftrightarrow x = \ln s.$$

Voltando na definição da conjugada

$$f^*(s) = s \ln s - e^{\ln s} = s \ln s - s.$$

Ilustramos a função $f(x) = e^x$ e sua conjugada nas Figuras 2.2 e 2.3, respectivamente.

A seguir abordaremos a interpretação geométrica da função conjugada. A equação da reta tangente em $(x_0, f(x_0))$ com inclinação s é

$$y - f(x_0) = s(x - x_0).$$

Quando $x = 0$

$$y - f(x_0) = -sx_0,$$

ou seja,

$$y = -(sx_0 - f(x_0)).$$

Usando (2.2) temos que

$$y = -f^*(s).$$

Isto significa que a reta, com inclinação s que tangencia f, intercepta o eixo das ordenadas em $(0, -f^*(s))$, como ilustrado na Figura 2.4.

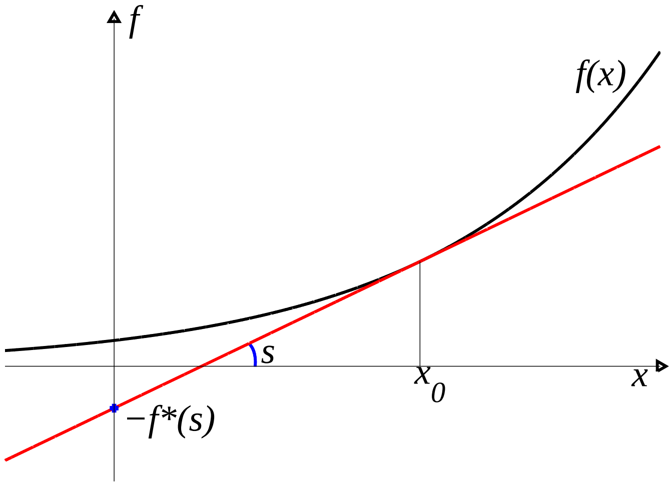

Figura 2.4: Função conjugada em um ponto.

Adiante veremos as propriedades de função conjugada e suas demonstrações.

2.2 Propriedades

Considere f e g duas funções. As seguintes afirmações são verdadeiras.

Proposição 2.2 *A função conjugada $f : I\!R \to I\!R \cup \{\infty\}$ é convexa.*

Prova. Considere $\alpha \in (0,1)$ e $s_1, s_2 \in I\!R$ quaisquer. Então

$$\begin{aligned}
f^*((1-\alpha)s_1+\alpha s_2) &= \sup_x\{\langle(1-\alpha)s_1+\alpha s_2, x\rangle - f(x)\} \\
&= \sup_x\{(1-\alpha)\langle s_1, x\rangle + \alpha\langle s_2, x\rangle - f(x)\} \\
&= \sup_x\{(1-\alpha)\langle s_1, x\rangle + \alpha\langle s_2, x\rangle - ((1-\alpha)f(x)+\alpha f(x))\} \\
&= \sup_x\{(1-\alpha)(\langle s_1, x\rangle - f(x)) + \alpha(\langle s_2, x\rangle - f(x))\} \\
&\leq (1-\alpha)\sup_x\{\langle s_1, x\rangle - f(x)\} + \alpha\sup_x\{\langle s_2, x\rangle - f(x)\} \\
&= (1-\alpha)f^*(s_1) + \alpha f^*(s_2)
\end{aligned}$$

completando assim a prova. □

Proposição 2.3 *A conjugada da função* $g(x) = f(x) + \alpha$ *é*

$$g^*(s) = f^*(s) - \alpha.$$

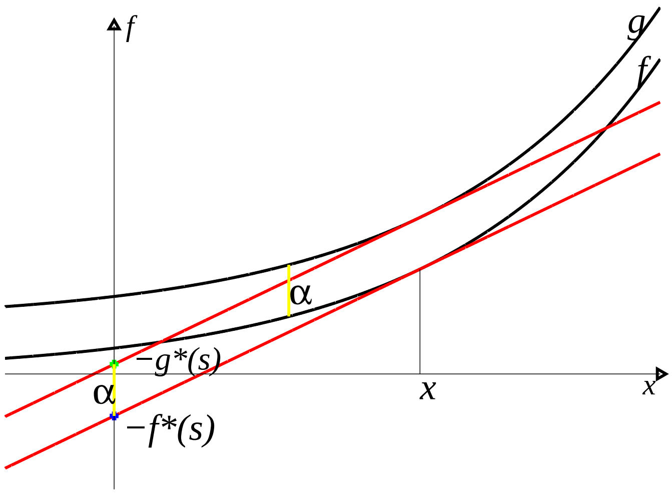

Figura 2.5: Ilustração da Proposição 2.3

Prova. Pela Definição 2.1 a conjugada de uma função g é

$$g^*(s) = \sup_x\{\langle x, s\rangle - g(x)\},$$

como $g(x) = f(x) + \alpha$

$$\begin{aligned}
g^*(s) &= \sup_x\{\langle x, s\rangle - (f(x)+\alpha)\} \\
&= \sup_x\{\langle x, s\rangle - f(x) - \alpha\} \\
&= \sup_x\{(\langle x, s\rangle - f(x)) - \alpha\}.
\end{aligned}$$

Pela Proposição 1.11, temos que

$$g^*(s) = \sup_x\{\langle x, s\rangle - f(x)\} - \alpha.$$

Porém a conjugada de f é $f^*(s) = \sup\limits_x \{\langle x, s\rangle - f(x)\}$, então

$$g^*(s) = f^*(s) - \alpha.$$

$\square$

Proposição 2.4 *A conjugada da função $g(x) = \alpha f(x)$, com $\alpha > 0$, é*

$$g^*(s) = \alpha f^*\left(\frac{s}{\alpha}\right).$$

Prova. A conjugada de g é dada por

$$\begin{aligned} g^*(s) &= \sup_x \{\langle x, s\rangle - g(x)\} \\ &= \sup_x \{\langle x, s\rangle - \alpha f(x)\} \\ &= \sup_x \left\{\alpha \left(\langle x, \frac{s}{\alpha}\rangle - f(x)\right)\right\}. \end{aligned}$$

Usando a Proposição 1.11, temos que

$$g^*(s) = \alpha \sup_x \left\{\langle x, \frac{s}{\alpha}\rangle - f(x)\right\} = \alpha f^*\left(\frac{s}{\alpha}\right).$$

$\square$

Proposição 2.5 *A conjugada da função $g(x) = f(\alpha x)$, com $\alpha \neq 0$, é*

$$g^*(s) = \alpha f^*\left(\frac{s}{\alpha}\right).$$

Prova. Temos que

$$g^*(s) = \sup_x \{\langle x, s\rangle - g(x)\} = \sup_x \{\langle x, s\rangle - f(\alpha x)\}.$$

Tomando $\alpha x = u$, ficamos com

$$\begin{aligned} g^*(s) &= \sup_u \left\{\langle \frac{u}{\alpha}, s\rangle - f(u)\right\} \\ &= \sup_u \left\{\langle u, \frac{s}{\alpha}\rangle - f(u)\right\} \\ &= f^*\left(\frac{s}{\alpha}\right). \end{aligned}$$

$\square$

Proposição 2.6 *A conjugada da função $g(x) = f(x - x_0)$ é*

$$g^*(s) = f^*(s) + \langle x_0, s\rangle$$

com $x_0 \in I\!R^n$.

Prova. Sabemos que

$$g^*(s) = \sup_x \{\langle x, s\rangle - g(x)\},$$

logo

$$g^*(s) = \sup_x \{\langle x, s\rangle - f(x - x_0)\}$$

$$g^*(s) - \langle x_0, s\rangle = \sup_x \{\langle x, s\rangle - f(x - x_0)\} - \langle x_0, s\rangle.$$

Pela Proposição 1.11

$$\begin{aligned} g^*(s) - \langle x_0, s\rangle &= \sup_x \{\langle x, s\rangle - f(x - x_0) - \langle x_0, s\rangle\} \\ &= \sup_x \{\langle (x - x_0), s\rangle - f(x - x_0)\} \\ &= f^*(s), \end{aligned}$$

assim

$$g^*(s) = f^*(s) + \langle x_0, s\rangle.$$

□

A Proposição 2.6 translada a função f horizontalmente e, então, a conjugada de f se modifica somando $\langle s, x_0\rangle$. Vemos, na Figura 2.6, que um triângulo se forma no gráfico. Um dos catetos é x_0 e o outro é sx_0 e que o ângulo interno α do triângulo coincide com a inclinação das retas que tangenciam f e g.

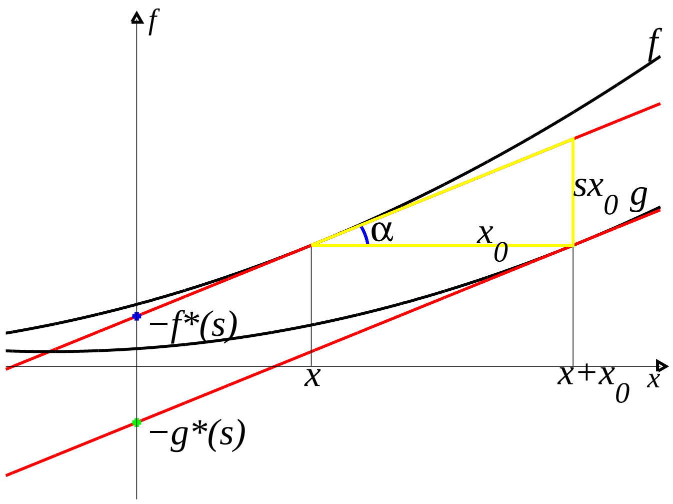

Figura 2.6: Ilustração da Proposição 2.6

Proposição 2.7 *A conjugada da função* $g(x) = f(x) + xs_0$ *é*

$$g^*(s) = f^*(s - s_0).$$

Prova. Temos que

$$g^*(s) = \sup_x \{\langle x, s\rangle - g(x)\}$$

logo

$$\begin{aligned} g^*(s) &= \sup_x\{\langle x, s\rangle - (f(x) + \langle x, s_0\rangle)\} \\ &= \sup_x\{\langle x, s\rangle - \langle x, s_0\rangle - f(x)\} \\ &= \sup_x\{\langle x, s - s_0\rangle - f(x)\} \\ &= f^*(s - s_0). \end{aligned}$$

□

Proposição 2.8 *Consideremos f_1 e f_2 duas funções convexas. Se $f_1 \leq f_2$ então $f_1^* \geq f_2^*$.*

Prova. Se f_1 satisfaz a Definição 2.1, então

$$f_1^*(s) = \sup_x\{\langle \bar{x_1}, s\rangle - f_1(\bar{x_1})\} \geq \langle x, s\rangle - f_1(x)$$

para todo x. Em especial para $x = \bar{x_2}$ tal que $f_2(\bar{x_2}) = s$ e satisfazendo a Definição 2.1, teremos

$$f_1^*(s) = \sup_x\{\langle \bar{x_1}, s\rangle - f_1(\bar{x_1})\} \geq \langle \bar{x_2}, s\rangle - f_2(\bar{x_2}) = \sup_x\{\langle \bar{x_2}, s\rangle - f_2(\bar{x_2})\} = f_2^*(s),$$

ou seja,

$$f_1^*(s) \geq f_2^*(s).$$

□

Proposição 2.9 *Se $dom f_1 \cap dom f_2 \neq \emptyset$ e $\alpha \in (0, 1)$, então*

$$[\alpha f_1 + (1 - \alpha) f_2]^* \leq \alpha f_1^* + (1 - \alpha) f_2^*.$$

Prova. Sejam $g(x) = \alpha f_1 + (1 - \alpha) f_2$, com f_1, f_2 satisfazendo (2.2), e $g'(\bar{x}) = s$, então temos que

$$f_1^*(s) = \sup_x\{\langle \bar{x_1}, s\rangle - f_1(\bar{x_1})\} \geq \langle \bar{x}, s\rangle - f_1(\bar{x}) \tag{2.3}$$

e

$$f_2^*(s) = \sup_x\{\langle \bar{x_2}, s\rangle - f_2(\bar{x_2})\} \geq \langle \bar{x}, s\rangle - f_2(\bar{x}). \tag{2.4}$$

Multiplicando (2.3) por α e (2.4) por $(1 - \alpha)$, temos que

$$\alpha f_1^*(s) \geq \alpha\langle \bar{x}, s\rangle - \alpha f_1(\bar{x})$$

e

$$(1 - \alpha) f_2^*(s) \geq (1 - \alpha)\langle \bar{x}, s\rangle - (1 - \alpha) f_2(\bar{x}).$$

Somando membro a membro obtemos

$$\alpha f_1^* + (1 - \alpha) f_2^* \geq \alpha(\langle \bar{x}, s\rangle - f_1(\bar{x})) + (1 - \alpha)(\langle \bar{x}, s\rangle - f_2(\bar{x}))$$

$$\begin{aligned}
\alpha f_1^* + (1-\alpha) f_2^* &\geq \langle \bar{x}, s\rangle - \underbrace{(\alpha f_1(\bar{x}) + (1-\alpha) f_2(\bar{x}))}_{g(\bar{x})} \\
\alpha f_1^* + (1-\alpha) f_2^* &\geq \langle \bar{x}, s\rangle - g(\bar{x}) \\
\langle \bar{x}, s\rangle - g(\bar{x}) &\leq \alpha f_1^* + (1-\alpha) f_2^* \\
g^*(s) = \sup_x \{\langle \bar{x}, s\rangle - g(\bar{x})\} = \langle \bar{x}, s\rangle - g(\bar{x}) &\leq \alpha f_1^* + (1-\alpha) f_2^*.
\end{aligned}$$

Como temos que $g(x) = \alpha f_1 + (1-\alpha) f_2$, assim

$$[\alpha f_1 + (1-\alpha) f_2]^* \leq \alpha f_1^* + (1-\alpha) f_2^*.$$

□

Proposição 2.10 *Considere $f : I\!R \to I\!R \cup \{\infty\}$ definida por $f(x) = ax^2 + bx + c$, com a, b e c constantes e $a > 0$. Então a sua conjugada é*

$$f^*(s) = \frac{(s-b)^2}{4a} - c.$$

Prova. Consideremos x e s satisfazendo (2.2), então $f'(x) = s$. Derivando a função $f(x) = ax^2 + bx + c$ ficamos com

$$f'(x) = 2ax + b,$$

ou seja,

$$f'(x) = 2ax + b = s \Leftrightarrow x = \frac{(s-b)}{2a}.$$

Assumindo que no ponto x a $f^*(s)$ atinge o seu supremo e substituindo a função f na expressão da conjugada, obtemos

$$\begin{aligned}
f^*(s) &= xs - f(x) \\
&= x(2ax + b) - (ax^2 + bx + c) \\
&= 2ax^2 + bx - 2ax^2 - bx - c \\
&= ax^2 - c.
\end{aligned}$$

Como $x = \dfrac{(s-b)}{2a}$, temos

$$f^*(s) = a\frac{(s-b)^2}{4a^2} - c = \frac{(s-b)^2}{4a} - c,$$

completando a demonstração. □

Proposição 2.11 *Considere $f : I\!R^n \to I\!R \cup \{\infty\}$ uma função semi-contínua inferiormente com o $dom(f)$ um conjunto compacto Q. Se $f^*(s) < \infty$, então*

$$f^*(s) = \max_x \{\langle x, s\rangle - f(x) \mid x \in Q\}.$$

Prova. Como $f(x) = \infty$ para $x \notin Q$, temos que

$$\begin{aligned} f^*(s) &= \sup_x \{\langle x, s\rangle - f(x) \mid x \in I\!R^n\} \\ &= \sup_x \{\langle x, s\rangle - f(x) \mid x \in Q\}. \end{aligned}$$

Pela semi-continuidade inferior de f e compacidade de Q, segue que

$$f^*(s) = \max_x \{\langle x, s\rangle - f(x) \mid x \in Q\},$$

completando a prova. □

A próxima seção estabelece uma expressão para a conjugada de uma função quadrática.

2.3 Aproximação quadrática da conjugada

Nesta seção discutimos uma relação entre a aproximação quadrática da função conjugada e a conjugada da aproximação quadrática de uma função convexa. Aqui consideraremos um função $f : I\!R \to I\!R \cup \{\infty\}$.

Enunciamos a seguir o teorema clássico da expansão de Taylor de 2^a ordem, que estabelece um modelo quadrático de uma função em torno de um ponto.

Teorema 2.12 *Sejam $f : I\!R \to I\!R \cup \{\infty\}$, derivável até 3^a ordem, e $x, x_0 \in I\!R$. Então existe x_1 no intervalo aberto de extremos x e x_0 tal que*

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2}(x - x_0)^2 + \frac{f'''(x_1)}{3!}(x - x_0)^3.$$

Prova. [7]. □

Neste teorema a parte associada a terceira derivada corresponde ao erro da aproximação.

Proposição 2.13 *Considere $f : I\!R \to I\!R \cup \{\infty\}$ uma função convexa de classe C^2, com sua conjugada de classe C^2 e $x_0 \in I\!R$. A conjugada de sua aproximação quadrática em torno de x_0 é dada por*

$$q^*(s) = \frac{(s - f'(x_0))^2}{2f''(x_0)} - f(x_0) + sx_0$$

e coincide com a aproximação quadrática da conjugada de f em torno de $s_0 = f'(x_0)$.

Prova. Primeiro calculemos a conjugada da aproximação quadrática. Seja $q(x)$ a aproximação quadrática em torno do ponto x_0. Fazendo $q(x) = h(x - x_0)$ e $y = x - x_0$, temos que

$$h(y) = f(x_0) + f'(x_0)y + \frac{1}{2}f''(x_0)y^2.$$

Como f é convexa, $f''(x_0) > 0$. Pela Proposição 2.10, a conjugada de h é:

$$h^*(s) = \frac{(s - f'(x_0))^2}{2f''(x_0)} - f(x_0).$$

Pela Proposição 2.6, obtemos

$$q^*(s) = \frac{(s - f'(x_0))^2}{2f''(x_0)} - f(x_0) + sx_0.$$

Agora veremos a aproximação quadrática da função conjugada em torno de $s_0 = f'(x_0)$. Por (2.2) temos que

$$f^*(s_0) = s_0 x_0 - f(x_0).$$

Note que, como x depende de s, temos que

$$(f^*)'(s) = x'(s)s + x(s) - f'(x(s))x'(s) = x \tag{2.5}$$

uma vez que $f'(x(s)) = s$. Derivando novamente,

$$(f^*)''(s) = x'(s). \tag{2.6}$$

Como $x(s_0) = x_0$ e $f'(x_0) = s_0$, temos de (2.5) que

$$(f^*)'(s_0) = x_0.$$

Por outro lado, ao derivarmos a expressão $f'(x) = s$ em relação a s obtemos

$$f''(x(s_0))x'(s_0) = 1.$$

Comparando com (2.6), temos que

$$(f^*)''(s_0) = \frac{1}{f''(x(s_0))}.$$

A aproximação quadrática de f^* em torno de s_0 é

$$\begin{aligned} q(s) &= (f^*)(s_0) + (f^*)'(s_0)(s - s_0) + \frac{1}{2}(f^*)''(s_0)(s - s_0)^2 \\ &= (x_0 s_0 - f(x_0)) + x_0(s - s_0) + \left(\frac{1}{2f''(x_0)}\right)(s - s_0)^2. \end{aligned}$$

Rearranjando a expressão, ficamos com

$$q(s) = x_0 s - f(x_0) + \frac{1}{2f''(x_0)}(s - s_0)^2.$$

Portanto, como $s_0 = f'(x_0)$, temos

$$q(s) = \frac{(s - f'(x_0))^2}{2f''(x_0)} - f(x_0) + x_0 s,$$

completando a demonstração. $\square$

Assim concluímos que se calcularmos a conjugada da aproximação quadrática ou calcularmos aproximação quadrática da conjugada obteremos o mesmo resultado.

$$
\begin{tikzcd}
 & q(x) \arrow[r] & q^*(s) \arrow[dr] & \\
f(x) \arrow[ur] \arrow[dr] & & & \text{resultado} \\
 & f^*(s) \arrow[r] & q(s) \arrow[ur] &
\end{tikzcd}
$$

Exemplo 2 *Considere $f(x) = e^x$. A aproximação quadrática desta função em torno de $x_0 = 1$ é dada por*

$$
\begin{aligned}
q(x) &= f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2}(x - x_0)^2 \\
&= e + e(x-1) + \frac{e}{2}(x-1)^2 \\
&= e + ex - e + \frac{e(x^2 - 2x + 1)}{2} \\
&= \frac{ex^2}{2} + \frac{e}{2}.
\end{aligned}
$$

Aplicando a Propriedade 2.10, obtemos

$$q^*(s) = \frac{s^2}{2e} - \frac{e}{2}.$$

Por outro lado, a conjugada de $f(x) = e^x$, calculada no Exemplo 1, é $f^*(s) = slns - s$. Para calcularmos a aproximação quadrática desta função devemos escolher s_0 de maneira que $s_0 = f'(x_0)$, ou seja, $s_0 = e$. Logo a aproximação quadrática da conjugada é

$$
\begin{aligned}
q(s) &= f^*(s_0) + (f^*)'(s_0)(s - s_0) + \frac{(f^*)''(s_0)}{2}(s - s_0)^2 \\
&= (elne - e) + (lne)(s - e) + \frac{1}{2e}(s - e)^2 \\
&= (s - e) + \frac{s^2 - 2es + e^2}{2e} \\
&= \frac{s^2}{2e} - \frac{e}{2}.
\end{aligned}
$$

Assim podemos observar que $q^*(s)$ e $q(s)$ são iguais.

Na Figura 2.7 ilustramos a conjugada da aproximação quadrática em um ponto e na Figura 2.8 a aproximação quadráticada conjugada.

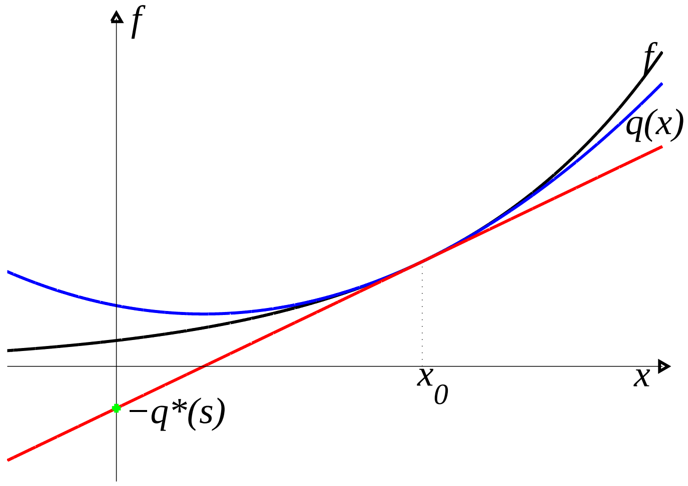

Figura 2.7: Conjugada da aproximação quadrática.

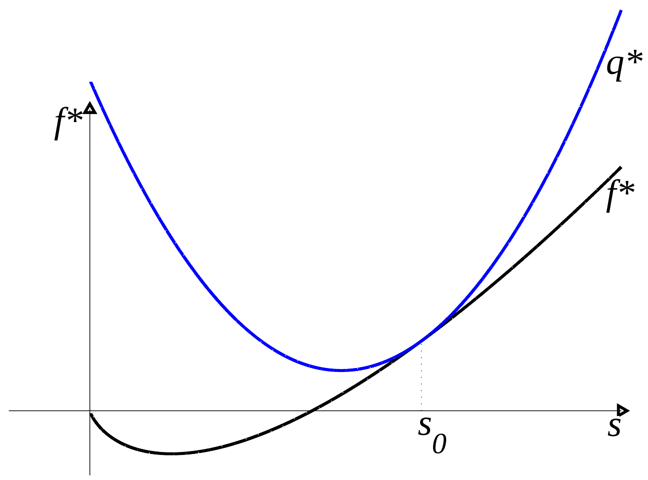

Figura 2.8: Aproximação quadrática da conjugada.

Capítulo 3

O Problema de Programação Não Linear

Neste capítulo apresentamos o problema geral de programação não linear diferenciável e suas condições de otimalidade e de qualificação. Veremos mais profundamente as condições de otimalidade de Karush-Kuhn-Tucker, mas exitem outras como as de Fritz-John. Provaremos o teorema de KKT usando uma condição fraca, mas difícil de ser verificada, que é a igualdade, no ponto ótimo, do polar do cone viável linearizado e do polar cone tangente. Veremos também algumas condições de qualificação como Slater, independência linear dos gradientes, Mangasarian-Fromovitz, posto constante, dependência linear positiva, quase-normalidade e quase-regularidade e as implicações entre elas.

Concentraremos nossa atenção no seguinte problema:

$$\begin{array}{cl} \text{minimizar} & f(x) \\ \text{sujeito a} & h(x) = 0 \\ & g(x) \leq 0 \end{array} \tag{3.1}$$

com $f : I\!R^n \to I\!R$, $h : I\!R^n \to I\!R^m$ e $g : I\!R^n \to I\!R^p$ continuamente diferenciáveis.

O conjunto

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

é o conjunto ou região viável.

Definição 3.1 *Uma restrição de desigualdade $g_i(x) \leq 0$ é dita ativa num ponto $\bar{x}$, se $g_i(\bar{x}) = 0$, e inativa em $\bar{x}$, se $g_i(\bar{x}) < 0$.*

Denotaremos o conjunto de índices das retrições de desigualdade ativas por $A(\bar{x})$, isto é,

$$A(\bar{x}) = \{i \mid g_i(\bar{x}) = 0\}.$$

Existem dois tipos de minimizadores, o local e o global. Todo minimizador global é local, mas nem todo minimizador local é global.

Definição 3.2 *Um ponto $x^* \in \Omega$ é minimizador global de f, se e somente se, $f(x^*) \leq f(x)$ para todo x em Ω.*

Definição 3.3 *Um ponto $x^* \in \Omega$ é dito minimizador local de f, se e somente se, existe uma vizinhança V de x^* tal que $f(x^*) \leq f(x)$ para todo x em $\Omega \cap V$.*

3.1 Cones importantes

Apresentamos a seguir alguns cones importantes que estão relacionadas a uma linearização do conjunto viável. Veremos também algumas relações entres estes cones. Indicaremos por $\nabla h(\bar{x})$ a transposta da jacobiana de h no ponto $\bar{x}$. As demonstrações dos Lemas e Teoremas desta seção podem ser encontradas em [2, 5, 13].

Definição 3.4 *Dado um ponto $\bar{x} \in \Omega$ e o conjunto $A(\bar{x})$, definimos o cone viável linearizado de Ω a partir de $\bar{x}$, denotado por $D(\bar{x})$, como*

$$D(\bar{x}) = \left\{ d \in I\!R^n \mid \nabla h(\bar{x})^t d = 0 \quad e \quad \nabla g_j(\bar{x})^t d \leq 0, \forall j \in A(\bar{x}) \right\}.$$

Segue da definição que $d \in D(\bar{x})$ é tangente às restrições de igualdade e d faz ângulo reto ou obtuso com os gradientes das restrições de desigualdade ativas. Claramente $D(\bar{x})$ é um cone não vazio, pois $0 \in D(\bar{x})$. Além disso, $D(\bar{x})$ é convexo e fechado. Na Figura 3.1 temos o conjunto viável e o cone viável linearizado transladado no ponto x^*.

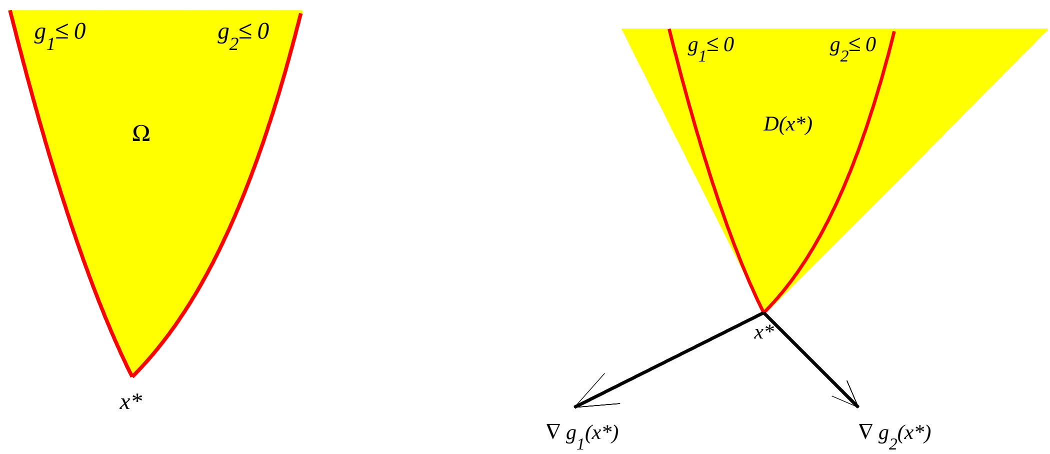

Figura 3.1: O conjunto viável e o cone viável linearizado transladado.

Considere $\bar{x} \in \Omega$ e o conjunto definido por

$$G(\bar{x}) = \left\{ \sum_{i=1}^{m} \lambda_i \nabla h_i(\bar{x}) + \sum_{j \in A(\bar{x})} \mu_j \nabla g_j(\bar{x}) \mid \mu_j \geq 0, \ \forall j \in A(\bar{x}) \right\}. \qquad (3.2)$$

Existe uma relação entre $D(\bar{x})$ e $G(\bar{x})$ antes, porém, veremos uma propriedade de $G(\bar{x})$.

Lema 3.5 *Seja $G(\bar{x})$ definido pela Equação* (3.2). *Então $G(\bar{x})$ é um cone convexo fechado.*

Lema 3.6 *Dado $\bar{x} \in \Omega$, temos que $D(\bar{x}) = P(G(\bar{x}))$.*

Antes de ver a definição de cone tangente precisamos da definição de direção tangente.

Definição 3.7 *Uma direção $d \in I\!R^n$ é tangente a $\Omega \subset I\!R^n$ a partir de $\bar{x} \in \Omega$, se e somente se, existe uma seqüência de pontos viáveis $(x^k) \in \Omega$ tal que $x^k \to \bar{x}$ e*

$$\frac{x^k - \bar{x}}{\|x^k - \bar{x}\|} \to \frac{d}{\|d\|}.$$

Podemos escrever o cone tangente, a partir da definição de direção tangente, como

$$T(\bar{x}) = \{0\} \cup \left\{ d \in I\!R^n \text{ com } d \neq 0 \mid \exists (x^k) \subset \Omega,\ x^k \to \bar{x} \quad \text{tal que} \quad \frac{x^k - \bar{x}}{\|x^k - \bar{x}\|} \to \frac{d}{\|d\|} \right\}.$$

Ao contrário do cone viável linearizado, o cone tangente não é necessariamente convexo. O cone tangente, assim como o cone viável linearizado, é uma aproximação do conjunto viável em torno do ponto dado. Este cone será útil para fornecer algumas relações entre condições de otimalidade e de qualificação. No sentido intuitivo, dizemos que d "penetra" em Ω ou o "tangencia". Outra forma de definir o cone tangente é a de Bouligand. Estas duas formas de definir cone tangente são equivalentes.

Teorema 3.8 *O cone tangente possui a seguinte forma equivalente*

$$T(\bar{x}) = \left\{ d \in I\!R^n \;\middle|\; \begin{array}{c} \exists (t_k) \subset I\!R_+,\ (t_k) \to 0_+, \\ \exists (d^k) \subset I\!R^n,\ (d^k) \to d, \textit{tais que} \\ \bar{x} + t_k d^k \ \in \ \Omega \,\textit{para todo}\ k \in I\!N \end{array} \right\}$$

Prova. [2] □

Podemos ver uma propriedade sobre o cone tangente enunciada no próximo lema.

Lema 3.9 *Considere $\Omega \in I\!R^n$ e $\bar{x} \in \Omega$. Então $T(\bar{x})$ é um cone fechado.*

As relações entre os cones tangente e viável linearizado é dada abaixo.

Lema 3.10 *Considere $\bar{x} \in \Omega$. Então $T(\bar{x}) \subset D(\bar{x})$.*

Pelos Lemas 3.10 e 1.16, temos que $P(D(\bar{x})) \subset P(T(\bar{x}))$. Se $P(D(\bar{x})) = P(T(\bar{x}))$ e $T(\bar{x})$ é convexo, pelo Lema de Farkas 1.19, temos que $D(\bar{x}) = T(\bar{x})$.

Outro resultado importante é que o valor da função objetivo aumenta ao longo de qualquer direção tangente a um ponto $\bar{x} \in \Omega$.

Teorema 3.11 *Se x^* é um minimizador local do Problema* (3.1)*, então*

$$\nabla f(x^*)^t d \geq 0, \quad \forall d \in T(x^*).$$

3.2 Condições de otimalidade

Consideraremos agora as condições que devem ser satisfeitas quando um ponto $x^* \in \mathbb{R}^n$ é um minimizador local do problema (3.1). Estas condições são chamadas condições necessárias de otimalidade.

Inicialmente veremos uma condição necessária de otimalidade para um problema irrestrito, isto é, quando $\Omega = \mathbb{R}^n$.

Teorema 3.12 *Seja $f : \mathbb{R}^n \to \mathbb{R}$, $f \in C^1$. Se x^* é um minimizador local de f em $\mathbb{R}^n$, então*

$$\nabla f(x^*) = 0.$$

Prova. [4] □

Reconhecer que um minimizador local é também um minimizador global é uma tarefa difícil, e a função objetivo deve ter características especiais. O caso mais simples é quando a função objetivo é convexa.

Teorema 3.13 (Teorema de minimização convexa) *Sejam $\Omega \subset \mathbb{R}^n$ um conjunto convexo e $f : \Omega \to \mathbb{R}$ uma função convexa em Ω. Então:*

(i) Todo minimizador local do problema é global e, além disso, o conjunto dos minimizadores é convexo;

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

Prova. [13] □

Inicialmente enunciaremos a condição de otimalidade de Karush-Kuhn-Tucker (KKT) com uma hipótese fraca, mas difícil de ser verificada, $P(D(\bar{x})) = P(T(\bar{x}))$. Tal hipótese, dita condição de qualificação, será discutida mais adiante.

Teorema 3.14 (Condições de KKT) *Considere $x^* \in \Omega$ um minimizador local do Problema (3.1). Se $P(T(x^*)) = P(D(x^*))$, então existem $\lambda^* \in \mathbb{R}^m$ e $\mu^* \in \mathbb{R}^p$ tais que:*

$$-\nabla f(x^*) = \sum_{i=1}^{m} \lambda_i^* \nabla h_i(x^*) + \sum_{i=1}^{p} \mu_i^* \nabla g_i(x^*),$$

$$\begin{array}{lcll} \mu_i^* & \geq & 0, & i = 1, \ldots, p, \\ \mu_i^* g_i(x^*) & = & 0, & i = 1, \ldots, p, \\ h_i(x^*) & = & 0, & i = 1, \ldots, m, \\ g_i(x^*) & \leq & 0, & i = 1, \ldots, p. \end{array}$$

Prova. Considere x^* um minimizador local do Problema (3.1).Temos, pelo Teorema 3.11, que

$$-\nabla f(x^*)^t d \leq 0, \quad \forall d \in T(x^*).$$

Pela Definição 1.15 de cone polar e pela hipótese de que $P(T(x^*)) = P(D(x^*))$, obtemos

$$-\nabla f(x^*) \in P(T(x^*)) = P(D(x^*)).$$

Como o Lema 3.6 nos diz que $D(x^*) = P(G(x^*))$, assim vemos

$$-\nabla f(x^*) \in P(P(G(x^*))).$$

Pelo Lema 3.5 sabemos que $G(x^*)$ é um cone convexo fechado. Então pelo Lema de Farkas 1.19 obtemos

$$-\nabla f(x^*) \in G(x^*),$$

ou seja, existem λ_i com $i = 1, \ldots, m$ e μ_j com $j \in A(x^*)$, tais que

$$-\nabla f(x^*) = \sum_{i=1}^{m} \lambda_i \nabla h_i(x^*) + \sum_{j \in A(x^*)} \mu_j \nabla g_j(x^*)$$

com $\mu_j \geq 0$ para todo $j \in A(x^*)$. Como o número de elementos de $(A(x^*))$ é menor do que ou igual a p, para encontrar λ^* e μ^* basta completar o vetor μ^* com zeros tanto quantos forem necessários, isto é, basta definir $\lambda_i^* = \lambda_i$ para todo $i = 1, \ldots, m$ e

$$\mu_j^* = \begin{cases} \mu_j & \forall j \in A(x^*) \\ 0 & \text{caso contrário.} \end{cases}$$

$\square$

A condição $P(T(x^*)) = P(D(x^*))$ foi introduzida por Guignard em [9] e é a mais fraca possível, o que foi provado em 1971 por Gould e Tolle, vide [6].

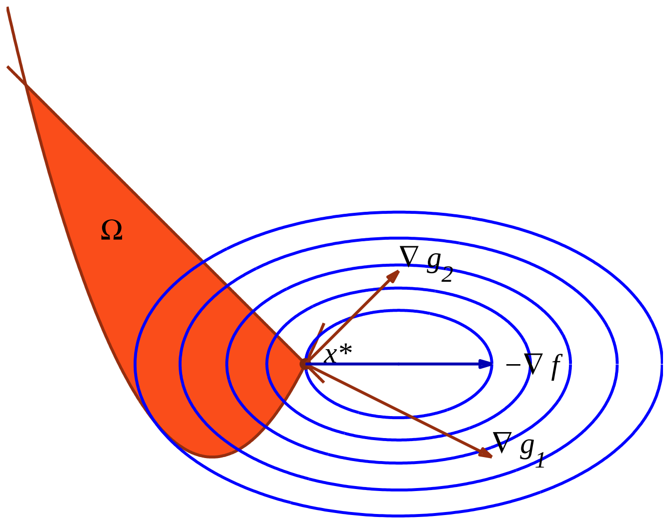

Figura 3.2: Condições de Karush-Kuhn-Tucker

A Figura 3.2 ilustra a situação em que há somente restrições de desigualdade. Neste caso, as condições de KKT significam que o oposto do gradiente da função objetivo pertence ao cone gerado pelos gradientes das restrições ativas.

3.3 Condições de qualificação

Condições de qualificação são ferramentas úteis na análise de convergência de métodos de otimização. A grosso modo, uma condição de qualificação pode ser entendida como uma propriedade que garante que as condições de KKT sejam válidas num minimizador local.

Definição 3.15 *Dizemos que as restrições $h(x) = 0$ e $g(x) \leq 0$ cumprem uma condição de qualificação em $x^* \in \Omega$ quando, dada qualquer função diferenciável que tenha mínimo em x^*, relativamente a Ω, sejam satisfeitas as condições de otimalidade de KKT.*

Observamos anteriormente que vale KKT em um minimizador x^*, quando $P(T(x^*)) = P(D(x^*))$. Pela definição acima, esta condição é de qualificação.

Antes de expormos as diferentes condições de qualificação veremos uma importante definição.

Definição 3.16 *Considere $A = \{a_1, \ldots, a_l\}$ e $B = \{b_1, \ldots, b_r\}$ dois subconjuntos finitos em $I\!R^n$ tais que $A \cup B \neq \emptyset$. Dizemos que (A, B) é positivo-linearmente dependente (PLD) se existem $\alpha \in I\!R^l$ e $\beta \in I\!R^r$ tais que $\beta \geq 0$, $(\alpha, \beta) \neq 0$, e*

$$\sum_{i=1}^{l} \alpha_i a_i + \sum_{j=1}^{r} \beta_j b_j = 0.$$

Se (A, B) não é PLD, dizemos então que é positivo-linearmente independente (PLI).

A seguir exporemos as diferentes condições de qualificações estudadas, porém, aqui, não demonstraremos as implicações entre estas condições. Em [2] há um tratamento do assunto muito amplo. Teremos sempre em mente que $\bar{x} \in \Omega$.

Definição 3.17 (Condição de Qualificação de Slater) *Dizemos que a condição de qualificação de Slater vale se h é afim, g é convexa e existe $\tilde{x} \in \Omega$ tal que*

$$h(\tilde{x}) = 0 \quad e \quad g(\tilde{x}) < 0.$$

Definição 3.18 (Condição de Qualificação de Independência Linear - LICQ) *Dizemos que a condição de qualificação de indepedência linear (LICQ) é satisfeita em $\bar{x}$ quando o conjunto dos gradientes das restrições de desigualdade ativas e das restrições de igualdade são linearmente independentes, isto é,*

$$\{\nabla h_i(\bar{x})\}_{i \in \{1,\ldots,m\}} \cup \{\nabla g_j(\bar{x})\}_{j \in A(\bar{x})} \text{ é } LI.$$

Definição 3.19 (Condição de Qualificação de Mangasarian-Fromovitz-MFCQ) *Dizemos que a condição de qualificação de Mangasarian-Fromovitz (MFCQ) é satisfeita em $\bar{x}$ quando os gradientes das restrições de igualdade são linearmente independentes e existir um vetor $d \in I\!R^n$ tal que:*

$$\begin{aligned} \nabla h(\overline{x})^T d &= 0 \quad e \\ \nabla g_j(\overline{x})^T d &< 0, \text{ para todo } j \in A(\overline{x}). \end{aligned}$$

Definição 3.20 (Condição de Qualificação Posto Constante - CRCQ) *Dizemos que a condição de qualificação de posto constante (CRCQ) é satisfeita em $\bar{x}$ se existe uma vizinhança V de $\bar{x}$ tal que para todo $I \subset \{1, \ldots, m\}$, $J \subset A(\bar{x})$, o conjunto de vetores gradientes*

$$\{\nabla h_i(y)\}_{i \in I} \cup \{\nabla g_j(y)\}_{j \in J}$$

tem posto constante para todo $y \in V$.

Definição 3.21 (Condição de Qualificação Dependência Linear Positiva Constante - CPLD) *Dizemos que a condição de qualificação CPLD é satisfeita em $\bar{x}$ se para todo $I_0 \subset \{1, \ldots, m\}$, $J_0 \subset A(\bar{x})$ com os gradientes*

$$(\{\nabla h_i(\bar{x})\}_{i \in I_0}, \{\nabla g_j(\bar{x})\}_{j \in J_0})$$

positivo-linearmente dependentes, existir uma vizinhança V de $\bar{x}$ tal que os gradientes

$$(\{\nabla h_i(y)\}_{i \in I_0}, \{\nabla g_j(y)\}_{j \in J_0})$$

sejam positivo-linearmente dependentes para todo $y \in V$.

Em [20] a Condição de Qualificação Dependência Linear Positiva Constante foi usada para provar a convergência de um método de lagrangiano aumentado.

Definição 3.22 (Condição de Qualificação Quase-Normalidade) *Dizemos que a condição de qualificação quase-normalidade é satisfeita em $\bar{x}$ se não existem escalares λ_i com $i \in \{1, \ldots, m\}$ e μ_j com $j \in A(\bar{x})$ tais que:*

(i)

$$\sum_{i=1}^{m} \lambda_i \nabla h_i(\bar{x}) + \sum_{j \in A(\bar{x})} \mu_j \nabla g_j(\bar{x}) = 0;$$

(ii) $\mu_j \geq 0$;

(iii) λ_i *e* μ_j *são não todos nulos;*

(iv) Em toda vizinhança V de $\bar{x}$ existe um ponto $x \in V$ tal que $\lambda_i h_i(x) > 0$ para todo i com $\lambda_i \neq 0$ e $\mu_j g_j(x) > 0$ para todo j com $\mu_j \neq 0$.

Definição 3.23 (Condição de Qualificação Quase-Regularidade) *Dizemos que a condição de qualificação quase-regularidade é satisfeita em $\bar{x}$ quando ocorre a igualdade $T(\bar{x}) = D(\bar{x})$.*

Definição 3.24 (Condição de Qualificação de Guignard) *Dizemos que a condição de qualificação de Guignard é satisfeita em $\bar{x}$ quando ocorre a igualdade $P(T(\bar{x})) = P(D(\bar{x}))$.*

Segue abaixo um gráfico ilustrando as relações de implicação entre as condições de qualificação.

$$\begin{tikzcd}
\boxed{\text{Slater}} \arrow[dr, Rightarrow] & & \boxed{\text{LICQ}} \arrow[dl, Rightarrow] \arrow[dr, Rightarrow] & \\
& \boxed{\text{MFCQ}} \arrow[dr, Rightarrow] & & \boxed{\text{CRCQ}} \arrow[dl, Rightarrow] \\
& & \boxed{\text{CPLD}} \arrow[d, Rightarrow] & \\
& & \boxed{\text{Quase-normalidade}} \arrow[d, Rightarrow] & \\
& & \boxed{\text{Quase-regularidade}} \arrow[d, Rightarrow] & \\
& & \boxed{\text{P(T)=P(D)}} &
\end{tikzcd}$$

O contra exemplo a seguir foi retirado de [2].

Contra-exemplo 1 *A condição* $P(T(\bar{x})) = P(D(\bar{x}))$ *não implica na condição quase-regularidade.*

Considere as funções $h_1 : I\!R^2 \to I\!R$, $g_i : I\!R^2 \to I\!R$ $(i = 1, 2)$ definidas por

$$\begin{aligned} h_1(x) &= x_1 x_2 \\ g_1(x) &= -x_1 \\ g_2(x) &= -x_2, \end{aligned}$$

o conjunto viável $\Omega = \{x \in I\!R^2 \mid h_1(x) = 0, g_i(x) \leq 0 \quad i = 1, 2\}$ e o ponto viável $\bar{x} = (0, 0)^T$.

Temos que

$$\nabla h_1(\bar{x}) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \nabla g_1(\bar{x}) = \begin{pmatrix} -1 \\ 0 \end{pmatrix} \quad \text{e} \quad \nabla g_2(\bar{x}) = \begin{pmatrix} 0 \\ -1 \end{pmatrix}.$$

Vemos facilmente que $D(\bar{x}) = \{(d_1, d_2) \mid d_1 \geq 0, d_2 \geq 0\}$. Para obter $T(\bar{x})$, note inicialmente que $(1, 0)^T$ e $(0, 1)^T$ pertencem a $T(\bar{x})$. Considere agora $d = (d_1, d_2)$ uma direção tangente não nula arbitrária. Temos $d_1 \geq 0$ e $d_2 \geq 0$, pois pelo

Lema 3.10, sabemos que $T(\bar{x}) \subset D(\bar{x})$. Seja $x^k = (x_1^k, x_2^k)$ uma seqüência viável convergindo para $\bar{x}$, tal que $\dfrac{x^k - \bar{x}}{\|x^k - \bar{x}\|} \to \dfrac{d}{\|d\|}$. Desse modo temos

$$\frac{x_1^k}{\|x^k\|} \to \frac{d_1}{\|d\|} \quad \text{e} \quad \frac{x_2^k}{\|x^k\|} \to \frac{d_2}{\|d\|}. \tag{3.3}$$

Observe que $d_1 = 0$ ou $d_2 = 0$, pois do contrário, isto é, se fosse $d_1 > 0$ e $d_2 > 0$, a relação (3.3) implicaria $x_1^k > 0$ e $x_2^k > 0$ para k suficientemente grande, contradizendo o fato de que $x_1^k x_2^k = 0$. Desta forma, segue que $T(\bar{x}) = \{(d_1, d_2) \mid d_1 \geq 0, d_2 \geq 0, d_1 d_2 = 0\}$, concluíndo assim que $T(\bar{x}) \neq D(\bar{x})$.

Para verificar a igualdade entre o polar de $T(\bar{x})$ e o polar de $D(\bar{x})$, basta utilizar a definição de polar, obtendo assim $P(T(\bar{x})) = P(D(\bar{x})) = \{(d_1, d_2) \mid d_1 \leq 0, d_2 \leq 0\}$. Veja Figura 3.3.

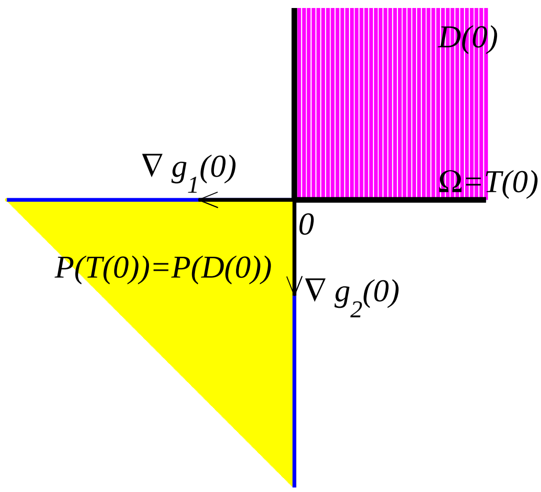

Figura 3.3: Contra-exemplo 1.

Capítulo 4

Aplicações na Economia

Neste capítulo procuramos aplicar na Economia os conceitos estudados nos capítulos anteriores de função conjugada e das condições de otimalidade de Karush-Kuhn-Tucker (KKT).

Apresentaremos alguns modelos relacionados à Teoria da Firma que recaem em problemas de programação não linear [11, 15]. Veremos três diferentes modelos da teoria da Firma: a concorrência perfeita, o monopólio e o oligopólio. Discutiremos então as condições necessárias de otimalidade destes modelos.

Estaremos interessados em quanto uma organização deve produzir e qual o preço a ser praticado. Queremos então uma relação entre os insumos de entrada, a quantidade a ser produzida e o preço do produto.

Assumiremos que a organização produz um único produto a partir de diversos insumos. A quantidade de insumo i usado na fabricação do produto denotaremos por x_i. O *espaço de entrada*, ou espaço de entrada dos insumos é denotado por

$$I = \{x = (x_1, \ldots, x_n)^t \mid x_i \geq 0\}.$$

A relação entre a quantidade de insumos e a produção que é realizada com estes insumos é dada pela *função de produção*

$$f(x) = f(x_1, \ldots, x_n) = q$$

onde q é a quantidade produzida. Assumiremos sempre que a função de produção é continuamente diferenciável.

A função de produção satisfaz dois axiomas.

Axioma 1. O primeiro é que existe um subconjunto do espaço de entrada I, chamado de *região econômica*, em que cada acréscimo nos insumos a produção não diminui. Assim, se x^1 e x^2 são dois pontos desta região

$$x^1 \geq x^2 \Rightarrow f(x^1) \geq f(x^2).$$

Logo a derivada parcial da função de produção, chamada de *produção marginal*, é não negativa, ou seja,

$$\frac{\partial f}{\partial x_i}(x) = MP_i(x) \geq 0.$$

Definindo o vetor produção marginal como:

$$MP(x) = \frac{\partial f}{\partial x}(x) = (MP_1(x), \ldots, MP_n(x)), \tag{4.1}$$

a região econômica é o subconjunto do espaço de entrada

$$\{x \in I \mid MP(x) \geq 0\}.$$

Axioma 2. O segundo axioma diz que existe um subconjunto convexo da região econômica, em que a matriz Hessiana da função de produção

$$H = H(x) = \frac{\partial^2 f}{\partial x^2}(x) = \begin{pmatrix} \dfrac{\partial^2 f}{\partial x_1^2}(x) & \cdots & \dfrac{\partial^2 f}{\partial x_1 \partial x_n}(x) \\ \vdots & \ddots & \vdots \\ \dfrac{\partial^2 f}{\partial x_n \partial x_1}(x) & \cdots & \dfrac{\partial^2 f}{\partial x_n^2}(x) \end{pmatrix}$$

é negativa definida.

De acordo com estes dois axiomas existe uma região convexa do espaço de entrada chamada *região relevante* R, definida por:

$$R = \{x \in I \mid MP(x) \geq 0 \; e \; H(x) \text{ negativa definida}\}. \tag{4.2}$$

4.1 Concorrência perfeita

A estrutura de mercado definida como concorrência perfeita tem algumas hipóteses básicas:

(i) grande número de empresas;

(ii) produto homogêneo;

(iii) livre entrada e saída de empresas;

(iv) maximização de lucros;

(v) livre circulação de informação.

A teoria neoclássica da firma diz que o objetivo da organização é maximizar o lucro pela escolha dos insumos.

O lucro Π é a diferença entre a receita total RT e o custo total de produção CT:

$$\Pi = RT - CT,$$

onde a receita total é o produto do preço do produto p pela quantidade produzida q:

$$RT = pq = pf(x),$$

e o custo total é o somatório do produto dos preços dos insumos pela quantidade de insumos:

$$CT = \sum_{i=1}^{n} w_i x_i = w^t x.$$

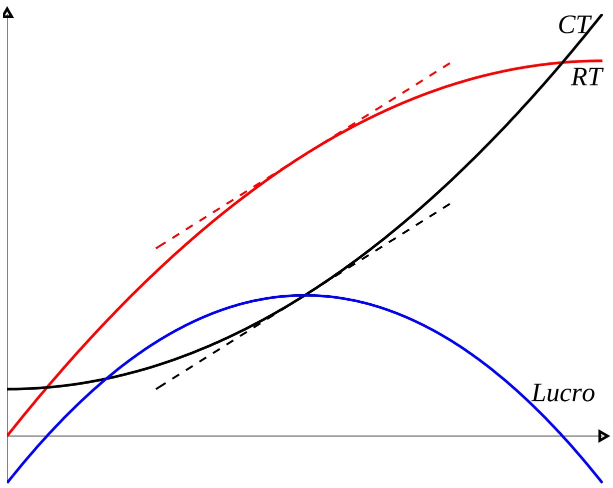

Figura 4.1: Gráfico do lucro Π, receita total RT e custo total CT.

Na Figura 4.1 vemos exemplos de curvas da Receita Total, do Custo Total e do Lucro.

Logo o problema da firma de longo prazo é

$$\begin{array}{ll} \max\limits_{x} & \Pi(x) = pf(x) - w^t x \\ \text{s. a} & x \geq 0. \end{array} \tag{4.3}$$

Porém (4.3) não está de acordo com o problema proposto (3.1), logo temos que modificá-lo.

$$\begin{array}{ll} \min\limits_{x} & -\Pi(x) = -pf(x) + w^t x \\ \text{s. a} & -x \leq 0 \end{array}$$

Aplicando o Teorema 3.14 de KKT , temos que se $x \geq 0$ é solução de (4.3), então existe $\mu \in I\!R^n$ tal que

$$p\frac{\partial f}{\partial x}(x) - w = -\mu$$

ou equivalentemente,

$$p\frac{\partial f}{\partial x}(x) - w + \mu = 0.$$

Como $\mu \geq 0$, temos que

$$p\frac{\partial f}{\partial x}(x) - w \leq 0. \tag{4.4}$$

Por outro lado, KKT ainda afirma que

$$\mu g(x) = 0,$$

e como $\mu = -\left(p\frac{\partial f}{\partial x}(x) - w\right)$ e $g(x) = -x$, obtemos

$$\left(p\frac{\partial f}{\partial x}(x) - w\right)x = 0.$$

De (4.1) e de (4.4) temos para $i = 1, \ldots, n$, que

$$p\frac{\partial f}{\partial x_i}(x) = pMP_i(x) \leq w_i.$$

Logo

$$\begin{cases} pMP_i(x) = w_i & \text{se} \quad x_i > 0 \\ x_i = 0 & \text{se} \quad pMP_i(x) < w_i \end{cases}$$

onde $pMP_i(x)$ é o valor da produção marginal no ponto x. Assumindo que todos os insumos são realmente usados, $x > 0$ e as condições de primeira ordem são

$$p\frac{\partial f}{\partial x_i}(x) = pMP_i(x) = w_i.$$

As n condições de primeira ordem

$$\psi_i(x) = p\frac{\partial f}{\partial x_i}(x) - w_i = 0, \ i = 1, \ldots, n$$

podem ser resolvidas para encontrar as entradas ótimas, se a Jacobiana da matriz

$$J = \begin{pmatrix} \frac{\partial \psi_1}{\partial x_1}(x) & \cdots & \frac{\partial \psi_1}{\partial x_n}(x) \\ \vdots & \ddots & \vdots \\ \frac{\partial \psi_n}{\partial x_1}(x) & \cdots & \frac{\partial \psi_n}{\partial x_n}(x) \end{pmatrix} = pH$$

é não singular. Assumindo que o vetor de entrada x está na região relevante (4.2), a Jacobiana da matriz é não singular e os valores ótimos das entradas podem ser obtidos. Logo os valores ótimos das entradas são *funções de demanda de entrada*:

$$x^* = x^*(p, w).$$

4.2 Monopólio

O monopólio é a estrutura em que há apenas um produtor. As hipóteses básicas do modelo de monopólio são:

(i) um único produtor;

(ii) produto sem substitutos próximos;

(iii) barreiras à entrada;

(iv) maximização de lucros.

Uma das principais diferenças entre concorrência perfeita e o monopólio é que no monopólio o produtor pode variar o preço do produto produzido variando a quantidade produzida, ou seja, $p = p(q)$. Geralmente, quanto mais o monopolista vende mais barato o preço fica, logo

$$\frac{dp}{dq} < 0.$$

A receita total é definida como

$$RT(q) = p(q)q$$

e a *receita marginal* é

$$MR(q) = \frac{dRT}{dq} = p + \frac{dp}{dq}q.$$

Outra diferença entre o monopólio e a concorrência é que o monopolista pode influenciar o preço dos insumos, $w_i = w_i(x_i),\ i = 1, \ldots, n$. E, em geral, o preço aumenta quanto mais é comprado,

$$\frac{\partial w_i}{\partial x_i} > 0,\ i = 1, \ldots, n.$$

Como o custo total do insumo i é

$$CT_i(x_i) = w_i(x_i)x_i,$$

o *custo marginal* do insumo i é

$$MC_i(x_i) = \frac{dCT_i}{dx_i}(x_i) = w_i + \frac{\partial w_i}{\partial x_i}x_i.$$

Assim o problema da firma para monopólio é

$$\begin{array}{ll} \max\limits_{q,x} & \Pi = p(q)q - \sum\limits_{i=1}^{n} w_i(x_i)x_i \\ \text{s. a} & q = f(x_1, \ldots, x_n) \end{array} \tag{4.5}$$

mudando (4.5) para a forma (3.1)

$$\begin{array}{ll} \min\limits_{q,x} & -\Pi = -p(q)q + \sum\limits_{i=1}^{n} w_i(x_i)x_i \\ \text{s. a} & q - f(x_1, \ldots, x_n) = 0. \end{array}$$

Aplicando KKT temos que se $x \geq 0$ é solução de (4.5), então existe $\lambda \in I\!R^n$ tal que

$$\nabla\Pi = \lambda\nabla h$$

$$\nabla\Pi - \lambda\nabla h = 0$$

$$\left(\begin{array}{c} p + \dfrac{dp}{dq}q - \lambda \\ -w_i - \dfrac{\partial w_i}{\partial x_i}x_i + \lambda\dfrac{\partial f}{\partial x_i}(x) \end{array}\right) = \left(\begin{array}{c} 0 \\ 0 \end{array}\right)$$

com $i = 1, \ldots, n$.
Então temos que as condições necessárias são

$$p + \frac{dp}{dq}q = \lambda$$

$$w_i + \frac{\partial w_i}{\partial x_i}x_i = \lambda\frac{\partial f}{\partial x_i}(x),\ i = 1, \ldots, n.$$

Entretanto

$$MR(q) = p + \frac{dp}{dq}q,$$

$$MC_i(x_i) = w_i + \frac{\partial w_i}{\partial x_i} x_i$$

e

$$\frac{\partial f}{\partial x_i}(x) = MP_i(x).$$

Assim

$$\lambda = MR(q)$$

e

$$MC_i(x_i) = \lambda MP_i(x), \ i = 1, \ldots, n.$$

$$MC_i(x_i) = MR(q) MP_i(x), \ i = 1, \ldots, n. \tag{4.6}$$

Definindo o custo marginal como

$$MC(q) = \frac{dCT}{dq}(q),$$

ou seja,

$$MC(f(x_1, \ldots, x_n))$$

e aplicando o Teorema das Funções Implícitas, vide [8], obtemos

$$MC(f(x_1, \ldots, x_n)) = \frac{MC_i(x_i)}{MP_i(x)}. \tag{4.7}$$

De (4.6) e (4.7) temos que a receita marginal tem que ser igual ao custo marginal

$$MR = MC. \tag{4.8}$$

Esta condição é a condição necessária de otimalidade. Sabemos que a derivada em um ponto é a inclinação da reta tangente naquele ponto, logo (4.8) nos diz que inclinação da reta tangente do custo total e da receita total tem que ser iguais no ponto ótimo. Isso pode ser verificado na Figura 4.1.

4.3 Oligopólio

Na estrutura do mercado em que há poucas organizações competindo em um mesmo segmento do mercado existe o oligopólio. Para escolher políticas ótimas para as organizações, cada uma delas deve escolher políticas que não só afetam à elas próprias, mas a seus pares.

Aqui nos deteremos no caso em que existem somente duas organizações, entretanto para mais organizações é análogo. A função de produção de cada uma das organizações descreveremos por

$$q^1 = f^1(x_1^1, x_2^1, \ldots, x_n^1)$$

$$q^2 = f^2(x_1^2, x_2^2, \ldots, x_n^2).$$

É evidente que o preço das mercadorias produzidas depende não só da quantidade que uma firma irá produzir e sim da quantidade que todas irão produzir, então neste caso

$$p = p(q^1, q^2).$$

Assim como no monopólio o preço das mercadorias produzidas tende a decrescer quanto mais mercadorias são produzidas

$$\frac{\partial p}{\partial q^1} < 0$$

e

$$\frac{\partial p}{\partial q^2} < 0.$$

O preço dos insumos denotaremos por

$$w_i = w_i(x_i^1, x_i^2) \quad i = 1, \dots, n$$

e também irão depender das compras de ambas as organizações, e estes preços tendem a crescer quanto mais insumos são comprados

$$\frac{\partial w_i}{\partial x_i^1} > 0 \quad i = 1, \dots, n$$

e

$$\frac{\partial w_i}{\partial x_i^2} > 0 \quad i = 1, \dots, n.$$

Consideraremos a primeira firma como a organização líder e que influência nas quantidades que a segunda firma irá produzir, assim o problema para a primeira firma é

$$\begin{array}{ll} \max\limits_{q^1, x} & \Pi = p(q^1, q^2)q^1 - \sum\limits_{i=1}^{n} w_i(x_i^1, x_i^2)x_i^1 \\ \text{s. a} & q^1 = f^1(x_1^1, \dots, x_n^1) \end{array} \tag{4.9}$$

arrumando (4.9) para a forma padrão

$$\begin{array}{ll} \min\limits_{q^1, x} & -\Pi = -p(q^1, q^2)q^1 + \sum\limits_{i=1}^{n} w_i(x_i^1, x_i^2)x_i^1 \\ \text{s. a} & f^1(x_1^1, \dots, x_n^1) - q^1 = 0. \end{array}$$

Aplicando KKT temos que

$$\begin{pmatrix} p(q^1, q^2) + q^1 \dfrac{\partial p}{\partial q^1} + q^1 \dfrac{\partial p}{\partial q^2}\dfrac{\partial q^2}{\partial q^1} - \lambda \\ -w_i(x_i^1, x_i^2) - x_i^1 \dfrac{\partial w_i}{\partial x_i^1} - x_i^1 \dfrac{\partial w_i}{\partial x_i^2}\dfrac{\partial x_i^2}{\partial x_i^1} + \lambda \dfrac{\partial f^1}{\partial x_i^1} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

com $i = 1, \dots, n$. Isolando λ

$$\lambda = p(q^1, q^2) + q^1 \frac{\partial p}{\partial q^1} + q^1 \frac{\partial p}{\partial q^2}\frac{\partial q^2}{\partial q^1}$$

substituindo na outra equação equação

$$w_i(x_i^1, x_i^2) + x_i^1 \frac{\partial w_i}{\partial x_i^1} + x_i^1 \frac{\partial w_i}{\partial x_i^2}\frac{\partial x_i^2}{\partial x_i^1} = \left(p(q^1, q^2) + q^1 \frac{\partial p}{\partial q^1} + q^1 \frac{\partial p}{\partial q^2}\frac{\partial q^2}{\partial q^1} \right) \frac{\partial f^1}{\partial x_i^1}$$

para $i = 1, \dots, n$. Os termos

$$\frac{\partial q^2}{\partial q^1} \quad \text{e} \quad \frac{\partial x_i^2}{\partial x_i^1} \quad i = 1, \dots, n$$

são chamados de variáveis conjucturais, elas indicam as mudanças entre as decisões da firma 2 em relação a firma 1.

4.4 Função conjugada na economia

Considere q a quantidade produzida de uma determinada mercadoria que será vendida a um preço unitário p. A receita total é

$$RT = pq.$$

Considere c um função de custo de produção dependente da quantidade q produzida da mercadoria. O lucro é, então, dado por:

$$\Pi(q) = pq - c(q).$$

Como desejamos maximizar o lucro, podemos relacionar à função conjugada (ver Cap 2) de c:

$$c^*(p) = \sup_q \{pq - c(q)\},$$

a fim de discutir o preço p a ser fixado. Em leis de mercado, procura-se definir o preço p de forma a minimizar a função conjugada c^*.

Se a função lucro cumprir as hipóteses da Proposição 2.11, podemos derivar a função lucro com relação a q para encontrar o máximo desta função.

$$\nabla\Pi(q) = \nabla(pq - c(q)) = 0,$$

ou seja,

$$\nabla pq = \nabla c(q).$$

No entanto podemos observar que ∇pq é a receita marginal (RM) e que $\nabla c(q)$ é o custo marginal (CM), logo

$$MR = MC.$$

Conclusão

Neste trabalho apresentamos a função conjugada de Fenchel, Condições de Otimalidade e algumas aplicações na economia.

Vimos a definição da função conjugada e diversas de suas propriedades. Entre elas podemos destacar que a função conjugada é sempre convexa e que a conjugada da aproximação quadrática é igual à aproximação quadrática da conjugada. Também vimos a interpretação geométrica de algumas dessas propriedades.

A seguir contemplamos uma análise detalhada das condições de KKT usando uma abordagem cônica, que facilita a visualização de certos aspectos da teoria relacionados à KKT. Provamos KKT utilizando uma hipótese fraca, mas difícil de ser verificada, que é a igualdade, no ponto ótimo, do polar do cone viável linearizado e do polar do cone tangente. Neste contexto o Lema de Farkas foi de suma importância, pois o utilizamos indiretamente para provar o Teorema de KKT. Vimos também que as condições de Karush-Kuhn-Tucker são válidas se uma condição de qualificação é satisfeita. Para isto apresentamos diversas condições de qualificação, entre elas a de Slater, independência linear dos gradientes, Mangasarian-Fromovitz, posto constante, dependência linear positiva, quase-normalidade e quase-regularidade. Observamos em seguida as relações de implicações entre estas condições.

Por fim abordamos três diferentes modelos da economia relacionados à teoria da firma: a concorrência perfeita, o monopólio, e o oligopólio. Vimos que estes modelos podem ser encarados como problemas de programação não linear e então aplicamos as condições de KKT a eles. Também vimos brevemente a aplicação da função conjugada na economia.

A parte desta monografia relacionada à função conjugada foi apresentada, sob forma de pôster, nos eventos "Foz 2006 - Congresso de Matemática e suas Aplicações" e "XIV EVINCI - Evento de Iniciação Científica da UFPR". A parte relacionada às condições de otimalidade e de qualificação foi apresentada em forma de pôster no "XI ERMAC - Encontro Regional de Matemática Aplicada e Computacional" realizado em Curitiba; como comunicação oral no "XII ERMAC" realizado em Foz do Iguaçu e no "XVI EVINCI", onde obteve o primeiro lugar na banca examinadora de Matemática Aplicada.

Referências Bibliográficas

[1] M. S. Bazaraa, H. D. Sherali, and C. M. Shetty. *Nonlinear Programming Theory and Algorithms*. John Wiley, New York, 2nd edition, 1993.

[2] R. G. Eustáquio. Condições de otimalidade e de qualificação para problemas de programação não linear. Dissertação de mestrado, Programa de Métodos Numéricos em Engenharia. Universidade Federal do Paraná, 2007.

[3] W. Fenchel. On conjugated convex functions. *Canadian Journal Mathemetics*, pages 73–77, 1949.

[4] A. Friedlander. *Elementos de programação não-linear*. Editora da Unicamp, 1994.

[5] C. C. Gonzaga. Um curso de programação não linear. Brasil, 2004. Notas de aula não publicadas.

[6] F. J. Gould and J. W. Tolle. A necessary and sufficient qualification for constrained optimization. *SIAM Journal on Applied Mathematics*, 20:164–172, 1971.

[7] H. L. Guidorizzi. *Um curso de cálculo*, volume 1. Editora LTC, 5th edition.

[8] H. L. Guidorizzi. *Um curso de cálculo*, volume 2. Editora LTC, 5th edition.

[9] M. Guignard. Generalized Kuhn-Tucker conditions for mathematical programming problems in a Banach space. *SIAM Journal on Control*, 7:232–241, 1969.

[10] J-B. Hiriart-Urruty and C. Lemarechal. *Convex Analysis and Minimization Algorithms I*. Springer-Verlag, New York, 1993.

[11] M. D. Intriligator. *Mathematical Optimization and Economic Theory*. SIAM, 2002.

[12] A. Iusem. Métodos de ponto proximal em otimização. 20^0 Colóquio Brasileiro de Matemática, Rio de Janeiro, Brazil, 1995. In Portuguese.

[13] A. Izmailov and M. Solodov. *Otimização 1*. IMPA, Rio de Janeiro, 2005.

[14] H. W. Kuhn and A. W. Tucker. Nonlinear programming. In J. Neyman, editor, *Proceendings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, pages 481–492. University of California Press, Berkeley, CA, 1951.

[15] D. Kupfer and L. Hasenclever. *Economia Industrial: fundamentos teóricos e práticos no Brasil*. Rio de Janeiro, 2002.

[16] S. J. Leon. *Álgebra Linear com Aplicações.* Editora LTC, Rio de Janeiro, 4th edition, 1998.

[17] E. L. Lima. *Um curso de análise*, volume 1. IMPA, Rio de Janeiro, 8th edition, 1994.

[18] E. L. Lima. *Um curso de análise*, volume 2. IMPA, Rio de Janeiro, 8th edition, 1994.

[19] J. M. Martínez and S. A. Santos. Métodos computacionais de otimização. 20.^{0} Colóquio Brasileiro de Matemática - IMPA, 1995.

[20] M. L. Schuverdt. *Métodos de Langrangiano aumentado com convergência utilizando a condição de dependência linear positiva constante.* PhD thesis, Instituto de Matemática Estatística e Computação Científica, Universidade Estadual de Campinas, Brasil, 2006.