

Renata Aparecida Galletti

**Problemas Diferenciáveis de Programação
não Linear com Tempo Contínuo**



Problemas Diferenciáveis de
Programação Não Linear
com Tempo Contínuo

Renata Aparecida Galletti

**Problemas Diferenciáveis
de Programação não linear
com tempo contínuo.**

Renata Aparecida Galletti

**Dissertação de Mestrado
Pós-Graduação em Matemática Aplicada e
Computacional
MAC-044**



São José do Rio Preto

1999



Problemas Diferenciáveis de Programação Não Linear com Tempo Contínuo

Renata Aparecida Galleti

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

Orientador: Prof. Dr. Geraldo Nunes Silva

São José do Rio Preto
1999



872/99



AGRADECIMENTOS

A realização deste trabalho só foi possível graças a colaboração de muitas pessoas.

Manifesto minha gratidão a todas elas e de forma particular:

Ao Prof. Dr. Geraldo Nunes Silva pela orientação e paciência na elaboração deste trabalho.

Aos professores e funcionários, que de alguma forma contribuíram durante este período.

A todos os colegas pós-graduandos pelo companheirismo.

A meu noivo, Jorge, que jamais deixou de incentivar meu trabalho, pelo apoio, compreensão e carinho.

A meus pais, Santo e Zélia, primeiros mestres na arte de viver, minha gratidão eterna pelo apoio nos momentos mais difíceis.

A Atilio, Regina e Meire pela compreensão e apoio financeiro.

A toda minha família, pelo incentivo e carinho durante este trabalho.

A Deus, presente em todos os momentos de minha vida, vivo lhe agradeço.

Aos meus pais, Santo e Zélia

Ao meu noivo, Jorge

dedico.

AGRADECIMENTOS

A realização deste trabalho só foi possível graças a colaboração de muitas pessoas. Manifesto minha gratidão a todas elas e de forma particular:

Ao Prof. Dr. Geraldo Nunes Silva pela orientação e paciência na elaboração deste trabalho.

Aos professores e funcionários, que de alguma forma contribuíram durante este período.

A todos os colegas pós-graduandos pelo companheirismo.

A meu noivo, Jorge, que jamais deixou de incentivar meu trabalho, pelo apoio, compreensão e carinho.

A meus pais, Santo e Zélia, primeiros mestres na arte de viver, minha gratidão eterna pelo apoio nos momentos mais difíceis.

A Atilio, Regina e Meire pela compreensão e apoio financeiro.

A toda minha família, pelo incentivo e carinho durante este trabalho.

A Deus, presente em todos os momentos de minha vida, elevo meu pensamento.

RESUMO

Neste trabalho, estuda-se condições necessárias e suficientes de otimalidade de Fritz-John e Karush-Kuhn-Tucker para problemas de programação não linear com tempo contínuo. Estuda-se também a relação entre o critério de otimalidade de ponto estacionário e ponto de sela de Fritz-John e usando conceitos de perturbação e estabilidade, desenvolve-se condições de otimalidade de ponto estacionário, de ponto de sela de Karush-Kuhn-Tucker e dualidade Lagrangiana para uma classe de problemas de programação convexa com tempo contínuo. Um resultado auxiliar sobre alternativas, o Teorema de Gordon generalizado, é demonstrado no Apêndice A.

“ Mas é preciso ter força
É preciso ter raça
É preciso ter sonho sempre
Quem traz na pele essa marca
Possui a estranha mania
de ter Fé na vida ... ”
Milton Nascimento.

RESUMO

Neste trabalho, estuda-se condições necessárias e suficientes de otimalidade de Fritz-John e Karush-Kuhn-Tucker para problemas de programação não linear com tempo contínuo. Estuda-se também, a relação entre o critério de otimalidade de ponto estacionário e ponto de sela de Fritz-John e usando conceitos de perturbação e estabilidade, desenvolve-se condições de otimalidade de ponto estacionário, de ponto de sela de Karush-Kuhn-Tucker e dualidade Lagrangiana para uma classe de problemas de programação convexa com tempo contínuo. Um resultado auxiliar sobre alternativas, o Teorema de Gordan generalizado, é demonstrado no Apêndice A.



Índice

Differentiable Problems of Continuous-Time Nonlinear Programming.

Renata Aparecida Galleti

Abstract

This work, concerns of necessary and sufficient optimality conditions of both Fritz-John and Karush-Kuhn-Tucker type for continuous-time nonlinear programming problems. This work also studies the relationship between Fritz-John saddlepoint and stationary-point optimality criteria and using concepts of perturbation and stability, develops Karush-Kuhn-Tucker saddlepoint, stationary-point optimality conditions and Lagrangian duality for a class of continuous-time convex programming problems. An auxiliary result about alternatives, the generalized Gordan theorem, is showed in Appendix A.



Índice

Notações	9
Introdução	11
1 Preliminares	14
1.1 Mensurabilidade e Medida de Lebesgue	14
2 Condições Necessárias para Otimalidade	19
2.1 Introdução	19
2.2 Condições de Otimalidade	22
2.2.1 Condições necessárias de otimalidade de Fritz-John	25
2.2.2 Condições necessárias de otimalidade de Karush-Kuhn-Tucker	27
3 Condições Suficientes para Otimalidade	30
3.1 Convexidade Generalizada	30
3.2 Condições suficientes de otimalidade de Fritz-John	37
3.3 Condições suficientes de otimalidade de Karush-Kuhn-Tucker	40
4 Dualidade em programação não linear com tempo contínuo	50
4.1 Relação entre o critério de otimalidade de ponto estacionário e ponto de sela.	50
4.2 Condições de Otimalidade e Dualidade Lagrangiana em Programação não linear com tempo contínuo.	54
4.2.1 Os problemas primal e dual	54
4.2.2 Resultados Auxiliares	57



4.2.3	Condições de otimalidade	65
4.2.4	Dualidade Lagrangiana	71
	Conclusão	76
	Referências Bibliográficas	77
	Apêndice A	80

$A(T)$ conjunto dos pontos ativos

C cone convexo

(D) problema dual

$\partial f(x)$ subdiferencial de f em x

$\frac{\partial f}{\partial x}$ derivada parcial de f em relação a variável x

$Df(x)$ derivada de Fréchet de f em x

$D^2 f(x)$ segunda derivada de Fréchet de f em x

$D(\phi, Z)$ conjunto de todas as direções da derivada de ϕ em Z

ess sup supremo essencial

$(E)_*$ medida exterior de E

$|E|$ medida de Lebesgue

f^* versão λ de f

F -derivada / derivada de Fréchet

$F(S, Z)$ conjunto das direções factíveis de S em Z

F conjunto de todas as soluções factíveis

∇f gradiente de f

$\nabla_x f$ gradiente de f em relação a variável x

$F(f, E)$ gráfico de f ao longo do conjunto E

I^- intervalo fechado

$I_1 \times I_2$ produto cartesiano de I_1 e I_2

$\inf$ infimo

$\text{int} X$ interior de X

Notações

$A(\bar{x})$ conjunto dos pontos ativos

C cone convexo

(D) problema dual

$\partial f(x)$ subdiferencial de f em x

$\frac{\partial f}{\partial x}$ derivada parcial de f com relação a variável x

$Df(x)$ derivada de Fréchet de f em x

$D^2 f(x)$ segunda derivada de Fréchet de f em x

$D(\phi_i, \bar{x})$ conjunto de todas as direções de descida de ϕ em $\bar{x}$

ess sup supremo essencial

$|E|_e$ medida exterior de E

$|E|$ medida de Lebesgue

f^x sessão x de f

F-derivada derivada de Fréchet

$\mathcal{F}(S, \bar{x})$ conjunto das direções factíveis de S em $\bar{x}$

$\mathcal{F}$ conjunto de todas as soluções factíveis

∇f gradiente de f

$\nabla_x f$ gradiente de f com relação a variável x

$\Gamma(f, E)$ gráfico de f ao longo do conjunto E

I intervalo fechado

$I_1 \times I_2$ produto cartesiano de I_1 e I_2

$\inf$ infimo

$\text{int}X$ interior de X



$\int_E f(x)dx$ integral de Lebesgue de f no conjunto E

$\lim$ limite

$\mathcal{L}(E)$ conjunto das funções Lebesgue integráveis em E

$L(\cdot, \lambda)$ função Lagrangiana

$\Lambda_1^m[0, T]$ espaço das funções vetoriais essencialmente limitadas com norma $\|\cdot\|_1$

$\mathcal{L}(X, Y)$ classe de todas as aplicações lineares contínuas

$\mathcal{L}_\infty^n[0, T]$ espaço das funções vetoriais essencialmente limitadas com norma $\|\cdot\|_\infty$

$\max$ maximizar

$\min$ minimizar

$\|\cdot\|$ norma Euclidiana

$\|\cdot\|_1$ norma um

$\|\cdot\|_\infty$ norma infinito

(P) problema primal

$p(\cdot)$ função perturbação

q. s. quase sempre

$\mathbb{R}$ conjunto dos números reais

$\mathbb{R}^n$ espaço euclidiano n -dimensional

$R(f, E)$ região inferior a função f ao longo do conjunto E

$\mathcal{R}$ região convexa de um espaço Euclidiano

$\sup$ supremo

$U(x)$ vizinhança de x

$\emptyset$ conjunto vazio

x' transposto de x

$x'y$ produto interno de x e y

Y conjunto factível do problema perturbado

Z^* dual topológico de Z

$\sim$ relação de equivalência

Introdução

Programação com tempo contínuo originou-se nos estudos de Bellman [3, 4]. Considerando como uma generalização para um problema de programação linear, formulou um correspondente problema dual, obtendo o teorema de dualidade fraca e sugeriu alguns procedimentos computacionais.

A formulação de Bellman e a teoria da dualidade foi estendida para problemas gerais de programação linear com tempo contínuo e também para problemas de programação não linear com tempo contínuo.

Condições de otimalidade de Karush-Kuhn-Tucker em programação com tempo contínuo foi primeiro estudado por Hanson e Mond [12], considerando um problema não linear com restrições lineares. Seu método para obter condições de otimalidade foi indireto visto que a função objetivo é aproximada por funções lineares, possibilitando aplicar os resultados de dualidade linear de Levinson [14] para obter um teorema de dualidade para problemas não lineares e então deduzir as condições de Karush-Kuhn-Tucker como consequência deste teorema de dualidade não linear.

Farr e Hanson [6], introduziram restrições não lineares. Supondo algumas condições de positividade semelhantes as de Levinson [14], aplicando o método de linearização e usando os resultados de dualidade linear de Grinold [11], obtiveram dualidade para problemas lineares e com condições adicionais para a função f , obtiveram um teorema de dualidade para problemas não lineares. Os resultados de Farr e Hanson [6], foram aperfeiçoados e generalizados por Reiland e Hanson [19].

Recentemente, condições de otimalidade para problemas não lineares com tempo contínuo foram obtidos por métodos diretos. Em [1], certas suposições de regularidade foram



usadas para obter condições de Karush-Kuhn-Tucker para problemas de programação convexa.

Geoffrion, seguido de Gale [7], usando uma excelente mistura de intuição geométrica e raciocínio analítico, desenvolveram uma compreensiva e elegante teoria de dualidade para programação não linear de dimensão finita.

Reiland [20], aplicou a versão com tempo contínuo das restrições de qualificação de Zangwill [27], introduzidas em [19] e o teorema de Farkas para dimensão infinita, obtendo condições de otimalidade e dualidade para problemas diferenciáveis com tempo contínuo.

Problemas não convexos sem diferenciabilidade foram considerados em [17, 22], onde foi discutido condições de otimalidade de primeira e segunda ordem e dualidade.

Neste trabalho, estuda-se condições de otimalidade e dualidade em programação não linear com tempo contínuo, no caso diferenciável.

A motivação para este trabalho é organizar os melhores resultados existentes na literatura sobre condições de otimalidade e dualidade para programação não linear com tempo contínuo diferenciável. Assim a construção de algoritmos computacionais, a análise e a interpretação de modelos dinâmicos de política econômica ficam bastante facilitados, pois o material necessário poderá ser encontrado num único local.

No capítulo 1, recorda-se algumas definições e resultados básicos que serão necessários para o desenvolvimento deste trabalho.

No capítulo 2, recorda-se as definições de diferenciabilidade de Fréchet, adota-se um ponto de vista geométrico e desenvolve-se condições necessárias de otimalidade em programação não linear com tempo contínuo de Fritz-John e Karush-Kuhn-Tucker [25].

No capítulo 3, desenvolve-se condições suficientes de otimalidade em programação não linear com tempo contínuo de Fritz-John e Karush-Kuhn-Tucker para problemas com restrições de desigualdades e igualdades não lineares com suposição de convexidade generalizada [26]. Os teoremas são generalizações de resultados de suficiência de primeira ordem na área de programação não linear determinística de dimensão finita e produzem condições suficientes de otimalidade para problemas de controle ótimo.

No capítulo 4, estuda-se a relação entre condições de otimalidade de ponto estacionário



e ponto de sela, e considera-se uma extensão de dualidade Lagrangiana clássica para uma certa classe de problemas de programação convexa com tempo contínuo e não diferenciável [23, 25].

Por fim, apresenta-se as conclusões do trabalho e as referências bibliográficas.

No Apêndice A, usando teoria da dualidade apresentada no capítulo 4, demonstra-se o Teorema de Gordan generalizado [24].

Preliminares

Neste capítulo recorda-se algumas definições e resultados básicos da teoria da medida e integração de Lebesgue [28] que posteriormente serão necessários para o desenvolvimento de condições de otimalidade e teoria de dualidade para programas não lineares com tempo contínuo.

1.1 Mensurabilidade e Medida de Lebesgue

Seja E um subconjunto do espaço euclidiano n -dimensional, denotado R^n .

Definição 1.1.1 *Seja I um intervalo fechado n -dimensional definido por*

$$I = \{x : a_j \leq x_j \leq b_j, j = 1, \dots, n\}$$

e seu volume $v(I) = \prod_{j=1}^n (b_j - a_j)$. Seja C uma coleção enumerável de intervalos I_k , tal que

$$E \subset \bigcup_{k=1}^{\infty} I_k \text{ e seja}$$

$$v(C) = \sum_{k=1}^{\infty} v(I_k).$$

Neste caso diz-se que C é uma cobertura de E . A medida exterior de E , denotada por $|E|_e$, é definida por

$$|E|_e = \inf v(C).$$

Capítulo 1

Preliminares

Neste capítulo recorda-se algumas definições e resultados básicos da teoria da medida e integração de Lebesgue [28] que posteriormente serão necessários para o desenvolvimento de condições de otimalidade e teoria de dualidade para programas não lineares com tempo contínuo.

1.1 Mensurabilidade e Medida de Lebesgue

Seja E um subconjunto do espaço euclidiano n -dimensional, denotado $\mathbb{R}^n$.

Definição 1.1.1 *Seja I um intervalo fechado n -dimensional definido por*

$$I = \{x : a_j \leq x_j \leq b_j, j = 1, \dots, n\}$$

e seu volume $v(I) = \prod_{j=1}^n (b_j - a_j)$. Seja $\mathcal{C}$ uma coleção enumerável de intervalos I_k , tal que $E \subset \bigcup_{k=1}^{\infty} I_k$ e seja

$$\sigma(\mathcal{C}) = \sum_{I_k \in \mathcal{C}} v(I_k).$$

Neste caso diz-se que $\mathcal{C}$ é uma cobertura de E . A medida exterior de E , denotada por $|E|_e$ é definida por

$$|E|_e = \inf \sigma(\mathcal{C}),$$



onde o ínfimo é tomado sobre todas as coberturas $\mathcal{C}$ de E . Assim, $0 \leq |E|_e \leq +\infty$.

Um subconjunto E é Lebesgue mensurável, ou simplesmente mensurável, se dado $\epsilon > 0$, existe um conjunto aberto G tal que

$$E \subset G \text{ e } |G - E|_e < \epsilon.$$

Se E é mensurável, sua medida exterior é chamada de medida de Lebesgue, ou simplesmente medida, e denotada por $|E|$, ou seja,

$$|E| = |E|_e, \text{ quando } E \text{ é mensurável.}$$

Definição 1.1.2 Seja f uma função a valor real definida em um conjunto $E \subset \mathbb{R}^n$, isto é, $-\infty \leq f(x) \leq +\infty$, $x \in E$. Então a função f é chamada Lebesgue mensurável em E , ou simplesmente mensurável, se para qualquer $a \in \mathbb{R}$, o conjunto

$$\{x \in E : f(x) > a\}$$

é um subconjunto mensurável de $\mathbb{R}^n$.

Definição 1.1.3 Seja f uma função não negativa, $0 \leq f \leq +\infty$, definida em um subconjunto mensurável $E \in \mathbb{R}^n$. Sejam

$$\Gamma(f, E) = \{(x, f(x)) \in \mathbb{R}^{n+1} : x \in E, f(x) < +\infty\},$$

$$R(f, E) = \{(x, y) \in \mathbb{R}^{n+1} : x \in E, 0 \leq y \leq f(x) \text{ se } f(x) < +\infty \\ \text{e } 0 \leq y < +\infty \text{ se } f(x) = +\infty\}.$$

$\Gamma(f, E)$ é chamado de gráfico de f ao longo do conjunto E e $R(f, E)$ a região inferior a função f ao longo do conjunto E . Se $R(f, E)$ é mensurável (como um subconjunto de $\mathbb{R}^{n+1}$), sua medida $|R(f, E)|_{n+1}$ é chamada de integral de Lebesgue de f no conjunto E , definida por

$$|R(f, E)|_{n+1} = \int_E f(x) dx.$$

Usualmente, uma das abreviações

$$\int_E f dx \text{ ou } \int_E f$$

é usada, e as vezes a notação por extenso

$$\int_E \dots \int f(x_1, \dots, x_n) dx_1 \dots dx_n$$

é conveniente.

Agora, se f é uma função mensurável arbitrária definida em um conjunto E , então $f = f^+ - f^-$, onde $f^+ = \max\{f, 0\}$ e $f^- = -\min\{f, 0\}$ são funções mensuráveis. Além disso, as integrais $\int_E f^+(x)dx$ e $\int_E f^-(x)dx$ existem e são não negativas, possivelmente tomando o valor $+\infty$. Quando pelo menos uma das integrais é finita, define-se

$$\int_E f(x)dx = \int_E f^+(x)dx - \int_E f^-(x)dx,$$

e diz-se que $\int_E f(x)dx$ existe.

Se $\int_E f$ existe, é claro que, $-\infty \leq \int_E f \leq +\infty$. Se $\int_E f$ existe e é finita, diz-se que f é Lebesgue integrável, ou simplesmente integrável, em E e denota-se por $f \in \mathcal{L}(E)$. Assim,

$$\mathcal{L}(E) = \left\{ f : \int_E f \text{ é finita} \right\}.$$

Se $\int_E f$ existe, então

$$\left| \int_E f \right| \leq \int_E f^+ + \int_E f^- = \int_E (f^+ + f^-),$$

por ([28], Teorema 5.14). Desde que $f^+ + f^- = |f|$, tem-se

$$\left| \int_E f dx \right| \leq \int_E |f| dx. \quad (1.1)$$

Definição 1.1.4 Sejam f e $\{f_k\}$ funções mensuráveis as quais são definidas quase sempre em um conjunto E , então $\{f_k\}$ converge quase sempre para a função f se para quase todo $x \in E$ tem-se

$$f_k(x) \longrightarrow f(x),$$

e denota-se por

$$f_k \longrightarrow f \text{ q. s. em } E.$$

Teorema 1.1.1 (Teorema da convergência dominada de Lebesgue)

Seja $\{f_k\}$ uma sequência de funções mensuráveis em E tal que $f_k \rightarrow f$ q. s. em E . Se existe $\phi \in \mathcal{L}(E)$ tal que $|f_k| \leq \phi$ q. s. em E para todo k , então $\int_E f_k \rightarrow \int_E f$.

Definição 1.1.5 O supremo essencial de $x(t)$ em $[0, T]$ é definido por:

$$\operatorname{ess\,sup}_{t \in [0, T]} x(t) = \inf \{ \alpha \in \mathbb{R} : |\{t \in [0, T] : x(t) > \alpha\}| = 0 \}.$$

Diz-se que $x(t)$ é essencialmente limitado em $[0, T]$ se

$$\operatorname{ess\,sup}_{t \in [0, T]} |x(t)| \text{ é finito.}$$

A classe de todas as funções que são essencialmente limitadas em $[0, T]$ é denotada por $\mathcal{L}_\infty[0, T]$.

Seja $x = (x_1, x_2, \dots, x_n)'$ sendo um ponto do n -dimensional intervalo I_1 definido por

$$I_1 = \{x = (x_1, x_2, \dots, x_n)' : a_i \leq x_i \leq b_i, i = 1, \dots, n\},$$

onde o vetor x' denota o transposto do vetor x e seja y sendo um ponto do m -dimensional intervalo I_2 definido por

$$I_2 = \{y = (y_1, y_2, \dots, y_m)' : c_j \leq y_j \leq d_j, j = 1, \dots, m\},$$

assim $I_1 \in \mathbb{R}^n$ e $I_2 \in \mathbb{R}^m$.

O produto cartesiano $I = I_1 \times I_2$ é um $(n + m)$ -dimensional intervalo consistindo de pontos da forma $(x_1, \dots, x_n, y_1, \dots, y_m)'$ definidos em I e podendo ser escritos como $f(x, y)$ e sua integral $\int_I f$ pode ser escrita como

$$\int_I f(x, y) dx dy.$$

O produto interno de dois vetores x e y é definido por

$$x' y = x_1 y_1 + x_2 y_2 + \dots + x_n y_n,$$

e a notação $\nabla f(x_0)$ é usada para representar o vetor

$$\left(\frac{\partial f(x)}{\partial x_1}, \frac{\partial f(x)}{\partial x_2}, \dots, \frac{\partial f(x)}{\partial x_n} \right)' \text{ calculada em } x = x_0.$$

Definição 1.1.6 Seja f uma função definida em $\mathbb{R}^n \times \mathbb{R}^m$, então dado $x \in \mathbb{R}^n$, a sessão x de f é a função em $\mathbb{R}^m : f^x(y) = f(x, y)$ e a sessão y de f é a função em $\mathbb{R}^n : f^y(x) = f(x, y)$.



Teorema 1.1.2 (Teorema de Fubini)

Seja $f(x, y) \in \mathcal{L}(I_1 \times I_2)$. Então $f^x \in \mathcal{L}(I_2)$ para quase todo $x \in \mathbb{R}^n$, e $f^y \in \mathcal{L}(I_1)$ para quase todo $y \in \mathbb{R}^m$. Além disso, define-se as funções

$$\phi(x) = \int_{\mathbb{R}^m} f^x dy$$

e

$$\psi(y) = \int_{\mathbb{R}^n} f^y dx,$$

para cada $x \in \mathbb{R}^n$ e $y \in \mathbb{R}^m$. Então ϕ e ψ são mensuráveis e

$$\int_{\mathbb{R}^n} \phi dx = \int_{\mathbb{R}^n \times \mathbb{R}^m} f dx dy = \int_{\mathbb{R}^m} \psi dy$$

com $\phi(x) \in \mathcal{L}(I_1)$ e $\psi(y) \in \mathcal{L}(I_2)$.

Neste capítulo, começa-se com condições geométricas e depois desenvolve-se condições necessárias de optimalidade de Fritz-John e Karush-Kuhn-Tucker em programação não linear com tempo contínuo, no caso diferenciável.

2.1 Introdução

Considere o problema Primal de programação não linear com tempo contínuo:

$$(P) \quad \begin{aligned} \min \quad & \phi(x) = \int_0^T f(x(t), t) dt, \\ \text{s. a.} \quad & g(x(t), t) \leq 0 \text{ q. s. em } [0, T], \\ & x \in X, \end{aligned}$$

onde X é um subconjunto convexo aberto não vazio de um espaço de Banach $\mathcal{L}_{\infty}^n[0, T]$ de todas as funções vetoriais n -dimensionais as quais são Lebesgue Mensuráveis essencialmente limitadas definidas sobre o intervalo compacto $[0, T] \subset \mathbb{R}$, com norma $\|\cdot\|_{\infty}$ definida por:

$$\|x\|_{\infty} = \max_{1 \leq j \leq n} \sup_{t \in [0, T]} |x_j(t)|.$$

Capítulo 2

Condições Necessárias para Otimalidade

Neste capítulo, começa-se com condições geométricas e depois desenvolve-se condições necessárias de otimalidade de Fritz-John e Karush-Kuhn-Tucker em programação não linear com tempo contínuo, no caso diferenciável.

2.1 Introdução

Considere o problema Primal de programação não linear com tempo contínuo:

$$\begin{aligned} \min \quad & \phi(x) = \int_0^T f(x(t), t) dt, \\ (P) \quad & \text{s. a. } g(x(t), t) \leq 0 \quad \text{q. s. em } [0, T], \\ & x \in X, \end{aligned}$$

onde X é um subconjunto convexo aberto não vazio de um espaço de Banach $\mathcal{L}_\infty^n[0, T]$ de todas as funções vetoriais n -dimensional as quais são Lebesgue Mensuráveis essencialmente limitadas definidas sobre o intervalo compacto $[0, T] \subset \mathbb{R}$, com norma $\|\cdot\|_\infty$ definida por:

$$\|x\|_\infty = \max_{1 \leq j \leq n} \text{ess sup}_{t \in [0, T]} |x_j(t)|.$$

Seja $\Lambda_1^m[0, T]$ o espaço de todas as funções vetoriais m -dimensionais, Lebesgue mensuráveis e essencialmente limitadas definidas sobre $[0, T]$ com norma $\|\cdot\|_1$ definida por:

$$\|y\|_1 = \max_{1 \leq j \leq m} \int_0^T |y_j(t)| dt.$$

Aqui, para $t \in [0, T]$, x_j é o j -ésimo componente de $x(t) \in \mathbb{R}^n$. Seja ϕ uma função a valor real definida em X e $g(x(t), t) := \Gamma(x)(t)$, onde Γ é uma aplicação com domínio em X tomando valores no espaço normado $\Lambda_1^1[0, T]$.

Definição 2.1.1 *Seja $\mathcal{I}$ o conjunto de todas as soluções factíveis do problema (P) (suponha não vazio) dado por:*

$$\mathcal{I} = \{x \in X : g(x(t), t) \leq 0 \text{ q. s. em } [0, T]\}.$$

Seja V um conjunto aberto em $\mathbb{R}^n$ contendo o conjunto

$$\{x(t) \in \mathbb{R}^n : x \in \mathcal{I}, t \in [0, T]\}.$$

Assim f e g são funções a valores reais definidas em $V \times [0, T]$. Suponha que a função $t \rightarrow f(x(t), t)$ seja Lebesgue Mensurável e integrável em $x \in X$.

Definição 2.1.2 *Seja $A(\bar{x})$ o conjunto que corresponde aos pontos $t \in [0, T]$ que em g é ativo e é dado por:*

$$A(\bar{x}) = \{t \in [0, T] : g(\bar{x}(t), t) = 0\}, \forall \bar{x} \in \mathcal{I}.$$

Definição 2.1.3 *Seja S um subconjunto de um Espaço de Banach Real E e $\bar{x}$ pertencente ao fecho de S . Então um ponto $y \in E$ é chamado de direção factível de S em $\bar{x}$ se existe um número real $\delta > 0$ tal que*

$$\bar{x} + \lambda y \in S \text{ para todo } 0 < \lambda \leq \delta.$$

O conjunto de todas as direções factíveis de S em $\bar{x}$ será denotado por:

$$\mathcal{F}(S, \bar{x}).$$



Definição 2.1.4 *Seja a função $\psi : E \longrightarrow \mathbb{R}$ e $\bar{x} \in E$. Então um ponto $y \in E$ é chamado de direção de descida de ψ em $\bar{x}$ se existe um número real $\delta > 0$ tal que*

$$\psi(\bar{x} + \lambda y) < \psi(\bar{x}) \text{ para todo } 0 < \lambda < \delta.$$

O conjunto de todas as direções de descida de ψ em $\bar{x}$ será denotado por $\mathcal{D}(\psi, \bar{x})$.

$$\mathcal{D}(\psi, \bar{x}) = \{y(t) : \psi(\bar{x} + \lambda y)(t) < \psi(\bar{x})(t), t \in A(\bar{x})\}.$$

Obs: Todos os vetores são vetores colunas exceto se transpostos e todas as integrais são no sentido de Lebesgue. O transposto de u é denotado por u' .

Teorema 2.1.1 [24], (Teorema de Gordan Generalizado)

Seja X um subconjunto convexo não vazio de $\mathcal{L}_{\infty}^n[0, T]$. Seja a aplicação $p : V \times [0, T] \longrightarrow \mathbb{R}^m$, definida por $p(x(t), t) = \pi(x)(t)$, onde π é uma aplicação de X em $\Lambda_1^m[0, T]$ e suponha que p é convexo com relação a sua primeira variável sobre V por todo $[0, T]$. Então

(a) *existe $x \in X$ tal que*

$$p(x(t), t) < 0 \text{ q. s. em } [0, T]$$

ou

(b) *existe uma função m -vetorial não nula $u \in \mathcal{L}_{\infty}^m[0, T]$, $u(t) \geq 0$ q. s. em $[0, T]$ tal que*

$$\int_0^T u'(t)p(x(t), t)dt \geq 0 \text{ para todo } x \in X.$$

Mas nunca ambos.

A demonstração deste teorema exige teoria da dualidade que será apresentado no capítulo 4. Assim opta-se por colocá-la no Apêndice A.

2.2 Condições de Otimalidade

Nesta seção, estuda-se a versão com tempo contínuo do critério de otimalidade necessário de Fritz-John, primeiramente caracterizando a noção de otimalidade local em termos da ausência de direções factíveis que melhorem a função objetivo.

Começa-se recordando a definição da derivada de Fréchet que será necessária para o desenvolvimento de condições de otimalidade.

Definição 2.2.1 *Sejam $f : U(x) \subseteq X \rightarrow Y$ sendo uma aplicação dada, X e Y : espaços de Banach e $U(x)$ uma vizinhança de x . A aplicação f é diferenciável no sentido de Fréchet (F -diferenciável) em x se e somente se existe uma aplicação $T \in \mathcal{L}(X, Y)$ tal que*

$$f(x + y) - f(x) = Th + o(h),$$

onde $\lim_{h \rightarrow 0} \frac{o(h)}{\|h\|} = 0$.

Aqui $\mathcal{L}(X, Y)$ denota a classe de todas as aplicações lineares contínuas $T : X \rightarrow Y$. Se esta aplicação T existe, então T é chamada de F -derivada de f em x . Denota-se a derivada de Fréchet de f em x por:

$$Df(x) := T.$$

Seja $\phi : \mathcal{L}_{\infty}^n[0, T] \rightarrow \mathbb{R}$ uma função F -diferenciável, então o cone de direções de descida em $\bar{x}$ é o conjunto

$$\{h \in \mathcal{L}_{\infty}^n[0, T] : D\phi(\bar{x})h < 0\}.$$

Se um ponto h no cone das direções factíveis do conjunto factível $\mathcal{IF}$ do problema (P) satisfaz a desigualdade

$$D\phi(\bar{x})h = \int_0^T \nabla f'(x(t), t)h(t)dt < 0$$

então pode-se achar $z \in \mathcal{L}_{\infty}^n[0, T]$ em uma vizinhança de $\bar{x}$ tal que $f(z) < f(\bar{x})$. Isto significa que se $\bar{x}$ é um ponto de mínimo local do problema (P) e h satisfaz a desigualdade acima então h não pode ser uma direção factível pois contraria a definição de condição necessária de otimalidade local do problema (P) . Formalmente tem-se o seguinte resultado:

Teorema 2.2.1 Suponha que ϕ é Fréchet diferenciável em $\bar{x} \in \mathbb{F}$. Se $\bar{x}$ é uma solução ótima local do problema (P), então

$$\mathcal{F}(\mathbb{F}, \bar{x}) \cap \mathcal{D}(\phi, \bar{x}) = \emptyset. \quad (2.1)$$

Demonstração :

Se $y \in \mathcal{F}(\mathbb{F}, \bar{x}) \cap \mathcal{D}(\phi, \bar{x})$ então existe $\delta_1, \delta_2 > 0$ tal que

$$\bar{x} + \lambda y \in \mathbb{F} \text{ para todo } 0 < \lambda < \delta_1$$

e

$$\phi(\bar{x} + \lambda y) < \phi(\bar{x}) \text{ para todo } 0 < \lambda < \delta_2.$$

Assim para $\delta = \min\{\delta_1, \delta_2\}$ existirá uma solução factível $\bar{x} + \lambda y$, $0 < \lambda < \delta$ com melhor valor da função objetivo. O que é uma contradição pela hipótese de que $\bar{x}$ é mínimo local.

Assim, (2.1) vale. ■

Teorema 2.2.2 Seja $\bar{x} \in \mathbb{F}$, e suponha que ϕ e Γ , são Fréchet diferenciáveis em $\bar{x}$, e que Γ é contínua em $\bar{x}$. Se $\bar{x}$ é uma solução ótima local do problema (P), então

$$\mathcal{D}(\phi, \bar{x}) \cap \mathcal{D}(\Gamma, \bar{x}) = \emptyset. \quad (2.2)$$

Demonstração :

Primeiro deve-se mostrar que o cone

$$\mathcal{D}(\Gamma, \bar{x}),$$

é um subconjunto de $\mathcal{F}(\mathbb{F}, \bar{x})$, e usar o Teorema (2.2.1). Seja

$$y \in \mathcal{D}(\Gamma, \bar{x}).$$

Já que $\bar{x} \in X$ e X é aberto, então existe um $\delta_0 > 0$ tal que

$$\bar{x} + \lambda y \in X, \text{ para todo } 0 < \lambda < \delta_0.$$

Já que

$$y \in \mathcal{D}(\Gamma, \bar{x}),$$

então existe um $\delta_1 > 0$ tal que

$$\Gamma(\bar{x} + \lambda y)(t) < \Gamma(\bar{x})(t), \text{ para todo } 0 < \lambda < \delta_1.$$

Por definição tem-se que

$$\Gamma(\bar{x})(t) = g(\bar{x}(t), t).$$

Como $\bar{x} \in \mathcal{IF}$ tem-se

$$g(\bar{x}(t), t) \leq 0 \text{ q. s. em } [0, T].$$

Assim

$$\Gamma(\bar{x} + \lambda y)(t) < 0 \text{ q. s. em } [0, T], \text{ para todo } 0 < \lambda < \delta_1.$$

Finalmente, como Γ é contínua em $\bar{x}$, pode-se achar $\delta_1' > 0$ tal que

$$\Gamma(\bar{x} + \lambda y)(t) \leq 0 \text{ q. s. em } [0, T], \text{ para todo } 0 < \lambda < \delta_1'.$$

Agora, tomando

$$\delta = \min\{\delta_0, \delta_1'\}$$

tem-se que os pontos da forma

$$\bar{x} + \lambda y, \quad 0 < \lambda < \delta$$

são factíveis para o problema (P). Assim

$$y \in \mathcal{F}(\mathcal{IF}, \bar{x})$$

e portanto

$$\mathcal{D}(\Gamma, \bar{x}) \subset \mathcal{F}(\mathcal{IF}, \bar{x}).$$

Já que $\bar{x}$ é uma solução ótima local do problema (P), pelo Teorema (2.2.1), tem-se que

$$\mathcal{F}(\mathcal{IF}, \bar{x}) \cap \mathcal{D}(\phi, \bar{x}) = \emptyset$$

e assim (2.2) vale. ■

Agora, pelo Teorema (2.2.2) e o Teorema de Gordan (2.1.1) pode-se determinar condições necessárias de otimalidade de Fritz-John.

2.2.1 Condições necessárias de otimalidade de Fritz-John

Teorema 2.2.3 *Seja $\bar{x} \in IF$. Suponha que f e g são continuamente diferenciáveis com relação a sua primeira variável em $\bar{x}$ por todo $[0, T]$ e que a função $t \rightarrow \nabla f(\bar{x}(t), t)$ e $t \rightarrow \nabla g(\bar{x}(t), t)h(t) \equiv D\Gamma(\bar{x})(h)(t)$ são Lebesgue integráveis em $[0, T]$ para todo $h \in \mathcal{L}_\infty^n[0, T]$. Se $\bar{x}$ é uma solução ótima local do problema (P), então existe $\bar{u}_0 \in \mathbb{R}$, $\bar{u} \in \mathcal{L}_\infty[0, T]$ tal que*

$$\int_0^T [\bar{u}_0 \nabla f'(\bar{x}(t), t) + \bar{u}(t) \nabla g(\bar{x}(t), t)] h(t) dt = 0 \text{ para todo } h \in \mathcal{L}_\infty^n[0, T]; \quad (2.3)$$

$$\bar{u}(t) g(\bar{x}(t), t) = 0 \text{ q. s. em } [0, T]; \quad (2.4)$$

$$\bar{u}_0 \geq 0, \bar{u}(t) \geq 0 \text{ q. s. em } [0, T]; \quad (2.5)$$

$$(\bar{u}_0, \bar{u}(t)) \neq 0 \text{ q. s. em } [0, T]. \quad (2.6)$$

Demonstração :

Já que $\bar{x}$ é solução ótima do problema (P), o sistema de desigualdade linear

$$\begin{cases} \int_0^T \nabla f'(\bar{x}(t), t) h(t) dt < 0, \\ \nabla g(\bar{x}(t), t) h(t) < 0 \text{ para } t \in A(\bar{x}), \end{cases} \quad (2.7)$$

não tem solução $h \in \mathcal{L}_\infty^n[0, T]$ pois

$$D\phi(\bar{x})h = \int_0^T \nabla f'(\bar{x}(t), t) h(t) dt < 0 \implies h \in \mathcal{D}(\phi, \bar{x})$$

e

$$D\Gamma(\bar{x})h(t) = \nabla g(\bar{x}(t), t)h(t) < 0 \implies h \in \mathcal{D}(\Gamma, \bar{x}).$$

Pelo Teorema (2.2.2) tem-se

$$\mathcal{D}(\phi, \bar{x}) \cap \mathcal{D}(\Gamma, \bar{x}) = \emptyset.$$

Assim o sistema de desigualdade linear (2.7) não tem solução $h \in \mathcal{L}_\infty^n[0, T]$. Logo, pelo Teorema (2.1.1), existe $\bar{u}_0 \in \mathbb{R}$, $\bar{u} \in \mathcal{L}_\infty[0, T]$ tal que (2.5) e (2.6) são satisfeitos e

$$\int_0^T \bar{u}_0 \nabla f'(\bar{x}(t), t) h(t) dt + \int_{A(\bar{x})} \bar{u}(t) \nabla g(\bar{x}(t), t) h(t) dt \geq 0 \text{ para todo } h \in \mathcal{L}_\infty^n[0, T].$$

Agora, define-se

$$\bar{u}(t) = \begin{cases} \bar{u}(t) & \text{se } t \in A(\bar{x}), \\ 0 & \text{se } t \notin A(\bar{x}). \end{cases} \quad (2.8)$$

Assim, a desigualdade acima pode ser escrita da seguinte forma:

$$\int_0^T [u_0 \nabla f'(\bar{x}(t), t) + \bar{u}(t) \nabla g(\bar{x}(t), t)] h(t) dt \geq 0 \text{ para todo } h \in \mathcal{L}_\infty^n[0, T].$$

Já que a desigualdade acima é satisfeita para todo $h \in \mathcal{L}_\infty^n[0, T]$, então deverá ser satisfeita para a igualdade. Assim,

$$\int_0^T [u_0 \nabla f'(\bar{x}(t), t) + \bar{u}(t) \nabla g(\bar{x}(t), t)] h(t) dt = 0 \text{ para todo } h \in \mathcal{L}_\infty^n[0, T].$$

Sabe-se também que:

$$\bar{u}(t)g(\bar{x}(t), t) = 0 \text{ q. s. em } [0, T], \forall t.$$

De fato:

(a) Se $t \in A(\bar{x})$ então $g(\bar{x}(t), t) = 0$

Assim, $\bar{u}(t)g(\bar{x}(t), t) = 0$ q. s. em $[0, T]$.

(b) Se $t \notin A(\bar{x})$ então por definição tem-se que $\bar{u}(t) = 0$

Assim, $\bar{u}(t)g(\bar{x}(t), t) = 0$ q. s. em $[0, T]$. ■

Uma melhor ilustração das condições de otimalidade vistas acima pode ser obtida considerando uma forma explícita da função g . Seja

$$\begin{aligned} \min \quad & \phi(x) = \int_0^T f(x(t), t) dt, \\ (\bar{P}) \quad & \text{s. a. } P(x(t), t) \leq r(t) + \int_0^T Q(x(s), t, s) ds \text{ q. s. em } [0, T], \\ & x \in X. \end{aligned} \quad (2.9)$$

X e f são definidos como no problema (P). P , r e Q são funções definidas em $V \times [0, T]$, $[0, T]$ e $V \times [0, T] \times [0, T]$ a valores real. Assim sob suposições apropriadas para o problema $\bar{P}$, (2.3) torna-se:

$$\int_0^T [\bar{u}_0 \nabla f'(\bar{x}(t), t) h(t) + \bar{u}(t) \nabla P'(\bar{x}(t), t) h(t) - \int_0^T \bar{u}(t) \nabla Q'(\bar{x}(s), t, s) h(s) ds] dt = 0$$

para todo $h \in \mathcal{L}_{\infty}^n[0, T]$. Aplicando o Teorema de Fubini (1.1.2) tem-se:

$$\int_0^T h'(t) [\bar{u}_0 \nabla f(\bar{x}(t), t) + \bar{u}(t) \nabla P(\bar{x}(t), t) - \int_t^T \bar{u}(s) \nabla Q(\bar{x}(t), s, t) ds] dt = 0$$

para todo $h \in \mathcal{L}_{\infty}^n[0, T]$. Assim a expressão dentro do colchetes é igual a zero e (2.3) - (2.6) do problema $\bar{P}$ tornam-se:

$$\bar{u}_0 \nabla f(\bar{x}(t), t) + \bar{u}(t) \nabla P(\bar{x}(t), t) - \int_t^T \bar{u}(s) \nabla Q(\bar{x}(t), s, t) ds = 0 \quad \text{q. s. em } [0, T];$$

$$\bar{u}(t) [P(\bar{x}(t), t) - r(t) - \int_0^t Q(\bar{x}(s), t, s) ds] = 0 \quad \text{q. s. em } [0, T];$$

$$\bar{u}_0 \geq 0, \bar{u}(t) \geq 0 \quad \text{q. s. em } [0, T];$$

$$(\bar{u}_0, \bar{u}(t)) \neq 0 \quad \text{q. s. em } [0, T].$$

2.2.2 Condições necessárias de otimalidade de Karush-Kuhn-Tucker

Na demonstração do Teorema de Fritz John não há uma condição que garanta que o multiplicador de Lagrange ($\bar{u}_0$) associado com a derivada da função objetivo seja diferente de zero. Se for zero, então as informações contidas na derivada da função objetivo são perdidas e consequentemente as condições de Fritz John serão de pequeno valor na busca da solução ótima. Para evitar este problema, define-se algumas condições de regularidade do problema. Essas condições são chamadas de qualificação. Usa-se condições de qualificação de Slater no caso da versão com tempo contínuo das condições de Karush-Kuhn-Tucker.

Sejam X um conjunto convexo de $\mathcal{L}_{\infty}^n[0, T]$, e g uma função convexa com relação a sua primeira variável sobre V por todo $[0, T]$. Então g satisfaz a condição de qualificação de Slater (em X) se existe $\hat{x} \in X$ tal que

$$g(\hat{x}(t), t) < 0 \quad \text{q. s. em } [0, T]. \quad (2.9)$$

Pelo Teorema(2.1.1) a condição de qualificação de Slater é equivalente a afirmação de que não existe $u \in \mathcal{L}_{\infty}[0, T]$, $u(t) \geq 0$, $u(t) \neq 0$ q. s. em $[0, T]$ tal que

$$\int_0^T u(t) g(x(t), t) dt \geq 0 \quad \text{para todo } x \in X,$$

esta condição é conhecida como condição de qualificação de Karlin.

O teorema a seguir é conhecido como Teorema de Karush-Kuhn-Tucker.



Teorema 2.2.4 *Seja $\bar{x} \in \mathbb{F}$ e suponha em adição com as hipóteses do Teorema (2.2.3), a função g é convexa com relação a sua primeira variável sobre V por todo $[0, T]$ e satisfaz a restrição de qualificação de Slater (2.9) em X . Se $\bar{x}$ é a solução ótima local do problema (P) então existe $\tilde{u} \in \mathcal{L}_\infty[0, T]$ tal que*

$$\int_0^T [\nabla f'(\bar{x}(t), t) + \tilde{u}(t) \nabla g'(\bar{x}(t), t)] h(t) dt = 0 \text{ para todo } h \in \mathcal{L}_\infty^n[0, T];$$

$$\tilde{u}(t) g(\bar{x}(t), t) = 0 \text{ q. s. em } [0, T];$$

$$\tilde{u}(t) \geq 0 \text{ q. s. em } [0, T].$$

Demonstração :

Pelo Teorema (2.2.3), existe $\bar{u}_0 \in \mathbb{R}$, $\bar{u} \in \mathcal{L}_\infty[0, T]$, tal que (2.3) - (2.6) vale. Se $\bar{u}_0 = 0$, então (2.3) é reduzido a

$$\int_0^T \bar{u}(t) \nabla g(\bar{x}(t), t) h(t) dt = 0 \text{ para todo } h \in \mathcal{L}_\infty^n[0, T]. \quad (2.10)$$

Por hipótese tem-se que g satisfaz a condição de qualificação de Slater (2.9) em X por todo $[0, T]$, então existe um $\hat{x} \in X$ tal que

$$g(\hat{x}(t), t) < 0 \text{ q. s. em } [0, T].$$

Pela convexidade de $g(\cdot, t)$ em V por todo $[0, T]$, tem-se

$$g(\hat{x}(t), t) - g(\bar{x}(t), t) \geq \nabla g(\bar{x}(t), t) [\hat{x}(t) - \bar{x}(t)] \text{ q. s. em } [0, T].$$

Multiplicando por $\bar{u}(t) \geq 0$ q. s. em $[0, T]$ em ambos os lados e integrando, obtem-se

$$\int_0^T \bar{u}(t) g(\hat{x}(t), t) dt - \int_0^T \bar{u}(t) g(\bar{x}(t), t) dt \geq \int_0^T \bar{u}(t) \nabla g(\bar{x}(t), t) [\hat{x}(t) - \bar{x}(t)] dt.$$

Portanto

$$\int_0^T \bar{u}(t) g(\hat{x}(t), t) dt \geq \int_0^T \bar{u}(t) \nabla g'(\bar{x}(t), t) [\hat{x}(t) - \bar{x}(t)] dt.$$

Considerando (2.10) com $h = \hat{x} - \bar{x}$, esta desigualdade mostra que

$$\int_0^T \bar{u}(t) g(\hat{x}(t), t) dt \geq 0$$

e pelo Teorema (2.1.1) implica que

$$g(\hat{x}(t), t) < 0 \text{ q. s. em } [0, T]$$

é impossível, contradizendo a restrição de qualificação de Slater (2.9).

Assim

$$\bar{u}_0 \neq 0.$$

Agora, dividindo (2.3) por $\bar{u}_0$ e tomando

$$\tilde{u}(t) = \frac{\bar{u}(t)}{\bar{u}_0}$$

obtem-se

$$\int_0^T [\nabla f'(\bar{x}(t), t) + \tilde{u}(t) \nabla g(\bar{x}(t), t)] h(t) dt = 0 \text{ para todo } h \in \mathcal{L}_\infty^n[0, T];$$

$$\tilde{u}(t) \geq 0 \text{ q. s. em } [0, T].$$

Agora, resta mostrar que

$$\tilde{u}(t)g(\bar{x}(t), t) = 0 \text{ q. s. em } [0, T].$$

De fato:

Sabe-se que

$$\tilde{u}(t) = \frac{\bar{u}(t)}{\bar{u}_0}$$

e considerando $\bar{u}(t)$ como foi definido no Teorema (2.2.3), tem-se:

(a) Se $t \in A(\bar{x})$ então $g(\bar{x}(t), t) = 0$. Assim, $\tilde{u}(t)g(\bar{x}(t), t) = 0$ q. s. em $[0, T]$.

(b) Se $t \notin A(\bar{x})$ então por definição $\bar{u}(t) = 0$. Assim, $\tilde{u}(t) = \frac{\bar{u}(t)}{\bar{u}_0} = \frac{0}{\bar{u}_0} = 0$.

Desta forma, tem-se que $\tilde{u}(t)g(\bar{x}(t), t) = 0$ q. s. em $[0, T]$. ■



Capítulo 3

Condições Suficientes para Otimalidade

Neste capítulo, desenvolve-se condições suficientes de otimalidade de Fritz-John e Karush-Kuhn-Tucker em programação não linear com tempo contínuo para problemas com restrições de desigualdades e igualdades não lineares com suposição de alguma convexidade generalizada.

3.1 Convexidade Generalizada

Vê-se algumas definições e resultados básicos de convexidade generalizada definidas por Ponstein [18], que posteriormente serão necessários para o desenvolvimento de condições suficientes de otimalidade. Seja $f(x)$ uma função real, contínua e diferenciável que tem derivadas parciais contínuas com relação ao vetor $x = (x_1, x_2, \dots, x_n)'$, pertencente à uma região convexa $\mathcal{R}$ do espaço Euclidiano $\mathbb{R}^n$. Suponha que λ e μ são escalares satisfazendo $\lambda, \mu \geq 0$ e $\lambda + \mu = 1$.

Definição 3.1.1 A função $f(x)$ é convexa se

$$(1a) \quad f(\lambda x_1 + \mu x_2) \leq \lambda f(x_1) + \mu f(x_2)$$

ou, equivalentemente, se

$$(1b) \quad f(x_2) \geq f(x_1) + \nabla f(x_1)(x_2 - x_1).$$



Vai-se provar a equivalência ($\sim$), ou seja, $(1a) \sim (1b)$.

De fato:

$(1a) \Rightarrow (1b)$

Suponha que f é convexa, então para todo $\lambda, \mu \geq 0$ e $\lambda + \mu = 1$ tem-se

$$f(\lambda x_1 + \mu x_2) \leq \lambda f(x_1) + \mu f(x_2).$$

Como $\lambda + \mu = 1$, então $\lambda = 1 - \mu$, $0 \leq \mu \leq 1$. Assim

$$f((1 - \mu)x_1 + \mu x_2) \leq (1 - \mu)f(x_1) + \mu f(x_2).$$

Para $0 < \mu \leq 1$ tem-se

$$\frac{f(x_1 + \mu(x_2 - x_1)) - f(x_1)}{\mu} \leq f(x_2) - f(x_1).$$

Aplicando o limite com $\mu \rightarrow 0$ pela direita, segue da definição de derivada direcional que

$$\nabla f(x_1)(x_2 - x_1) \leq f(x_2) - f(x_1).$$

Portanto

$$f(x_2) \geq f(x_1) + \nabla f(x_1)(x_2 - x_1).$$

$(1b) \Rightarrow (1a)$

Seja $z = \mu x_2 + (1 - \mu)x_1$, $z \in S$, $0 \leq \mu \leq 1$. Por hipótese tem-se

$$f(x_2) \geq f(z) + \nabla f(z)(x_2 - z) = f(z) + (1 - \mu)\nabla f(z)(x_2 - x_1), \quad (3.1)$$

$$f(x_1) \geq f(z) + \nabla f(z)(x_1 - z) = f(z) - \mu\nabla f(z)(x_2 - x_1). \quad (3.2)$$

Multiplicando (3.1) por $\mu \geq 0$, (3.2) por $(1 - \mu) \geq 0$ e somando, obtem-se

$$\mu f(x_2) + (1 - \mu)f(x_1) \geq \mu f(z) + (1 - \mu)f(z) = f(z) = f(\mu x_2 + (1 - \mu)x_1).$$

Assim

$$f(\mu x_2 + (1 - \mu)x_1) \leq \mu f(x_2) + (1 - \mu)f(x_1).$$

Como $\lambda = 1 - \mu$, tem-se

$$f(\lambda x_1 + \mu x_2) \leq \lambda f(x_1) + \mu f(x_2). \quad \blacksquare$$



Definição 3.1.2 A função $f(x)$ é estritamente convexa se

$$(2a) \quad f(\lambda x_1 + \mu x_2) < \lambda f(x_1) + \mu f(x_2)$$

onde $x_1 \neq x_2$, ou equivalentemente, se

$$(2b) \quad f(x_2) > f(x_1) + \nabla f(x_1)(x_2 - x_1)$$

com $x_1 \neq x_2$.

Vai-se provar que $(2a) \sim (2b)$.

De fato:

$$(2a) \Rightarrow (2b)$$

Da definição (3.1.1), existe $\nabla f(x_1)$ tal que

$$f(x_2) \geq f(x_1) + \nabla f(x_1)(x_2 - x_1), \quad (3.3)$$

para todo $x_1, x_2 \in S$.

Suponha que existe $\hat{x}_2 \neq x_1$ tal que

$$f(\hat{x}_2) = f(x_1) + \nabla f(x_1)(\hat{x}_2 - x_1).$$

Pela convexidade estrita de f em $0 \leq \lambda \leq 1$, tem-se

$$\begin{aligned} f(\lambda x_1 + (1 - \lambda)\hat{x}_2) &< \lambda f(x_1) + (1 - \lambda)f(\hat{x}_2) \\ &= \lambda f(x_1) + (1 - \lambda)f(x_1) + (1 - \lambda)\nabla f(x_1)(\hat{x}_2 - x_1) \\ &= f(x_1) + (1 - \lambda)\nabla f(x_1)(\hat{x}_2 - x_1). \end{aligned} \quad (3.4)$$

Seja $x_2 = \lambda x_1 + (1 - \lambda)\hat{x}_2$. Por (3.3) tem-se

$$f(\lambda x_1 + (1 - \lambda)\hat{x}_2) \geq f(x_1) + \nabla f(x_1)(\lambda x_1 + (1 - \lambda)\hat{x}_2 - x_1).$$

Assim

$$f(\lambda x_1 + (1 - \lambda)\hat{x}_2) \geq f(x_1) + (1 - \lambda)\nabla f(x_1)(\hat{x}_2 - x_1),$$

contradizendo (3.4). Então

$$f(x_2) > f(x_1) + \nabla f(x_1)(x_2 - x_1).$$

(2b) $\Rightarrow$ (2a)

Seja $z = \mu x_2 + (1 - \mu)x_1$, $z \in S$. Por hipótese tem-se

$$f(x_2) > f(z) + \nabla f(z)(x_2 - z) = f(z) + (1 - \mu)\nabla f(z)(x_2 - x_1), \quad (3.5)$$

$$f(x_1) > f(z) + \nabla f(z)(x_1 - z) = f(z) - \mu\nabla f(z)(x_2 - x_1). \quad (3.6)$$

Multiplicando (3.5) por $\mu \geq 0$, (3.6) por $(1 - \mu) \geq 0$ e somando, obtém-se

$$\mu f(x_2) + (1 - \mu)f(x_1) > \mu f(z) + (1 - \mu)f(z) = f(z) = f(\mu x_2 + (1 - \mu)x_1).$$

Assim

$$f(\mu x_2 + (1 - \mu)x_1) < \mu f(x_2) + (1 - \mu)f(x_1).$$

Como $\lambda = 1 - \mu$, tem-se

$$f(\lambda x_1 + \mu x_2) < \lambda f(x_1) + \mu f(x_2). \quad \blacksquare$$

Seja S um conjunto convexo em $\mathbb{R}^n$ e $f : S \rightarrow \mathbb{R}^n$ uma função convexa. Então o conjunto de nível

$$S_c = \{x \in S : f(x) \leq c\},$$

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

De fato:

Sejam $x_1, x_2 \in S_c$, então $x_1, x_2 \in S$, $f(x_1) \leq c$ e $f(x_2) \leq c$.

Sejam $\lambda \in (0, 1)$ e $x = \lambda x_1 + (1 - \lambda)x_2$. Pela convexidade de S , $x \in S$. Além disso, pela convexidade de f , tem-se

$$f(x) \leq \lambda f(x_1) + (1 - \lambda)f(x_2) \leq \lambda c + (1 - \lambda)c = c.$$

Assim, $x \in S_c$. Portanto, S_c é convexo. ■

Entretanto o fato do conjunto de nível ser convexo não implica que a função seja convexa.

Por exemplo, a função $f(x) = x^3$ não é convexa em $\mathbb{R}$, embora que o conjunto $S_c = \{x \in \mathbb{R} : x^3 \leq c\}$, seja convexo.

Para funções quase-convexas tem-se quatro definições equivalentes dadas a seguir.



Definição 3.1.3 A função $f(x)$ é quase-convexa se

$$(3a) \quad S_c = \{x/f(x) \leq c\}$$

é um conjunto convexo para todo $c \in \mathbb{R}$, ou se

$$(3b) \quad S_c = \{x/f(x) < c\}$$

é um conjunto convexo para todo $c \in \mathbb{R}$, ou se

$$(3c) \quad f(x_2) \leq f(x_1) \text{ então } f(\lambda x_1 + \mu x_2) \leq f(x_1),$$

ou se

$$(3d) \quad f(x_2) \leq f(x_1) \text{ então } (x_2 - x_1) \nabla f(x_1) \leq 0,$$

Vai-se provar que $(3a) \sim (3b)$

De fato:

$$(3a) \Rightarrow (3b)$$

A condição $(3b)$ segue imediatamente de $(3a)$.

$$(3b) \Rightarrow (3a)$$

Para qualquer $c \in \mathbb{R}$, seja $f(x_1) \leq c$ e $f(x_2) \leq c$. Por $(3b)$, tem-se

$$f(\lambda x_1 + \mu x_2) < c + \epsilon, \quad \forall \epsilon > 0.$$

Assim

$$f(\lambda x_1 + \mu x_2) \leq c.$$

Portanto, S_c é um conjunto convexo. ■

Vai-se provar que $(3a) \sim (3c)$

De fato:

$$(3a) \Rightarrow (3c)$$

Por $(3a)$, tem-se $f(x_1) \leq c$.

Seja $f(x_2) \leq f(x_1)$, então

$$f(\lambda x_1 + \mu x_2) \leq f(x_1) \leq c.$$

$$(3c) \Rightarrow (3a)$$

Sejam $x_1, x_2 \in S_c$. Por hipótese $f(x_2) \leq f(x_1)$, então

$$f(\lambda x_1 + \mu x_2) \leq f(x_1) \leq c.$$

Assim, $\lambda x_1 + \mu x_2 \in S_c$. Portanto, S_c é um conjunto convexo. ■

Vai-se provar que $(3c) \sim (3d)$

De fato:

$(3c) \Rightarrow (3d)$

Seja $\lambda = 1 - \mu$, então $\lambda x_1 + \mu x_2 = x_1 + \mu(x_2 - x_1)$. Suponha que $f(x_2) \leq f(x_1)$. Então por $(3c)$, tem-se $f(\lambda x_1 + \mu x_2) \leq f(x_1)$, ou seja

$$f(x_1 + \mu(x_2 - x_1)) \leq f(x_1).$$

Assim

$$\frac{f(x_1 + \mu(x_2 - x_1)) - f(x_1)}{\lambda} \leq 0.$$

O item $(3d)$ segue da definição de derivada direcional a aplicar o limite com λ tendendo a zero na desigualdade acima.

$(3d) \Rightarrow (3c)$

De fato:

Seja $f(x_2) \leq f(x_1)$ e suponha que

$$f(x_3) > f(x_1). \quad (3.7)$$

para alguma combinação convexa x_3 de x_1 e x_2 . Por $(3d)$, tem-se

$$(x_1 - x_3) \nabla f(x_3) \leq 0 \quad (3.8)$$

e

$$(x_2 - x_3) \nabla f(x_3) \leq 0. \quad (3.9)$$

De (3.8) e (3.9), segue que

$$(x_1 - x_2) \nabla f(x_3) \leq 0$$

e

$$(x_2 - x_1) \nabla f(x_3) \leq 0,$$

o que significa que

$$(x_2 - x_1) \nabla f(x_3) = 0.$$

Pode-se escrever x_3 como qualquer combinação x de x_1 e x_2 tal que

$$f(x_1) < f(x).$$

Agora, seja T definido por

$$T = \{x : f(x) \leq f(x_1), x = \lambda x_1 + \mu x_3\}.$$

Então existe $x_0 \in T$ mais próximo de x_3 tal que

$$f(x_3) - f(x_0) = (x_3 - x_0) \nabla f(x),$$

para alguma combinação convexa x de x_0 e x_3 .

Já que $f(x_1) < f(x)$ e x é uma combinação convexa de x_1 e x_2 tem-se

$$(x_2 - x_1) \nabla f(x) = 0.$$

Assim

$$(x_3 - x_0) \nabla f(x) = 0,$$

o que implica

$$f(x_1) \geq f(x_0) = f(x_3),$$

contradizendo (3.7). Então

$$f(x_2) \leq f(x_1) \text{ então } f(\lambda x_1 + \mu x_2) \leq f(x_1).$$

Definição 3.1.4 A função $f(x)$ é estritamente quase-convexa se

$$(4) \quad f(x_2) < f(x_1) \text{ então } f(\lambda x_1 + \mu x_2) < f(x_1),$$

com $\lambda, \mu > 0$.

Definição 3.1.5 A função $f(x)$ é pseudo-convexa se

$$(5a) \quad f(x_2) < f(x_1) \text{ então } (x_2 - x_1) \nabla f(x_1) < 0.$$

Isto é equivalente a dizer que

$$(5b) \quad (x_2 - x_1) \nabla f(x_1) \geq 0 \text{ então } f(x_2) \geq f(x_1).$$

Definição 3.1.6 A função $f(x)$ é estritamente pseudo-convexa se

$$(6) \qquad f(x_2) \leq f(x_1) \text{ então } (x_2 - x_1)\nabla f(x_1) < 0$$

com $x_1 \neq x_2$.

A tabela a seguir mostra a relação entre os diferentes tipos de convexidade. Por conveniência, usa-se as seguintes notações :

SC estritamente convexa,

C convexa,

SPC estritamente pseudo-convexa,

PC pseudo-convexa,

SQC estritamente quase-convexa,

QC quase-convexa,

+ a função satisfaz a condição correspondente na tabela,

– a função não satisfaz a condição correspondente na tabela.

função	região $\mathcal{R}$	SC	C	SPC	PC	SQC	QC
$f(x) = \begin{cases} 0 \\ -(x-1)^2 \end{cases}$	$0 \leq x \leq 1$	–	–	–	–	–	+
	$1 \leq x \leq 2$	–	–	–	–	–	–
$f(x,y) = -x^2$	$0 \leq x,y \leq 1$	–	–	–	–	+	+
$f(x,y) = -x^2 - x$	$0 \leq x,y \leq 1$	–	–	–	+	+	+
$f(x) = 0$	$0 \leq x \leq 1$	–	+	–	+	+	+
$f(x) = -x$	$0 \leq x \leq 1$	–	+	+	+	+	+
$f(x) = x^2$	$0 \leq x \leq 1$	+	+	+	+	+	+
$f(x) = -x^2 - x$	$0 \leq x \leq 1$	–	–	+	+	+	+
$f(x) = -x^2$	$0 \leq x \leq 1$	–	–	–	–	+	+

3.2 Condições suficientes de otimalidade de Fritz-John

Apenas nesta seção, acrescenta-se ao Problema (P), a restrição de igualdade $h(x(t),t) = 0$ q. s. em $[0,T]$. Os conjuntos V e F devem ser adaptados adequadamente



para incluir também a restrição de igualdade. Isto é :

$$IF = \{x \in X : g(x(t), t) \leq 0, h(x(t), t) = 0 \text{ q. s. em } [0, T]\},$$

$$V = \{x(t) \in \mathbb{R}^n : x \in IF\}.$$

obviamente, quando a restrição de igualdade h está ausente, volta-se ao problema original (P), com IF e V definidos como antes.

Teorema 3.2.1 *Sejam X convexo, $\bar{x} \in IF$, e suponha que f, g e h são continuamente diferenciáveis com relação a sua primeira variável em $\bar{x}(t)$ por todo $[0, T]$. Além disso, suponha que ϕ é pseudo-convexa em $\bar{x}$ (com relação a IF), e que g e h são estritamente pseudo-convexas em relação a suas primeiras variáveis em $\bar{x}(t)$ (com relação a V) por todo $[0, T]$. Se existe $u_0 \in \mathbb{R}$, $u \in \mathcal{L}_\infty[0, T]$ e $\bar{\mu} \in \mathcal{L}_\infty[0, T]$ tal que*

$$\int_0^T [u_0 \nabla f'(\bar{x}(t), t) + u(t) \nabla g(\bar{x}(t), t) + \bar{\mu}(t) \nabla h(\bar{x}(t), t)] y(t) dt = 0 \text{ para todo } y \in \mathcal{L}_\infty^n[0, T]; \quad (3.10)$$

$$u(t) = 0 \text{ q. s. em } [0, T], t \in [0, T] \setminus A(\bar{x}); \quad (3.11)$$

$$u_0 \geq 0, u(t) \geq 0, \bar{\mu}(t) \geq 0 \text{ q. s. em } [0, T]; \quad (3.12)$$

$$(u_0, u(t), \bar{\mu}(t)) \neq 0 \text{ q. s. em } [0, T]; \quad (3.13)$$

então $\bar{x}$ é uma solução ótima global do problema (P).

Demonstração :

Pelo Teorema (2.1.1), o sistema de desigualdade linear

$$\begin{cases} \int_0^T \nabla f'(\bar{x}(t), t) y(t) dt < 0; \\ \nabla g'(\bar{x}(t), t) y(t) < 0 \text{ q. s. em } A(\bar{x}); \\ \nabla h'(\bar{x}(t), t) y(t) < 0 \text{ q. s. em } [0, T]; \end{cases} \quad (3.14)$$

não tem solução $y \in \mathcal{L}_\infty^n[0, T]$. Consequentemente, o sistema

$$\begin{cases} \int_0^T f(x(t), t) dt - \int_0^T f(\bar{x}(t), t) dt < 0; \\ g(x(t), t) - g(\bar{x}(t), t) \leq 0 \text{ q. s. em } A(\bar{x}); \\ h(x(t), t) - h(\bar{x}(t), t) = 0 \text{ q. s. em } [0, T]; \end{cases} \quad (3.15)$$

não tem solução $x \in X$, pois se existe uma solução $\hat{x} \in X$, $\hat{x} \neq \bar{x}$, então pela pseudoconvexidade de $f(., t)$ tem-se:

$$f(\hat{x}(t), t) < f(\bar{x}(t), t) \implies \nabla f'(\bar{x}(t), t)(\hat{x}(t) - \bar{x}(t)) < 0,$$

e assim

$$\int_0^T \nabla f'(\bar{x}(t), t)(\hat{x} - \bar{x}) dt < 0. \quad (3.16)$$

Analogamente, da estrita pseudoconvexidade de $g(., t)$ em $\bar{x}(t)$ para $t \in A(\bar{x})$ e $h(., t)$ em $\bar{x}(t)$ por todo $[0, T]$ e das relações

$$g(\hat{x}(t), t) \leq g(\bar{x}(t), t) \text{ q. s. em } A(\bar{x})$$

e

$$h(\hat{x}(t), t) = h(\bar{x}(t), t) \text{ q. s. em } [0, T],$$

tem-se que

$$\nabla g(\bar{x}(t), t)(\hat{x}(t) - \bar{x}(t)) < 0 \text{ q. s. em } A(\bar{x}) \quad (3.17)$$

e

$$\nabla h(\bar{x}(t), t)(\hat{x}(t) - \bar{x}(t)) < 0 \text{ q. s. em } [0, T]. \quad (3.18)$$

Claramente, (3.16) - (3.18) contradiz (3.14) com $y = \hat{x} - \bar{x}$. Portanto, o sistema (3.15) não pode ter solução $x \in X$, para $t \in A(\bar{x})$. Como $u(t) = 0$ q. s. em $[0, T]$, para $t \in [0, T] \setminus A(\bar{x})$ pode-se concluir que o sistema (3.15) não tem solução $x \in X$, para $\forall t \in [0, T]$. Assim, $\bar{x}$ é solução ótima global do problema (P). ■

Corolário 3.2.1 *Sejam X convexo, $\bar{x} \in F$ e suponha que f , g e h são continuamente diferenciáveis com relação a sua primeira variável em $\bar{x}(t)$ por todo $[0, T]$. Além disso, suponha que ϕ é convexa em $\bar{x}(t)$ (com relação a F) e g e h são estritamente convexas com relação a sua primeira variável em $\bar{x}(t)$ (com relação a V) por todo $[0, T]$. Se existe $u_0 \in \mathbb{R}$, $u \in \mathcal{L}_\infty[0, T]$, $\bar{\mu} \in \mathcal{L}_\infty[0, T]$ tal que as condições (3.10) - (3.13) são satisfeitas, então $\bar{x}$ é uma solução ótima global do problema (P).*



3.3 Condições suficientes de otimalidade de Karush-Kuhn-Tucker

Neste teorema, as condições suficientes de otimalidade de Karush-Kuhn-Tucker diferem das condições suficientes de otimalidade de Fritz-John apenas pela ausência do multiplicador escalar u_0 que agora é tomado como sendo um. Também toma-se o problema (P) original, ou seja, sem a restrição de igualdade h .

Teorema 3.3.1 *Seja $\bar{x} \in \mathbb{F}$ e suponha que f e g são continuamente diferenciáveis com relação a sua primeira variável em $\bar{x}(t)$ por todo $[0, T]$. Além disso, suponha que ϕ é pseudo-convexa em $\bar{x}$ (com relação a $\mathbb{F}$) e que g é quase-convexa com relação a sua primeira variável em $\bar{x}(t)$ por todo $[0, T]$ (com relação a V). Se existe um $u \in \mathcal{L}_\infty[0, T]$ tal que $(\bar{x}, u)$ satisfaz as condições :*

$$\int_0^T [\nabla f'(\bar{x}(t), t) + u(t) \nabla g'(\bar{x}(t), t)] y(t) dt = 0 \text{ para todo } y \in \mathcal{L}_\infty^n[0, T]; \quad (3.19)$$

$$u(t) = 0 \text{ q. s. em } [0, T], \quad t \in [0, T] \setminus A(\bar{x}); \quad (3.20)$$

$$u(t) \geq 0 \text{ q. s. em } [0, T]; \quad (3.21)$$

então $\bar{x}$ é uma solução ótima global do problema (P).

Demonstração :

Já que para qualquer $x \in \mathbb{F}$

$$g(x(t), t) \leq 0 = g(\bar{x}(t), t) \text{ q. s. em } A(\bar{x}).$$

Pela quase-convexidade dada por hipótese, tem-se:

$$g(x(t), t) \leq g(\bar{x}(t), t) \text{ q. s. em } A(\bar{x}) \implies \nabla g'(\bar{x}(t), t)(x(t) - \bar{x}(t)) \leq 0 \text{ q. s. em } A(\bar{x}).$$

Como $u(t) = 0$ q. s. em $[0, T]$, para $t \in [0, T] \setminus A(\bar{x})$, pode-se concluir que

$$\int_0^T u(t) \nabla g'(\bar{x}(t), t)(x(t) - \bar{x}(t)) dt \leq 0 \text{ para todo } x \in \mathbb{F}.$$

Pela desigualdade acima e (3.19) com $y = x - \bar{x}$ implica que

$$\int_0^T \nabla f'(\bar{x}(t), t)(x(t) - \bar{x}(t))dt \geq 0 \text{ para todo } x \in IF.$$

Já que ϕ é pseudo-convexa em $\bar{x}$ (com relação a IF), a desigualdade anterior implica que

$$\int_0^T f(x(t), t)dt \geq \int_0^T f(\bar{x}(t), t)dt \text{ para todo } x \in IF.$$

Assim, $\bar{x}$ é uma solução ótima global do problema (P). ■

Teorema 3.3.2 *Seja $\bar{x} \in IF$ e suponha que f e g são continuamente diferenciáveis com relação a sua primeira variável em $\bar{x}(t)$ por todo $[0, T]$. Além disso, suponha que ϕ é pseudo-convexa em $\bar{x}$ (com relação a IF), se existe $u \in \mathcal{L}_\infty[0, T]$ tal que $(\bar{x}, u)$ satisfaz (3.19) - (3.21) e se a função $\psi(\cdot, u) : \mathcal{L}_\infty^n[0, T] \rightarrow \mathbb{R}$ definida por $\psi(x, u) = \int_0^T u(t)g(x(t), t)dt$ é quase-convexa em $\bar{x}$ (com respeito a IF), então $\bar{x}$ é solução ótima global do problema (P).*

Demonstração :

Para qualquer $x \in IF$, pelas condições (3.20) e (3.21) tem-se

$$\psi(x, u) \leq 0 = \psi(\bar{x}, u).$$

Pela quaseconvexidade de ψ em $\bar{x}$ tem-se

$$D\psi(\bar{x}, u)(x - \bar{x}) \leq 0 \text{ para todo } x \in IF.$$

Assim, tem-se

$$\int_0^T u(t) \nabla g(\bar{x}(t), t)(x(t) - \bar{x}(t))dt \leq 0 \text{ para todo } x \in IF.$$

Pela desigualdade acima e (3.19) com $y = x - \bar{x}$, tem-se

$$\int_0^T \nabla f'(\bar{x}(t), t)(x(t) - \bar{x}(t))dt \geq 0 \text{ para todo } x \in IF.$$

Pela pseudoconvexidade de $\phi(x) = \int_0^T f(x(t), t)dt$, tem-se

$$\int_0^T f(x(t), t)dt \geq \int_0^T f(\bar{x}(t), t)dt.$$

Assim

$$\phi(x) \geq \phi(\bar{x}) \text{ para todo } x \in \mathbb{F}.$$

Portanto, $\bar{x}$ é solução ótima global do problema (P). ■

Teorema 3.3.3 *Seja $\bar{x} \in \mathbb{F}$. Suponha que f e g são continuamente diferenciáveis com relação a sua primeira variável em $\bar{x}(t)$ por todo $[0, T]$. Se existe $u \in \mathcal{L}_\infty[0, T]$ tal que $(\bar{x}, u)$ satisfaz (3.19) - (3.21) e se a função de Lagrange $L(., u) : \mathcal{L}_\infty^n[0, T] \rightarrow \mathbb{R}$ definida por*

$$L(x, u) = \int_0^T [f(x(t), t) + u(t)g(x(t), t)]dt$$

é pseudo-convexa em $\bar{x}$ (com relação a $\mathbb{F}$), então $\bar{x}$ é uma solução ótima global do problema (P).

Demonstração :

A condição (3.19) com $y = x - \bar{x}$ é equivalente a

$$DL(\bar{x}, u)(x - \bar{x}) = 0 \text{ para todo } x \in \mathbb{F}.$$

Pela pseudoconvexidade de $L(\bar{x}, u)$ tem-se

$$L(\bar{x}, u) \leq L(x, u) \text{ para todo } x \in \mathbb{F}.$$

$$\int_0^T [f(\bar{x}(t), t) + u(t)g(\bar{x}(t), t)]dt \leq \int_0^T [f(x(t), t) + u(t)g(x(t), t)]dt \text{ para todo } x \in \mathbb{F}.$$

Como

$$g(\bar{x}(t), t) = 0$$

e

$$u(t) \geq 0 \text{ e } g(x(t), t) \leq 0 \implies u(t)g(x(t), t) \leq 0,$$

tem-se

$$\int_0^T f(\bar{x}(t), t)dt \leq \int_0^T f(x(t), t)dt \text{ para todo } x \in \mathbb{F}.$$

Portanto

$$\phi(\bar{x}) \leq \phi(x) \text{ para todo } x \in \mathbb{F}.$$

Assim, $\bar{x}$ é solução ótima do problema (P). ■

Até agora, viu-se condições suficientes de primeira ordem. Agora, no teorema a seguir vê-se uma condição suficiente de segunda ordem.



Teorema 3.3.4 *Sejam X convexo e $\bar{x} \in \mathbb{F}$. Suponha que f e g são duas vezes continuamente diferenciáveis com relação a sua primeira variável em $\bar{x}(t)$ por todo $[0, T]$. Além disso, suponha que existe um $u \in \mathcal{L}_\infty[0, T]$ tal que $(\bar{x}, u)$ satisfaz (3.19) - (3.21) e que existe uma constante $c > 0$ tal que*

$$D^2L(\bar{x}, u)(y, y) \geq c\|y\|_\infty^2 \text{ para todo } y \in \mathcal{L}_\infty^n[0, T]. \quad (3.22)$$

Se $L(\cdot, u)$ é quase-convexa em $\bar{x}$ (com relação a $\mathbb{F}$), então $\bar{x}$ é a solução ótima global do problema (P).

Demonstração :

Desde que $L(\cdot, u)$ é duas vezes continuamente diferenciável em $\bar{x}$ e pelo desenvolvimento de Taylor, pode-se escrever

$$L(x, u) - L(\bar{x}, u) = DL(\bar{x}, u)(x - \bar{x}) + \frac{1}{2}D^2L(\bar{x}, u)(x - \bar{x}, x - \bar{x}) + \|x - \bar{x}\|_\infty^2 \rho(\bar{x}; x - \bar{x}) \quad (3.23)$$

onde $x \in \mathbb{F}$ e $\rho(\bar{x}; x - \bar{x}) \rightarrow 0$ quando $x \rightarrow \bar{x}$.

Agora, usando (3.19) com $y = x - \bar{x}$ e (3.22), (3.23) pode ser escrita como:

$$L(x, u) - L(\bar{x}, u) = DL(\bar{x}, u)(x - \bar{x}) + \frac{1}{2}D^2L(\bar{x}, u)(x - \bar{x}, x - \bar{x}) + \|x - \bar{x}\|_\infty^2 \rho(\bar{x}; x - \bar{x}).$$

Assim

$$L(x, u) - L(\bar{x}, u) \geq \|x - \bar{x}\|_\infty^2 \left(\frac{c}{2} + \rho(\bar{x}; x - \bar{x}) \right) > 0, \forall x \text{ próximo de } \bar{x}, \text{ com } x \neq \bar{x}.$$

Portanto

$$L(x, u) > L(\bar{x}, u), \forall x \text{ próximo de } \bar{x}, x \neq \bar{x}.$$

Assim, $\bar{x}$ é um mínimo local estrito de $L(\cdot, u)$ sobre $\mathbb{F}$. Já que $L(\cdot, u)$ é quase-convexa em $\bar{x}$ (com respeito a $\mathbb{F}$), pelo Teorema (3.3.3), $\bar{x}$ é o mínimo global de $L(\cdot, u)$ sobre $\mathbb{F}$.

Assim

$$L(\bar{x}, u) < L(x, u) \text{ para todo } x \in \mathbb{F}.$$

Como viu-se nos Teoremas (3.3.2) e (3.3.3), a factibilidade de x implica que

$$\phi(\bar{x}) < \phi(x) \text{ para todo } x \in \mathbb{F}.$$

Assim, $\bar{x}$ é a solução ótima global do problema (P). ■

Teorema 3.3.5 *Seja X convexo e suponha que f e g são continuamente diferenciáveis com relação a sua primeira variável em V por todo $[0, T]$. Além disso, suponha que g é quase-convexa com relação a sua primeira variável em V por todo $[0, T]$, e que ϕ é quase-convexa em X . Suponha ainda que $\bar{x} \in \mathbb{F}$ e que existe $u \in \mathcal{L}_\infty[0, T]$ tal que $(\bar{x}, u)$ satisfaz (3.19) - (3.21). Se uma das alternativas abaixo for verdadeira:*

(a) *existe $\hat{x} \in X$ tal que*

$$\int_0^T \nabla f'(\bar{x}(t), t)(\hat{x}(t) - \bar{x}(t))dt > 0; \quad (3.24)$$

ou

(b) *f é duas vezes continuamente diferenciável com relação a sua primeira variável em V e $\nabla f(\bar{x}(t), t) \neq 0$ q. s. em $[0, T]$;*

então $\bar{x}$ é uma solução ótima global do problema (P).

Demonstração :

Primeiro note que pela demonstração do Teorema (3.3.1), pode-se afirmar que para qualquer $x \in \mathbb{F}$ a seguinte desigualdade vale:

$$\int_0^T \nabla f'(\bar{x}(t), t)(x(t) - \bar{x}(t))dt \geq 0. \quad (3.25)$$

(a) Sabe-se que V é um subconjunto convexo aberto em $\mathbb{R}^n$ contendo o conjunto: $\{x(t) : x \in X, t \in [0, T]\}$. Seja $x(t) \in V$ e $\hat{x}(t) \in V$ então a combinação convexa $(1 - r)x(t) + r\hat{x}(t)$ também pertence a V para todo $0 < r < 1$. Assim

$$\begin{aligned} \int_0^T \nabla f(\bar{x}(t), t)[(1 - r)x(t) + r\hat{x}(t) - \bar{x}(t)]dt &= \int_0^T \nabla f(\bar{x}(t), t)[x(t) - rx(t) + r\hat{x}(t) - \bar{x}(t)]dt \\ &= \int_0^T \nabla f(\bar{x}(t), t)[(x(t) - \bar{x}(t)) + r(\hat{x}(t) - x(t))]dt \\ &= \int_0^T \nabla f(\bar{x}(t), t)[x(t) - \bar{x}(t)]dt + r \int_0^T \nabla f(\bar{x}(t), t)[\hat{x}(t) - x(t)]dt. \end{aligned}$$

Por (3.24), (3.25) e $0 < r < 1$ tem-se

$$\int_0^T \nabla f(\bar{x}(t), t)[(1 - r)x(t) + r\hat{x}(t) - \bar{x}(t)]dt > 0 \text{ para todo } 0 < r < 1.$$

Pela quase-convexidade de ϕ em X , a desigualdade acima implica que

$$\int_0^T f(\bar{x}(t), t) dt < \int_0^T f((1-r)x(t) + r\hat{x}(t), t) dt \quad \text{para todo } 0 < r < 1. \quad (3.26)$$

Já que $f(., t)$ é contínua em V por todo $[0, T]$, aplicando o limite com $r \rightarrow 0$ tem-se:

$$\lim_{r \rightarrow 0} \int_0^T f(\bar{x}(t), t) dt \leq \lim_{r \rightarrow 0} \int_0^T f((1-r)x(t) + r\hat{x}(t), t) dt. \quad (3.27)$$

Como (3.26) vale para todo $r \in (0, 1)$, toma-se em particular uma sequência $r_n \rightarrow 0$.

Seja:

$$h_n(x, t) = f((1-r_n)x(t) + r_n\hat{x}(t), t) \rightarrow h(x, t) = f(x(t), t), \quad \text{quando } r_n \rightarrow 0.$$

Assim

$$h_n(x, t) \rightarrow h(x, t), \quad \text{quando } r_n \rightarrow 0.$$

Aplicando o Teorema da convergência dominada de Lebesgue (1.1.1) tem-se

$$\lim_{r \rightarrow 0} \int_0^T h_n dt = \int_0^T h dt.$$

$$\lim_{r \rightarrow 0} \int_0^T f((1-r)x(t) + r\hat{x}(t), t) dt = \int_0^T f(x(t), t) dt.$$

Assim (3.27) torna-se

$$\int_0^T f(\bar{x}(t), t) dt \leq \int_0^T f(x(t), t) dt, \quad \text{para todo } x \in \mathcal{F}.$$

Portanto, $\bar{x}$ é uma solução ótima global do problema (P).

(b) Para eliminar o caso (a), suponha que

$$\int_0^T \nabla f'(\bar{x}(t), t)(x(t) - \bar{x}(t)) dt \leq 0 \quad \text{para todo } x \in X.$$

Esta desigualdade e (3.25) implicam que

$$\int_0^T \nabla f'(\bar{x}(t), t)(x(t) - \bar{x}(t)) dt = 0 \quad \text{para todo } x \in \mathcal{F}. \quad (3.28)$$

Já que por hipótese

$$\nabla f'(\bar{x}(t), t)(z(t) - \bar{x}(t)) \neq 0 \quad \text{q. s. em } [0, T], \quad \text{para todo } z \in X,$$

$$z(t) \neq \bar{x}(t) \text{ q. s. em } [0, T],$$

tem-se

$$\int_0^T \nabla f'(\bar{x}(t), t)(z(t) - \bar{x}(t))dt < 0 \text{ para todo } z \in X \setminus \{\bar{x}\}. \quad (3.29)$$

Seja $x \in \mathbb{F}$, $x \neq \bar{x}$, e para qualquer $z \in X$, define-se $\rho = x - \bar{x}$ e $q = z - \bar{x}$. Assim, (3.28) e (3.29) podem ser escritos como

$$\int_0^T \nabla f'(\bar{x}(t), t)\rho(t)dt = 0; \quad (3.30)$$

$$\int_0^T \nabla f'(\bar{x}(t), t)q(t)dt < 0. \quad (3.31)$$

Quer-se provar que

$$\int_0^T f(\bar{x}(t), t)dt \leq \int_0^T f(x(t), t)dt \text{ para todo } x \in \mathbb{F}.$$

Suponha o contrário, que

$$\int_0^T f(\bar{x}(t), t)dt > \int_0^T f(x(t), t)dt \text{ para algum } x \in \mathbb{F}.$$

Já que ϕ é quase-convexa em X , a última desigualdade implica que

$$\int_0^T f((1-r)\bar{x}(t) + rx(t), t)dt \leq \int_0^T f(\bar{x}(t), t)dt \text{ para todo } 0 \leq r \leq 1.$$

Pela continuidade de $f(., t)$ em V por todo $[0, T]$, pode-se achar o maior valor de $0 \leq r^* \leq 1$ tal que

$$\int_0^T f(\bar{x}(t) + r^*p(t), t)dt = \int_0^T f(\bar{x}(t), t)dt. \quad (3.32)$$

De (3.31) e da definição de r^* , pode-se achar de forma apropriada duas seqüências de números reais $\{r_n\}$, $\{s_n\}$ tal que $r_n > r^*$, $s_n > 0$, $r_n \rightarrow r^*$, $s_n \rightarrow 0$ quando $n \rightarrow \infty$ e

$$\int_0^T f(\bar{x}(t) + r_np(t), t)dt = \int_0^T f(\bar{x}(t) + s_nq(t), t)dt. \quad (3.33)$$

Tem-se dois casos:

(I) Quando $r^* = 0$.

Já que ϕ é quase-convexa em X , (3.33) implica que

$$\int_0^T \nabla f'(\bar{x}(t) + s_nq(t), t)(\bar{x}(t) + r_np(t) - \bar{x}(t) - s_nq(t)) \leq 0.$$

Fazendo algumas operações algébricas chega-se a

$$\int_0^T \nabla f'(\bar{x}(t) + s_n q(t), t) q(t) dt \geq \frac{r_n}{s_n} \int_0^T \nabla f'(\bar{x}(t) + s_n q(t), t) p(t) dt$$

Usando (3.30), tem-se

$$\begin{aligned} \int_0^T \nabla f'(\bar{x}(t) + s_n q(t), t) q(t) dt &\geq \frac{r_n}{s_n} \int_0^T [\nabla f'(\bar{x}(t) + s_n q(t), t) - \nabla f'(\bar{x}(t), t)] p(t) dt \\ &\geq -\frac{r_n}{s_n} \|\nabla f(\bar{x}(t) + s_n q(t), t) - \nabla f(\bar{x}(t), t)\| \|p(t)\| T \end{aligned} \quad (3.34)$$

onde $\|\cdot\|$ denota a norma Euclidiana. Desde que $f(\cdot, t)$ é duas vezes continuamente diferenciável em V por todo $[0, T]$, pode-se achar uma constante $c > 0$ tal que

$$\begin{aligned} \|\nabla f(\bar{x}(t) + s_n q(t), t) - \nabla f(\bar{x}(t), t)\| &\leq \lim_{x \rightarrow \bar{x}} \|\nabla f(\bar{x}(t) + s_n q(t), t) - \nabla f(\bar{x}(t), t)\| \\ &= s_n D^2 f(\bar{x}) \|q(t)\| = s_n c \|q(t)\| \end{aligned}$$

onde $x(t) = \bar{x}(t) + s_n q(t)$. Assim (3.34) torna-se

$$\int_0^T \nabla f'(\bar{x}(t) + s_n q(t), t) q(t) dt \geq -\frac{r_n}{s_n} s_n c \|q(t)\| \|p(t)\| T = -r_n c T \|q(t)\| \|p(t)\|.$$

Aplicando o limite com $n \rightarrow \infty$ tem-se

$$\lim_{n \rightarrow \infty} \int_0^T \nabla f'(\bar{x}(t) + s_n q(t), t) q(t) dt \geq \lim_{n \rightarrow \infty} -r_n c T \|q(t)\| \|p(t)\| = 0.$$

Pelo Teorema da convergência dominada de Lebesgue (1.1.1) e (3.31) tem-se

$$\lim_{n \rightarrow \infty} \int_0^T \nabla f'(\bar{x}(t) + s_n q(t), t) q(t) dt = \int_0^T \nabla f'(\bar{x}(t), t) q(t) dt < 0.$$

Assim tem-se uma contradição no limite. Logo

$$\int_0^T f(\bar{x}(t), t) dt \leq \int_0^T f(x(t), t) dt \quad \text{para todo } x \in \mathcal{F}.$$

Portanto, $\bar{x}$ é uma solução ótima global do problema (P).

(II) Quando $0 < r^* \leq 1$.

Seja $\{a_n\} = \{1 - \frac{r^*}{r_n}\}$. Desde que

$$\bar{x}(t) + r^* p(t) + a_n s_n q(t) = \frac{r^*}{r_n} (\bar{x}(t) + r_n p(t)) + (1 - \frac{r^*}{r_n}) (\bar{x}(t) + s_n q(t)),$$

por (3.33) tem-se

$$\int_0^T f(\bar{x}(t) + r_n p(t), t) dt = \int_0^T f(\bar{x}(t) + s_n q(t), t) dt.$$

Lembrando que pela quase-convexidade de ϕ em X tem-se

$$\int_0^T f(x_2, t) dt = \int_0^T f(x_1, t) dt \implies \int_0^T f(ux_2 + \mu x_1) dt \leq \int_0^T f(x_1) dt.$$

Tomando

$$x_2 = \bar{x}(t) + r_n p(t), \quad x_1 = \bar{x}(t) + s_n q(t), \quad u = \frac{r^*}{r_n} \quad \text{e} \quad \mu = 1 - \frac{r^*}{r_n}$$

obtem-se

$$\begin{aligned} \int_0^T f(\bar{x}(t) + r^* p(t) + a_n s_n q(t), t) dt &= \int_0^T f\left(\frac{r^*}{r_n}(\bar{x}(t) + r_n p(t)) + \left(1 - \frac{r^*}{r_n}\right)(\bar{x}(t) + s_n q(t), t) dt \right. \\ &\leq \int_0^T f(\bar{x}(t) + s_n q(t), t) dt. \end{aligned}$$

Portanto

$$\int_0^T f(\bar{x}(t) + r^* p(t) + a_n s_n q(t), t) dt \leq \int_0^T f(\bar{x}(t) + s_n q(t), t) dt.$$

Usando a desigualdade acima, (3.32) e multiplicando por $\frac{1}{a_n s_n}$ em ambos os lados chega-se a

$$\begin{aligned} \int_0^T \frac{1}{a_n s_n} [f(\bar{x}(t) + r^* p(t) + a_n s_n q(t), t) - f(\bar{x}(t) + r^* p(t), t)] dt &\leq \frac{1}{a_n} \int_0^T \frac{1}{s_n} [f(\bar{x}(t) + \\ &s_n q(t), t) - f(\bar{x}(t), t)] dt. \end{aligned} \quad (3.35)$$

Aplicando o limite com $n \rightarrow \infty$ e usando o Teorema da convergência dominada de Lebesgue (1.1.1) obtem-se

$$\lim_{n \rightarrow \infty} \int_0^T \frac{1}{a_n s_n} [f(\bar{x}(t) + r^* p(t) + a_n s_n q(t), t) - f(\bar{x}(t) + r^* p(t), t)] dt = \int_0^T \nabla f'(\bar{x}(t) + r^* p(t), t) q(t) dt$$

e

$$\lim_{n \rightarrow \infty} \frac{1}{a_n} \int_0^T \frac{1}{s_n} [f(\bar{x}(t) + s_n q(t), t) - f(\bar{x}(t), t)] dt = \frac{1}{a_n} \int_0^T \nabla f(\bar{x}(t), t) q(t) dt.$$

Assim (3.35) torna-se

$$\int_0^T \nabla f'(\bar{x}(t) + r^* p(t), t) q(t) dt \leq \frac{1}{a_n} \int_0^T \nabla f(\bar{x}(t), t) q(t) dt.$$

Aplicando novamente o limite com $n \rightarrow \infty$ tem-se

$$\lim_{n \rightarrow \infty} \int_0^T \nabla f'(\bar{x}(t) + r^* p(t), t) q(t) dt \leq \lim_{n \rightarrow \infty} \frac{1}{a_n} \int_0^T \nabla f(\bar{x}(t), t) q(t) dt.$$

Portanto

$$\int_0^T \nabla f'(\bar{x}(t) + r^* p(t), t) q(t) dt \leq \lim_{n \rightarrow \infty} \frac{1}{a_n} \int_0^T \nabla f(\bar{x}(t), t) q(t) dt.$$

O que é uma contradição, pois por (3.31) tem-se que o lado direito da desigualdade acima leva a $-\infty$ e o lado esquerdo a integral tem valor finito. Assim, em ambos os casos quando $r^* = 0$ e $r^* > 0$ a suposição $\phi(x) < \phi(\bar{x})$ leva a uma contradição. Assim, $\bar{x}$ é uma solução ótima global do problema (P). ■

Teorema 3.3.6 *As afirmações dos Teoremas (3.3.1) a (3.3.5) continuaram válidas se em (3.3.1), y for substituído por $z - x$, onde $z \in \mathbb{F}$, e a igualdade por maior ou igual.*

4.1 Relação entre o critério de otimalidade de ponto estacionário e ponto de sela.

O problema (P) e o conjunto F são definidos como no Capítulo 3. O teorema abaixo estabelece a equivalência entre condições de otimalidade de Fritz-John de ponto estacionário e ponto de sela.

Teorema 4.1.1 [25] *Seja $\bar{x} \in F$ e suponha que f e g são continuamente diferenciáveis com relação a sua primeira variável em $\bar{x}(t)$ por todo $[0, T]$ e que as funções $t \mapsto \nabla f(\bar{x}(t), t)$ e $t \mapsto \nabla g(\bar{x}(t), t)h(t)$, são Lebesgue integráveis em $[0, T]$ para todo $h \in L_{\infty}^n[0, T]$. Além disso, supondo que ϕ é convexa em $\bar{x}$ (com relação a F), g é convexa com relação a sua primeira variável em $\bar{x}(t)$ (com relação a $\{x(t) \in \mathbb{R}^n : x \in F\}$) por todo $[0, T]$, é então,*



Capítulo 4

Dualidade em programação não linear com tempo contínuo

Neste capítulo, estuda-se a relação entre condições de otimalidade de ponto estacionário e ponto de sela de Fritz-John. Usando conceitos de perturbação e estabilidade, desenvolvem-se condições de otimalidade de Karush-Kuhn-Tucker e dualidade Lagrangiana para problemas de programação convexa com tempo contínuo.

4.1 Relação entre o critério de otimalidade de ponto estacionário e ponto de sela.

O problema (P) e o conjunto IF são definidos como no Capítulo 2. O teorema abaixo estabelecerá a equivalência entre condições de otimalidade de Fritz-John de ponto estacionário e ponto de sela.

Teorema 4.1.1 [25] *Seja $\bar{x} \in IF$ e suponha que f e g são continuamente diferenciáveis com relação a sua primeira variável em $\bar{x}(t)$ por todo $[0, T]$ e que as funções $t \rightarrow \nabla f(\bar{x}(t), t)$ e $t \rightarrow \nabla g'(\bar{x}(t), t)h(t)$, são Lebesgue integráveis em $[0, T]$ para todo $h \in \mathcal{L}_\infty^n[0, T]$. Além disso, supondo que ϕ é convexa em $\bar{x}$ (com relação a IF), g é convexa com relação a sua primeira variável em $\bar{x}(t)$ (com relação a $\{x(t) \in \mathbb{R}^n : x \in IF\}$) por todo $[0, T]$, e existe*



$\bar{u}_0 \in \mathbb{R}, \bar{u} \in \mathcal{L}_\infty[0, T]$ tal que

$$\int_0^T [\bar{u}_0 \nabla f(\bar{x}(t), t) + \bar{u}(t) \nabla g'(\bar{x}(t), t)] h(t) dt = 0 \text{ para todo } h \in \mathcal{L}_\infty^n[0, T]; \quad (4.1)$$

$$\bar{u}(t) g(\bar{x}(t), t) = 0 \quad \text{q. s. em } [0, T]; \quad (4.2)$$

$$\bar{u}_0 \geq 0, \bar{u}(t) \geq 0 \quad \text{q. s. em } [0, T]; \quad (4.3)$$

$$(\bar{u}_0, \bar{u}(t)) \neq 0 \quad \text{q. s. em } [0, T]. \quad (4.4)$$

Então

$$(\bar{x}, \bar{u}_0, \bar{u})$$

satisfaz a seguinte desigualdade de ponto de sela

$$L(\bar{x}, \bar{u}_0, u) \leq L(\bar{x}, \bar{u}_0, \bar{u}) \leq L(x, \bar{u}_0, \bar{u}) \quad (4.5)$$

para todo $x \in X, u_0 \in \mathbb{R}, u \in \mathcal{L}_\infty[0, T]$ com $u_0 \geq 0, u(t) \geq 0$ q. s. em $[0, T]$, onde a função Lagrangiana

$$L : \mathcal{L}_\infty^n[0, T] \times \mathbb{R} \times \mathcal{L}_\infty[0, T] \longrightarrow \mathbb{R}$$

é definida por:

$$L(x, u_0, u) = \int_0^T [u_0 f(x(t), t) + u(t) g(x(t), t)] dt.$$

Reciprocamente, suponha que

$$(\bar{x}, \bar{u}_0, \bar{u}) \in \mathcal{L}_\infty^n[0, T] \times \mathbb{R} \times \mathcal{L}_\infty[0, T]$$

com $\bar{x} \in X, \bar{u}_0 \geq 0, \bar{u}(t) \geq 0, (\bar{u}_0, \bar{u}(t)) \neq 0$ q. s. em $[0, T]$, satisfazendo a desigualdade de ponto de sela (4.5). Então $\bar{x}$ é uma solução factível do problema (P), e $(\bar{x}, \bar{u}_0, \bar{u})$ satisfaz as condições (4.1) e (4.2).

Observe que usando a definição Lagrangiana, pode-se escrever a desigualdade de ponto de sela da seguinte maneira:

$$\int_0^T [\bar{u}_0 f(\bar{x}(t), t) + u(t) g(\bar{x}(t), t)] dt \leq \int_0^T [\bar{u}_0 f(\bar{x}(t), t) + \bar{u}(t) g(\bar{x}(t), t)] dt \leq \int_0^T [\bar{u}_0 f(x(t), t) + \bar{u}(t) g(x(t), t)] dt.$$

Demonstração :

Suponha que $(\bar{x}, \bar{u}_0, \bar{u})$ satisfaz as condições de Fritz John (4.1) - (4.4). Pela convexidade de $f(\cdot, t)$ e $g(\cdot, t)$ tem-se

$$\int_0^T \bar{u}_0 f(x(t), t) dt \geq \int_0^T \bar{u}_0 f(\bar{x}(t), t) dt + \int_0^T \bar{u}_0 \nabla f'(\bar{x}(t), t) [x(t) - \bar{x}(t)] dt \quad (4.6)$$

e

$$g(x(t), t) \geq g(\bar{x}(t), t) + \nabla g(\bar{x}(t), t) [x(t) - \bar{x}(t)] \quad (4.7)$$

para todo $x \in X$, $t \in [0, T]$.

Multiplicando (4.7) por $\bar{u}(t)$ e integrando obtem-se

$$\int_0^T \bar{u}(t) g(x(t), t) dt \geq \int_0^T \bar{u}(t) g(\bar{x}(t), t) dt + \int_0^T \bar{u}(t) \nabla g(\bar{x}(t), t) [x(t) - \bar{x}(t)] dt.$$

Adicionando o resultado acima a inequação (4.6) obtem-se

$$\begin{aligned} \int_0^T [\bar{u}_0 f(x(t), t) + \bar{u}(t) g(x(t), t)] dt &\geq \int_0^T [\bar{u}_0 f(\bar{x}(t), t) + \bar{u}(t) g(\bar{x}(t), t)] dt \\ &+ \int_0^T [(\bar{u}(t) \nabla g(\bar{x}(t), t) + \bar{u}_0 \nabla f'(\bar{x}(t), t)) [x(t) - \bar{x}(t)]] dt. \end{aligned}$$

Por (4.1) e (4.2), a desigualdade acima se reduz a

$$\int_0^T [\bar{u}_0 f(\bar{x}(t), t) + \bar{u}(t) g(\bar{x}(t), t)] dt \leq \int_0^T [\bar{u}_0 f(x(t), t) + \bar{u}(t) g(x(t), t)] dt$$

que é a segunda parte da desigualdade (4.5).

A primeira parte da desigualdade (4.5) segue de (4.2) e pelo fato de $u(t)g(\bar{x}(t), t) \leq 0$ q. s. em $[0, T]$, para todo $u(t) \geq 0$ q. s. em $[0, T]$.

Assim $(\bar{x}, \bar{u}_0, \bar{u})$ satisfaz (4.5).

Para provar a recíproca, suponha que $(\bar{x}, \bar{u}_0, \bar{u})$ satisfaz (4.5), com $\bar{x} \in X$, $\bar{u}_0 \geq 0$, $\bar{u}(t) \geq 0$, $(\bar{u}_0, \bar{u}(t)) \neq 0$ q. s. em $[0, T]$. Pela primeira desigualdade tem-se

$$\int_0^T [u(t) - \bar{u}(t)] g(\bar{x}(t), t) dt \leq 0. \quad (4.8)$$

Vai-se definir os conjuntos A e B da seguinte maneira:

$$A = \{t \in [0, T] : g(\bar{x}(t), t) \leq 0\}$$



e

$$B = \{t \in [0, T] : g(\bar{x}(t), t) > 0\}.$$

Sabe-se que $A \cap B = \emptyset$ e $A \cup B = [0, T]$. Tomando

$$u(t) = \begin{cases} \bar{u}(t) & \text{para todo } t \in A, \\ \bar{u}(t) + 1 & \text{para todo } t \in B. \end{cases}$$

Por (4.8) tem-se

$$\int_0^T [u(t) - \bar{u}(t)]g(\bar{x}(t), t)dt \leq 0.$$

Para $t \in B$ tem-se

$$\int_B [\bar{u}(t) + 1 - \bar{u}(t)]g(\bar{x}(t), t)dt \leq 0,$$

então

$$\int_B g(\bar{x}(t), t)dt \leq 0.$$

Como $g(\bar{x}(t), t) > 0$ para todo $t \in B$ tem-se

$$|B| = 0$$

onde $|\cdot|$ é a medida de Lebesgue definida no campo sigma de todos os subconjuntos de $[0, T]$.

Assim conclui-se que

$$g(\bar{x}(t), t) \leq 0 \quad \text{q. s. em } [0, T].$$

E portanto, $\bar{x}$ é uma solução factível do problema (P). Já que

$$\bar{u}(t) \geq 0 \quad \text{q. s. em } [0, T]$$

tem-se

$$\bar{u}(t)g(\bar{x}(t), t) \leq 0 \quad \text{q. s. em } [0, T].$$

Se a desigualdade estrita vale em um subconjunto de $[0, T]$ com medida de Lebesgue positiva, então pela desigualdade de ponto de sela (4.5) com $u(t) = 0$ q. s. em $[0, T]$ tem-se

$$\int_0^T \bar{u}_0 f(\bar{x}(t), t)dt < \int_0^T \bar{u}_0 f(\bar{x}(t), t)dt,$$

o que é um absurdo. Assim:

$$\bar{u}(t)g(\bar{x}(t), t) = 0 \quad \text{q. s. em } [0, T].$$

Seja

$$\psi(\bar{x}) = L(\bar{x}, \bar{u}_0, \bar{u});$$

$$\nabla\psi(\bar{x}) = \nabla_x L(\bar{x}, \bar{u}_0, \bar{u}).$$

Como $\bar{x}$ é uma solução factível do problema (P) então :

$$\nabla\psi(\bar{x})h \leq 0 \text{ para todo } h \in \mathcal{L}_\infty^n[0, T];$$

$$\nabla_x L(\bar{x}, \bar{u}_0, \bar{u})h \leq 0 \text{ para todo } h \in \mathcal{L}_\infty^n[0, T].$$

Portanto

$$\int_0^T [\bar{u}_0 \nabla f(\bar{x}(t), t) + \bar{u}(t) \nabla g(\bar{x}(t), t)] h(t) dt \leq 0 \text{ para todo } h \in \mathcal{L}_\infty^n[0, T].$$

Já que a desigualdade acima é satisfeita para todo $h \in \mathcal{L}_\infty^n[0, T]$, então também será satisfeita para a igualdade. Assim:

$$\int_0^T [\bar{u}_0 \nabla f(\bar{x}(t), t) + \bar{u}(t) \nabla g(\bar{x}(t), t)] h(t) dt = 0 \text{ para todo } h \in \mathcal{L}_\infty^n[0, T]. \quad \blacksquare$$

4.2 Condições de Otimalidade e Dualidade Lagrangiana em Programação não linear com tempo contínuo.

4.2.1 Os problemas primal e dual

Considere o problema (P) como foi definido no Capítulo 2 e g uma função convexa com relação a sua primeira variável por todo $[0, T]$. A restrição de igualdade da forma:

$$h(x(t), t) = 0 \text{ q. s. em } [0, T]$$

onde $h(x(t), t) = \eta(x)(t)$, com $\eta : X \longrightarrow \Lambda_1^1[0, T]$, é linear com relação a sua primeira variável por todo $[0, T]$, podendo também ser usada no problema (P), pois a equação linear pode sempre ser substituída por um equivalente par de desigualdade linear sem alterar a convexidade.

Define-se o seguinte dual do problema (P):

$$(D) \quad \begin{aligned} &\max \quad \psi(u), \\ &\text{s. a. } u(t) \geq 0 \quad \text{q. s. em } [0, T], \end{aligned}$$

onde

$$\psi(u) = \inf_{x \in X} \int_0^T [f(x(t), t) + u(t)g(x(t), t)] dt \quad (4.9)$$

e $u \in \mathcal{L}_\infty[0, T]$.

Agora, deve-se apresentar algumas definições e introduzir algumas notações adicionais e terminologia que serão necessárias para o desenvolvimento da teoria da dualidade.

Definição 4.2.1 Um par $(x, u) \in \mathcal{L}_\infty^n[0, T] \times \mathcal{L}_\infty[0, T]$ é dito satisfazer as condições de otimalidade do Problema (P) se:

(a) x minimiza

$$\int_0^T [f(x(t), t) + u(t)g(x(t), t)] dt \quad \text{em } X;$$

(b) $u(t)g(x(t), t) = 0 \quad \text{q. s. em } [0, T];$

(c) $u(t) \geq 0 \quad \text{q. s. em } [0, T];$

(d) $g(x(t), t) \leq 0 \quad \text{q. s. em } [0, T].$

A condição (b) é conhecida como a condição de folgas complementares.

Definição 4.2.2 Um elemento $u \in \mathcal{L}_\infty[0, T]$ é um multiplicador ótimo do problema (P) se (x, u) satisfaz as condições de otimalidade para algum $x \in \mathcal{L}_\infty^n[0, T]$.

Definição 4.2.3 A função perturbação $p : \Lambda_1^1[0, T] \rightarrow \mathbb{R}$ associada com o problema (P) é definida por:

$$p(y) = \inf_{x \in X} \left\{ \int_0^T f(x(t), t) dt : g(x(t), t) \leq y(t) \quad \text{q. s. em } [0, T] \right\}.$$

A função p não é Fréchet diferenciável, mas como será provado depois, p é uma função convexa e assim possui muitas propriedades importantes. Além disso, p é uma função não crescente.

Definição 4.2.4 *Sejam Z um espaço normado real com dual topológico Z^* , e $h : Z \rightarrow [-\infty, \infty]$ uma função convexa tal que $h(\bar{z})$ é finita para algum $\bar{z} \in Z$. Então $z^* \in Z^*$ é o subgradiente de h em $\bar{z}$ se*

$$h(z) - h(\bar{z}) \geq z^*(z - \bar{z}) \text{ para todo } z \in Z.$$

O conjunto de todos os subgradientes de h em $\bar{z}$ é conhecido como o subdiferencial de h em $\bar{z}$ e é denotado por $\partial h(\bar{z})$.

Por subdiferenciabilidade e propriedades relativas a funções convexas os livros de Holmes [13] e Rockafellar [21] podem ser consultados.

Seja $\mathcal{L}_1[0, T]$ o espaço de todas as funções Lebesgue integráveis definidas em $[0, T]$, com norma $\|\cdot\|_1$ definida. Sabe-se por [5] que existe um isomorfismo isométrico entre o dual topológico $(\mathcal{L}_1[0, T])^*$ de $\mathcal{L}_1[0, T]$ e $\mathcal{L}_\infty[0, T]$, onde os elementos correspondentes y^* e u são relacionados pela identidade

$$y^*(y) = \int_0^T u(t)y(t)dt \text{ para todo } y \in \mathcal{L}_1[0, T]. \quad (4.10)$$

Já que $\Lambda_1^1[0, T]$ é um subespaço denso de $\mathcal{L}_1[0, T]$, segue que para cada elemento y^* em $(\Lambda_1^1[0, T])^*$ existe $u \in \mathcal{L}_\infty[0, T]$ tal que (4.10) é válida. Assim deve-se escolher uma definição equivalente a de subgradiente para a função perturbação p .

Definição 4.2.5 *Se p é finito em $\bar{y} \in \Lambda_1^1[0, T]$, diz-se que a função $u \in \mathcal{L}_\infty[0, T]$ é um subgradiente de p em $\bar{y}$ se e somente se*

$$p(y) - p(\bar{y}) \geq \int_0^T u(t)[y(t) - \bar{y}(t)]dt \text{ para todo } y \in \Lambda_1^1[0, T].$$

Definição 4.2.6 *O problema (P) é dito ser estável se $p(0)$ é finito e existe uma constante $m > 0$ tal que*

$$p(0) \leq p(y) + m\|y\|_1 \text{ para todo } y \in \Lambda_1^1[0, T].$$

Definição 4.2.7 O conjunto factível do problema perturbado é dado por:

$$Y = \{y \in \Lambda_1^1[0, T] : g(x(t), t) \leq y(t) \text{ q. s. em } [0, T] \text{ para algum } x \in X\}.$$

É óbvio que $p(y) = \infty$ se e somente se $y \notin Y$.

4.2.2 Resultados Auxiliares

Agora vê-se alguns resultados auxiliares que são necessários para determinar o critério de otimalidade e teoremas de dualidade.

Lema 4.2.1 Y é um conjunto convexo e p é uma função convexa em Y .

Demonstração : Seja

$$Y = \{y \in \Lambda_1^1[0, T] : g(x(t), t) \leq y(t) \text{ q. s. em } [0, T] \text{ para algum } x \in X\}.$$

Vai-se mostrar que Y é um conjunto convexo.

De fato:

Sabe-se que $g(\cdot, t)$ é uma função convexa e X é um conjunto convexo. Seja $y_1, y_2 \in Y$ então

$$\begin{aligned} \lambda y_1 + (1 - \lambda)y_2 &\geq \lambda g(x(t), t) + (1 - \lambda)g(x(t), t) \\ &= \lambda g(x(t), t) + g(x(t), t) - \lambda g(x(t), t) \\ &= g(x(t), t). \end{aligned}$$

Assim

$$\lambda y_1 + (1 - \lambda)y_2 \geq g(x(t), t)$$

Logo

$$\lambda y_1 + (1 - \lambda)y_2 \in Y.$$

Portanto, Y é um conjunto convexo.

Agora, vai-se provar que p é uma função convexa em Y .

De fato, sejam

$$y^1, y^2 \in Y, y = \lambda y^1 + (1 - \lambda)y^2, 0 < \lambda < 1 \text{ e } \epsilon > 0.$$

Pela definição de Y existe $x^1, x^2 \in X$ tal que

$$g(x(t), t) \leq y^i(t) \text{ q. s. em } [0, T]$$

e

$$\phi(x^i) < p(y^i) + \epsilon$$

para $i = 1, 2$.

Como X e Y são conjuntos convexos e ϕ é uma função convexa em X , tem-se

$$\begin{aligned} p(y) \leq \phi(\lambda x^1 + (1 - \lambda)x^2) &\leq \lambda\phi(x^1) + (1 - \lambda)\phi(x^2) \\ &< \lambda(p(y^1) + \epsilon) + (1 - \lambda)(p(y^2) + \epsilon). \end{aligned}$$

Já que ϵ é arbitrário, tem-se que p é uma função convexa em Y . ■

Lema 4.2.2 [23, 9] *Seja h uma função convexa em um conjunto convexo Y de um espaço normado real Z com norma $\|\cdot\|$ definida e supondo que h é finita em $\bar{y} \in Y$. Então h tem subgradiente em $\bar{y}$ ($\partial h(\bar{y}) \neq \emptyset$) se e somente se existe uma constante $m > 0$ tal que*

$$h(\bar{y}) \leq h(y) + m\|y - \bar{y}\| \text{ para todo } y \in Y. \quad (4.11)$$

Demonstração :

($\Rightarrow$)

Suponha que h tem um subgradiente y^* em $\bar{y} \in Y$, isto é,

$$h(y) - h(\bar{y}) \geq y^*(y - \bar{y}) \text{ para todo } y \in Y. \quad (4.12)$$

Assim

$$h(y) \geq h(\bar{y}) + y^*(y - \bar{y}) > -\infty \text{ em } y \in Y.$$

Multiplicando por -1 e dividindo por $\|y - \bar{y}\|$ em ambos os lados em (4.12), obtém-se

$$\frac{h(\bar{y}) - h(y)}{\|y - \bar{y}\|} \leq \frac{-y^*(y - \bar{y})}{\|y - \bar{y}\|} \text{ para todo } y \in Y, \text{ tal que } y \neq \bar{y}.$$

Sabe-se que

$$\frac{-y^*(y - \bar{y})}{\|y - \bar{y}\|} \leq \frac{\|y^*(y - \bar{y})\|}{\|y - \bar{y}\|} \leq \frac{\|y^*\| \|y - \bar{y}\|}{\|y - \bar{y}\|} = \|y^*\|.$$

Assim

$$\frac{h(\bar{y}) - h(y)}{\|y - \bar{y}\|} \leq \|y^*\|.$$

ou seja

$$h(\bar{y}) - h(y) \leq \|y^*\| \|y - \bar{y}\|. \quad (4.13)$$

Portanto, $m = \|y^*\|$. Se $\|y^*\| = 0$ então (4.13) é satisfeita para $\forall m > 0$. Por exemplo, tomando $m = 10$, se $\|y^*\| = 0$ tem-se

$$\frac{h(\bar{y}) - h(y)}{\|y - \bar{y}\|} \leq 0$$

e com $m = 10$, tem-se

$$\frac{h(\bar{y}) - h(y)}{\|y - \bar{y}\|} \leq 10$$

é satisfeito.

($\Leftarrow$)

Agora, suponha que existe um escalar $m > 0$ tal que a desigualdade em (4.11) é satisfeita. Deve-se mostrar que existe um subgradiente de h em $\bar{y}$. Desde que $h(y) > -\infty$ em Y , pode-se definir os conjuntos convexos

$$\Phi := \{(y, z) \in \mathbb{R}^2 : y \in Y \text{ e } h(\bar{y}) - h(y) \geq z\}$$

e

$$\Psi := \{(y, z) \in \mathbb{R}^2 : m\|y - \bar{y}\| < z \Rightarrow h(\bar{y}) - h(y) \leq m\|y - \bar{y}\| < z \Rightarrow h(\bar{y}) - h(y) < z\}.$$

Como Φ e Ψ são conjuntos convexos, é fácil ver que $\Phi \cap \Psi = \emptyset$ e que Ψ é um conjunto aberto. Por resultados elementares de separação de conjuntos convexos com intersecção vazia, obtem-se que Φ e Ψ podem ser separados por um hiperplano que não intercepta Ψ , isto é, existe y^* e escalares ρ e α tal que

$$y^*y + \rho z \geq \alpha \quad \text{para } (y, z) \in \Phi \quad (4.14)$$

e substituindo em (4.13) tem-se

$$y^*y + \rho z < \alpha \quad \text{para } (y, z) \in \Psi. \quad (4.15)$$

Agora, como $(\bar{y}, 0) \in \Phi$, pois,

$$0 = h(\bar{y}) - h(\bar{y}) \geq 0$$

e substituindo $(\bar{y}, 0)$ em (4.14) obtem-se

$$y^* \bar{y} \geq \alpha. \quad (4.16)$$

Seja $(\bar{y}, \epsilon)$ um par ordenado, com $\epsilon > 0$, então $(\bar{y}, \epsilon) \in \Psi$, pois,

$$0 = m \|\bar{y} - \bar{y}\| < \epsilon$$

e substituindo $(\bar{y}, \epsilon)$ em (4.15) tem-se

$$y^* \bar{y} + \rho \epsilon < \alpha.$$

Aplicando o limite com $\epsilon \rightarrow 0$ obtem-se

$$y^* \bar{y} \leq \alpha. \quad (4.17)$$

De (4.16) e (4.17) tem-se

$$\alpha = y^* \bar{y}.$$

Assim as desigualdades (4.14) e (4.15) tornam-se

$$y^*(y - \bar{y}) + \rho z \geq 0 \quad \text{para } (y, z) \in \Phi \quad (4.18)$$

e

$$y^*(y - \bar{y}) + \rho z < 0 \quad \text{para } (y, z) \in \Psi. \quad (4.19)$$

Se $(\bar{y}, 1) \in \Psi$ tem-se

$$y^* \bar{y} + \rho < \alpha = y^* \bar{y} \Rightarrow \rho < 0.$$

Tomando

$$\bar{y}^* = -\frac{y^*}{\rho} \Rightarrow y^* = -\bar{y}^* \rho$$

e substituindo em (4.18) tem-se

$$y^*(y - \bar{y}) + \rho z \geq 0 \Rightarrow -\bar{y}^* \rho (y - \bar{y}) + \rho z \geq 0.$$



Assim

$$\rho(-\bar{y}^*(y - \bar{y}) + z) \geq 0.$$

Como $\rho < 0$ tem-se

$$-\bar{y}^*(y - \bar{y}) + z \leq 0$$

Assim

$$-\bar{y}^*(y - \bar{y}) \leq -z \Rightarrow \bar{y}^*(y - \bar{y}) \geq z.$$

Assim

$$\bar{y}^*(y - \bar{y}) \geq z \text{ sempre que } (y, z) \in \Phi,$$

ou seja, sempre que $y \in Y$ e $h(\bar{y}) - h(y) \geq z$.

Tomando $z = h(\bar{y}) - h(y)$ tem-se

$$\bar{y}^*(y - \bar{y}) \geq h(\bar{y}) - h(y) \text{ sempre que } y \in Y.$$

Assim

$$h(y) - h(\bar{y}) \geq -\bar{y}^*(y - \bar{y}) \text{ sempre que } y \in Y.$$

Por definição tem-se que $-\bar{y}^*$ é o subgradiente de h em $\bar{y}$. ■

Lema 4.2.3 *Suponha que o problema (P) tem uma solução ótima. Então $\bar{u} \in \mathcal{L}_\infty[0, T]$ é uma função multiplicador ótimo do problema (P) se e somente se $-\bar{u} \in \partial p(0)$.*

Demonstração :

Seja $\bar{u}$ sendo uma função multiplicador ótimo do problema (P), então existe $\bar{x} \in X$ tal que $(\bar{x}, \bar{u})$ satisfaz as condições de otimalidade (a) - (d) da definição (4.2.1). As condições (a) e (b) levam a

$$\int_0^T f(x(t), t) dt \geq \int_0^T [f(\bar{x}(t), t) - \bar{u}(t)g(x(t), t)] dt \text{ para todo } x \in X.$$

Da condição (c), segue que

$$\int_0^T f(x(t), t) dt \geq \int_0^T [f(\bar{x}(t), t) - \bar{u}(t)y(t)] dt \text{ para todo } x \in X$$

tal que $g(x(t), t) \leq y(t)$ q. s. em $[0, T]$. Consequentemente

$$\inf_{x \in X} \left\{ \int_0^T f(x(t), t) dt : g(x(t), t) \leq y(t) \text{ q.s. em } [0, T] \right\} \geq \int_0^T [f(\bar{x}(t), t) - \bar{u}(t)y(t)] dt, \forall y \in Y.$$

Assim

$$p(y) \geq p(0) + \int_0^T [-\bar{u}(t)] y(t) dt \text{ para } y \in Y.$$

Portanto, $-\bar{u} \in \partial p(0)$.

Para provar a recíproca, suponha que $-\bar{u} \in \partial p(0)$ e $\bar{x}$ sendo uma solução ótima do problema (P). Deve-se provar que o par $(\bar{x}, \bar{u})$ satisfaz as condições de otimalidade (a) - (d). A condição (d) é satisfeita pois $\bar{x}$ é ótimo e assim factível. Suponha que a condição (c) não é satisfeita, ou seja, que $\bar{u}(t)$ é estritamente negativo em um subconjunto de $[0, T]$ com medida de Lebesgue positiva. Então tem-se

$$\int_0^T \bar{u}(t) dt < 0.$$

Já que por hipótese

$$p(y) \geq p(0) + \int_0^T [-\bar{u}(t)] y(t) dt \text{ para todo } y \in Y$$

e se escolher $y(t) = 1$ e lembrando que a função p é não crescente, obtem-se

$$\int_0^T \bar{u}(t) dt \geq p(0) - p(1) \geq 0$$

o que é uma contradição. Assim, $\bar{u}(t) \geq 0$ q. s. em $[0, T]$ e a condição (c) é satisfeita.

Para provar (b), de (c) e (d) tem-se

$$\bar{u}(t)g(\bar{x}(t), t) \leq 0 \text{ q. s. em } [0, T].$$

Entretanto, se a desigualdade estrita é satisfeita em um subconjunto de $[0, T]$ com medida de Lebesgue positiva, então

$$\int_0^T \bar{u}(t)g(\bar{x}(t), t) dt < 0. \quad (4.20)$$

Sejam

$$p(0) = \inf_{x \in X} \left\{ \int_0^T f(x(t), t) dt : g(x(t), t) \leq 0 \text{ q. s. em } [0, T] \right\}$$



e afirmação usando o Lema (4.2.3).

$$p(\Gamma(\bar{x})) = \inf_{x \in X} \left\{ \int_0^T f(x(t), t) dt : g(x(t), t) \leq \Gamma(\bar{x}) = g(x(t), t) \leq 0 \text{ q. s. em } [0, T] \right\}.$$

Assim

$$p(0) - p(\Gamma(\bar{x})) = 0.$$

Assim (4.20) contradiz

$$\int_0^T \bar{u}(t)g(\bar{x}(t), t)dt \geq p(0) - p(\Gamma(\bar{x})) = 0.$$

Desta forma deve-se ter

$$\bar{u}(t)g(\bar{x}(t), t) = 0 \text{ q. s. em } [0, T]$$

e portanto (b) é satisfeita.

Finalmente, para provar que a condição (a) é satisfeita, seja

$$y(t) = g(x(t), t) = \Gamma(x)(t).$$

Então tem-se

$$p(\Gamma(x)) \geq p(0) - \int_0^T \bar{u}(t)g(x(t), t)dt = \int_0^T [f(\bar{x}(t), t) - \bar{u}(t)g(x(t), t)]dt \text{ para todo } x \in X.$$

Assim, como $p(\Gamma(x)) = p(0) = \int_0^T f(\bar{x}(t), t)dt$, com $\bar{x}$ sendo a solução ótima, tem-se

$$\int_0^T f(x(t), t)dt \geq p(\Gamma(x)) \text{ para todo } x \in X,$$

e por (b), tem-se

$$\int_0^T [f(\bar{x}(t), t) + \bar{u}(t)g(\bar{x}(t), t)]dt \leq \int_0^T [f(x(t), t) + \bar{u}(t)g(x(t), t)]dt \text{ para todo } x \in X.$$

Portanto (a) é satisfeita. ■

Como consequência do Lema (4.2.3), observa-se que existe um único multiplicador ótimo se e somente se a função p é Fréchet diferenciável em zero. A função multiplicador ótimo pode ser útil em análise sensitiva porque pode dar certas estimativas para a escolha de valores da função objetivo resultante da perturbação de restrições. Deve-se justificar esta



afirmação usando o Lema (4.2.3).

Seja x^i uma solução ótima do problema:

$$(P^i) \quad \begin{aligned} \min \quad & \phi(x) = \int_0^T f(x(t), t) dt, \\ \text{s. a.} \quad & g(x(t), t) \leq y^i(t) \quad \text{q. s. em } [0, T], \quad i = 1, 2, \\ & x \in X. \end{aligned}$$

Se u^i , $i = 1, 2$ são as correspondentes funções multiplicadores ótimos, então tem-se

$$\begin{aligned} \int_0^T u^1(t)[y^1(t) - y^2(t)]dt &\leq \int_0^T f(x^2(t), t)dt - \int_0^T f(x^1(t), t)dt \\ &\leq \int_0^T u^2(t)[y^1(t) - y^2(t)]dt. \end{aligned} \quad (4.21)$$

Para verificar se a afirmação acima é verdadeira, suponha que q seja a função perturbação do problema convexo

$$\begin{aligned} \min \quad & \phi(x) = \int_0^T f(x(t), t)dt, \\ \text{s. a.} \quad & g(x(t), t) - y^2(t) \leq 0 \quad \text{q. s. em } [0, T], \\ & x \in X. \end{aligned}$$

Se p é a função perturbação do problema (P) então $p(y) = q(y - y^2)$.

De fato:

Com relação ao problema (P) , pela definição de função perturbação tem-se

$$p(y) = \inf_{x \in X} \left\{ \int_0^T f(x(t), t)dt : g(x(t), t) \leq y(t) \quad \text{q. s. em } [0, T] \right\}.$$

Com relação ao problema convexo, pode-se definir a função perturbação da seguinte maneira:

$$q(y) = \inf_{x \in X} \left\{ \int_0^T f(x(t), t)dt : g(x(t), t) - y^2(t) \leq y(t) \quad \text{q. s. em } [0, T] \right\}.$$

Assim pode-se verificar que

$$\begin{aligned} q(y - y^2) &= \inf_{x \in X} \left\{ \int_0^T f(x(t), t)dt : g(x(t), t) - y^2(t) \leq y(t) - y^2(t) \quad \text{q. s. em } [0, T] \right\} \\ &= \inf_{x \in X} \left\{ \int_0^T f(x(t), t)dt : g(x(t), t) \leq y(t) \quad \text{q. s. em } [0, T] \right\} = p(y). \end{aligned}$$

Assim

$$q(y - y^2) = p(y)$$

Pela igualdade acima, quando $y - y^2 = 0$ tem-se que $y = y^2$. Assim $q(0) = p(y^2)$ e portanto

$$\partial p(y^2) = \partial q(0).$$

Já que u^2 é uma função multiplicador ótimo do problema (P^2) , então pelo Lema (4.2.3) tem-se

$$-u^2 \in \partial q(0) = \partial p(y^2).$$

Agora, como x^i é a solução ótima do problema (P^i) tem-se

$$\int_0^T f(x^i(t), t) dt = p(y^i), \quad i = 1, 2.$$

Assim

$$\int_0^T f(x^2(t), t) dt - \int_0^T f(x^1(t), t) dt = -[p(y^1) - p(y^2)].$$

Pela definição de subgradiente tem-se

$$p(y^1) - p(y^2) \geq \int_0^T [-u^2(t)](y^1 - y^2) dt.$$

Assim

$$-[p(y^1) - p(y^2)] \leq \int_0^T u^2(t)(y^1 - y^2) dt.$$

Portanto

$$\int_0^T f(x^2(t), t) dt - \int_0^T f(x^1(t), t) dt \leq \int_0^T u^2(t)(y^1 - y^2) dt$$

que é o lado direito da equação (4.21).

De modo análogo tem-se

$$\int_0^T f(x^2(t), t) dt - \int_0^T f(x^1(t), t) dt = p(y^2) - p(y^1) \geq \int_0^T [-u^1(t)](y^2 - y^1) dt = \int_0^T u^1(y^1 - y^2) dt.$$

4.2.3 Condições de otimalidade

Nesta seção obtem-se um critério de otimalidade análogo ao critério de otimalidade de ponto estacionário e ponto de sela de Karush-Kuhn-Tucker com tempo contínuo para o problema (P) . Primeiro, deve-se examinar a relação entre a noção de estabilidade e a existência da função multiplicador ótimo.

Teorema 4.2.1 Suponha que o problema (P) tem solução ótima. Então uma função multiplicador ótimo existe se e somente se o problema (P) é estável, e $u \in \mathcal{L}_\infty[0, T]$ é uma função multiplicador ótimo se e somente se $-u \in \partial p(0)$.

Demonstração :

Pelo Lema (4.2.1), p é uma função convexa em Y e por hipótese, o problema (P) tem uma solução ótima e portanto $p(0)$ é finito. Assim, pelo Lema (4.2.2), a função p tem um subgradiente $u \in \mathcal{L}_\infty[0, T]$ em $y = 0$ se e somente se o problema (P) é estável.

Como o problema (P) tem uma solução ótima, pelo Lema (4.2.3), $\bar{u} \in \mathcal{L}_\infty[0, T]$ é uma função multiplicador ótimo do problema (P) se e somente se $-u \in \partial p(0)$. ■

No próximo teorema, deve-se estabelecer a equivalência entre as condições de otimalidade da definição (4.2.1) e a desigualdade de ponto de sela da função Lagrangiana L .

Teorema 4.2.2 Um par $(\bar{x}, \bar{u})$ satisfaz as condições de otimalidade (a) - (d) da definição (4.2.1) se e somente se $\bar{u}(t) \geq 0$ q. s. em $[0, T]$, $\bar{x} \in X$, e

$$L(\bar{x}, u) \leq L(\bar{x}, \bar{u}) \leq L(x, \bar{u})$$

para todo $u \in \mathcal{L}_\infty[0, T]$, $u(t) \geq 0$ q. s. em $[0, T]$ e $\forall x \in X$.

Demonstração :

($\Rightarrow$)

Suponha que $(\bar{x}, \bar{u})$ satisfaz as condições de otimalidade do problema (P) , então pela condição (a) da definição (4.2.1) tem-se que $\bar{x}$ minimiza

$$\int_0^T [f(x(t), t) + \bar{u}(t)g(x(t), t)]dt, \text{ em } X.$$

Assim

$$L(\bar{x}, \bar{u}) \leq L(x, \bar{u}).$$

Pela condição (d) da definição (4.2.1) tem-se

$$g(\bar{x}(t), t) \leq 0 \text{ q. s. em } [0, T].$$

Assim

$$\int_0^T u(t)g(\bar{x}(t), t)dt \leq 0, \text{ para } \forall u(t) \geq 0 \text{ q. s. em } [0, T]. \quad (4.22)$$

Pela condição (b) da definição (4.2.1) tem-se

$$u(t)g(\bar{x}(t), t) = 0 \text{ q. s. em } [0, T].$$

Assim

$$\int_0^T \bar{u}(t)g(\bar{x}(t), t)dt = 0 \text{ q. s. em } [0, T]. \quad (4.23)$$

De (4.22) e (4.23) tem-se

$$L(\bar{x}, u) \leq L(\bar{x}, \bar{u}).$$

($\Leftarrow$)

Por hipótese, a condição (c) é satisfeita.

Suponha que a condição de ponto de sela é satisfeita, então

$$L(\bar{x}, \bar{u}) \leq L(x, \bar{u}) \text{ para } \forall x \in X.$$

Assim $\bar{x}$ minimiza

$$\int_0^T [f(x(t), t) + \bar{u}(t)g(x(t), t)]dt, \text{ em } X.$$

Assim a condição (a) é satisfeita. De maneira análoga a demonstração do Teorema (4.1.1), conclui-se que as condições (b) e (d) são satisfeitas. ■

O teorema a seguir é conhecido como o Teorema de ponto de sela de Karush-Kuhn-Tucker.

Teorema 4.2.3 *Suponha que o problema (P) é estável. Então $\bar{x} \in X$ é uma solução ótima do problema (P) se e somente se existe $\bar{u} \in \mathcal{L}_\infty[0, T]$, $\bar{u}(t) \geq 0$ q. s. em $[0, T]$, tal que*

$$L(\bar{x}, u) \leq L(\bar{x}, \bar{u}) \leq L(x, \bar{u})$$

para todo $u \in \mathcal{L}_\infty[0, T]$, $u(t) \geq 0$ q. s. em $[0, T]$ e $\forall x \in X$.

Demonstração :

($\Rightarrow$)

Suponha que $\bar{x} \in X$ é uma solução ótima do problema (P). Desde que o problema (P) é estável, pelo Teorema (4.2.1), tem-se que existe $\bar{u} \in \mathcal{L}_\infty[0, T]$, que é a função multiplicador ótimo. Pela definição (4.2.2), tem-se que $(\bar{x}, \bar{u})$ satisfaz as condições (a) - (d) da definição

(4.2.1), o qual pelo Teorema (4.2.2) completa a demonstração.

($\Leftarrow$)

Suponha que as condições de ponto de sela são satisfeitas, então como na demonstração do Teorema (4.2.2), $\bar{x}$ é uma solução factível do problema (P) e $(\bar{x}, \bar{u})$ satisfaz a condição de folgas complementares. Já que $\bar{u}(t)g(\bar{x}(t), t) = 0$, $\bar{u}(t) \geq 0$ e $g(x(t), t) \leq 0$ q. s. em $[0, T]$ para cada solução factível do problema (P), pela segunda parte da desigualdade de ponto de sela tem-se

$$L(\bar{x}, \bar{u}) \leq L(x, \bar{u}),$$

ou seja,

$$\int_0^T [f(\bar{x}(t), t) + \bar{u}(t)g(\bar{x}(t), t)]dt \leq \int_0^T [f(x(t), t) + \bar{u}(t)g(x(t), t)]dt.$$

Então

$$\int_0^T f(\bar{x}(t), t)dt \leq \int_0^T f(x(t), t)dt,$$

isto é,

$$\phi(\bar{x}) \leq \phi(x)$$

para toda solução factível do problema (P).

Portanto, $\bar{x}$ é solução ótima do problema (P). ■

Agora vai-se mostrar que o resultado visto acima leva a uma forma diferencial das condições de Karush-Kuhn-Tucker na presença da diferenciabilidade. O teorema a seguir é conhecido como o Teorema de ponto estacionário de Karush-Kuhn-Tucker.

Teorema 4.2.4 *Suponha que o interior de X ($\text{int}X$) é diferente de vazio, e que f e g são continuamente diferenciáveis e convexas com relação a sua primeira variável em um conjunto aberto $V \subset \mathbb{R}^n$ contendo o conjunto $\{x(t) : x \in X, t \in [0, T]\}$ por todo $[0, T]$, e que as funções $t \rightarrow \nabla f(\bar{x}(t), t)$ e $t \rightarrow \nabla g'(\bar{x}(t), t)h(t) \equiv D\Gamma(\bar{x})(h)(t)$ (a derivada de Fréchet de Γ em $\bar{x}$ calculada em $h(t)$), são Lebesgue integráveis em $[0, T]$ para todo $h \in \mathcal{L}_\infty^n[0, T]$. Se o problema (P) é estável, então $(\bar{x}, \bar{u}) \in \mathcal{L}_\infty^n[0, T] \times \mathcal{L}_\infty[0, T]$ com $\bar{x} \in \text{int} X$ e $\bar{u}(t) \geq 0$ q. s. em $[0, T]$, satisfaz a condição de ponto de sela*

$$L(\bar{x}, u) \leq L(\bar{x}, \bar{u}) \leq L(x, \bar{u})$$

para todo $u \in \mathcal{L}_\infty[0, T]$, $u(t) \geq 0$ q. s. em $[0, T]$ e $\forall x \in X$, se e somente se $\bar{x}$ é uma solução factível do problema (P) e $(\bar{x}, \bar{u})$ satisfaz as condições de Karush-Kuhn-Tucker

$$\int_0^T [\nabla f'(\bar{x}(t), t) + \bar{u}(t) \nabla g'(\bar{x}(t), t)] h(t) dt = 0 \text{ para todo } h \in \mathcal{L}_\infty^n[0, T]; \quad (4.24)$$

$$\bar{u}(t) g(\bar{x}(t), t) = 0 \text{ q. s. em } [0, T].$$

Demonstração :

($\Rightarrow$)

Suponha que $(\bar{x}, \bar{u})$ com $\bar{x} \in \text{int } X$ e $\bar{u}(t) \geq 0$ q. s. em $[0, T]$ satisfaz as condições de ponto de sela, então como foi demonstrado no Teorema (4.2.2), $\bar{x}$ é uma solução factível do problema (P), e a condição de folgas complementares é satisfeita.

Já que $L(\bar{x}, \bar{u}) \leq L(x, \bar{u})$ para todo $x \in X$ e $\bar{x}$ é uma solução ótima do problema:

$$\begin{aligned} \min \quad & L(x, \bar{u}) = \int_0^T [f(x(t), t) + \bar{u}(t) g(x(t), t)] dt, \\ \text{s. a.} \quad & x \in X, \end{aligned}$$

o qual implica que a derivada de Fréchet da função $L(\cdot, \bar{u})$ se anula em $\bar{x}$, isto é

$$\int_0^T [\nabla f'(\bar{x}(t), t) + \bar{u}(t) \nabla g'(\bar{x}(t), t)] h(t) dt = 0 \text{ para todo } h \in \mathcal{L}_\infty^n[0, T]. \quad (4.25)$$

($\Leftarrow$)

Suponha que as condições de Karush-Kuhn-Tucker são satisfeitas, então pela hipótese de convexidade e diferenciabilidade de $f(\cdot, t)$ e $g(\cdot, t)$ tem-se

$$\int_0^T f(x(t), t) dt \geq \int_0^T \{f(\bar{x}(t), t) + \nabla f'(\bar{x}(t), t)[x(t) - \bar{x}(t)]\} dt, \quad (4.26)$$

$$g(x(t), t) \geq g(\bar{x}(t), t) + \nabla g'(\bar{x}(t), t)[x(t) - \bar{x}(t)], \quad (4.27)$$

q. s. em $[0, T]$, para todo $x \in X$.

Multiplicando a última desigualdade por $u(t) \geq 0$ q. s. em $[0, T]$, somando e integrando obtem-se

$$\int_0^T \bar{u}(t) g(x(t), t) dt \geq \int_0^T \{\bar{u}(t) g(\bar{x}(t), t) + \bar{u}(t) \nabla g'(\bar{x}(t), t)[x(t) - \bar{x}(t)]\} dt \text{ q. s. em } [0, T].$$

Adicionando o último resultado a (4.26), tem-se

$$\int_0^T \{f(x(t), t) + \bar{u}(t)g(x(t), t)\}dt \geq \int_0^T \{f(\bar{x}(t), t) + \bar{u}(t)g(\bar{x}(t), t)\}dt \\ + \int_0^T \{\nabla f'(\bar{x}(t), t) + \bar{u}(t)\nabla g'(\bar{x}(t), t)[x(t) - \bar{x}(t)]\}dt \quad \text{q. s. em } [0, T].$$

Tomando $h(t) = x(t) - \bar{x}(t)$, (4.24) implica que

$$\int_0^T \{f(x(t), t) + \bar{u}(t)g(x(t), t)\}dt \geq \int_0^T \{f(\bar{x}(t), t) + \bar{u}(t)g(\bar{x}(t), t)\}dt.$$

Portanto

$$L(\bar{x}, \bar{u}) \leq L(x, \bar{u}) \quad \text{para todo } x \in X.$$

Já que $g(\bar{x}(t), t) \leq 0$ e $u(t)g(\bar{x}(t), t) = 0$ q. s. em $[0, T]$, tem-se

$$L(\bar{x}, u) \leq L(\bar{x}, \bar{u}) \quad \text{para todo } u(t) \geq 0 \quad \text{q. s. em } [0, T].$$

Portanto, $(\bar{x}, \bar{u})$ satisfaz as condições de ponto de sela. ■

Observa-se que no teorema acima a propriedade de estabilidade para o problema (P) exige a validade das condições de Karush-Kuhn-Tucker, substituindo as condições de regularidade usualmente referidas como restrição de qualificação na literatura de programação matemática.

Uma melhor ilustração das condições de Karush-Kuhn-Tucker pode ser obtida considerando uma forma explícita da função g . Assim considere o seguinte problema :

$$\begin{aligned} \min \quad & \phi(x) = \int_0^T f(x(t), t)dt, \\ (\bar{P}) \quad & \text{s. a. } P(x(t), t) \leq r(t) + \int_0^t Q(x(s), t, s)ds \quad \text{q. s. em } [0, T], \\ & x \in X, \end{aligned}$$

onde f e X são definidos como no problema (P). P , r , e Q são funções a valores reais definidas em $V \times [0, T]$, $[0, T]$ e $V \times [0, T] \times [0, T]$, respectivamente. Então sob hipóteses apropriadas, (4.25) para o problema $(\bar{P})$, torna-se

$$\int_0^T [\nabla f'(\bar{x}(t), t)h(t) + \bar{u}(t)\nabla P'(\bar{x}(t), t)]h(t) - \int_0^T \bar{u}(t)\nabla Q'(\bar{x}(s), t, s)h(s)ds]dt = 0, \forall h \in \mathcal{L}_\infty^n[0, T].$$

Agora, aplicando o Teorema de Fubini para integral dupla nesta equação, obtem-se

$$\int_0^T h'(t) [\nabla f(\bar{x}(t), t) + \bar{u}(t) \nabla P(\bar{x}(t), t) - \int_t^T \bar{u}(s) \nabla Q(\bar{x}(t), t, s) ds] dt = 0, \quad \forall h \in \mathcal{L}_\infty^n[0, T],$$

o qual implica que a expressão dentro do colchetes é igual a zero q. s. em $[0, T]$. Por essa razão, a condição de otimalidade de ponto estacionário para o problema $(\bar{P})$ torna-se

$$\nabla f(\bar{x}(t), t) + \bar{u}(t) \nabla P(\bar{x}(t), t) - \int_t^T \bar{u}(s) \nabla Q(\bar{x}(t), t, s) ds = 0 \quad \text{q. s. em } [0, T],$$

$$\bar{u}'(t) [P(\bar{x}(t), t) - r(t) - \int_0^t Q(\bar{x}(s), t, s) ds] = 0 \quad \text{q. s. em } [0, T],$$

$$\bar{u}(t) \geq 0 \quad \text{q. s. em } [0, T].$$

4.2.4 Dualidade Lagrangiana

Nesta seção desenvolve-se uma dualidade conjunta entre o problema (P) e seu dual (D). O primeiro resultado, chamado de Teorema de dualidade fraca, é uma simples consequência da definição do problema dual. Contudo, verá-se que este teorema tem importantes implicações.

Teorema 4.2.5 (Teorema de dualidade fraca)

Sejam $\bar{x} \in \mathcal{L}_\infty^n[0, T]$ e $\bar{u} \in \mathcal{L}_\infty[0, T]$ sendo soluções factíveis do problema (P) e (D), respectivamente, então $\phi(\bar{x}) \geq \psi(\bar{u})$.

Demonstração :

Pela definição de ψ em (4.9), tem-se

$$\begin{aligned} \psi(\bar{u}) &= \inf_{x \in X} \int_0^T [f(x(t), t) + \bar{u}(t)g(x(t), t)] dt \\ &\leq \int_0^T f(\bar{x}(t), t) dt + \int_0^T \bar{u}(t)g(\bar{x}(t), t) dt \leq \phi(\bar{x}), \end{aligned}$$

desde que $\bar{u}(t) \geq 0$ e $g(\bar{x}(t), t) \leq 0$ q. s. em $[0, T]$. ■

Corolário 4.2.1 Se $\phi(\bar{x}) \leq \psi(\bar{u})$ para qualquer solução factível $\bar{x}$ do primal e qualquer solução $\bar{u}$ do dual, então $\bar{x}$ e $\bar{u}$ são soluções ótimas dos problemas (P) e (D), respectivamente.



Corolário 4.2.2 Se $\inf\{\phi(x) : x \in X, g(x(t), t) \leq 0 \text{ q. s. em } [0, T]\} = -\infty$ então $\psi(u) = -\infty$ para todo $u(t) \geq 0$ q. s. em $[0, T]$.

Corolário 4.2.3 Se $\sup\{\psi(u) : u \in \mathcal{L}_\infty[0, T], u(t) \geq 0 \text{ q. s. em } [0, T]\} = \infty$, então o problema primal não tem solução factível.

O Teorema de dualidade fraca será utilizado para demonstrar o lema a seguir, o qual é necessário para estabilizar o Teorema de dualidade forte, que verá-se posteriormente.

Lema 4.2.4 Suponha que o problema (P) tem uma solução ótima $\bar{x}$. Então $\bar{u} \in \mathcal{L}_\infty[0, T]$ é uma solução ótima do problema (D) e $\phi(\bar{x}) = \psi(\bar{u})$ se e somente se $-\bar{u} \in \partial p(0)$.

Demonstração :

($\Rightarrow$)

Seja $-\bar{u} \in \partial p(0)$, assim

$$p(y) \geq p(0) - \int_0^T \bar{u}(t) y(t) dt \text{ para todo } y \in Y.$$

Esta desigualdade, como viu-se na demonstração do Lema (4.2.3), implica que $\bar{u}(t) \geq 0$ q. s. em $[0, T]$, assim $\bar{u}$ é uma solução factível do problema (P). Ainda pela demonstração do Lema (4.2.3), tem-se

$$\int_0^T f(x(t), t) dt \geq p(\Gamma(x)) \geq p(0) - \int_0^T \bar{u}(t) g(x(t), t) dt \text{ para todo } x \in X,$$

onde $\Gamma(x)(t) = g(x(t), t)$. Assim

$$\int_0^T f(x(t), t) dt + \int_0^T \bar{u}(t) g(x(t), t) dt \geq p(0) \text{ para todo } x \in X.$$

Portanto

$$\psi(\bar{u}) = \inf_{x \in X} \int_0^T [f(x(t), t) + \bar{u}(t) g(x(t), t)] dt \geq p(0).$$

Já que $p(0) = \phi(\bar{x})$, tem-se que $\psi(\bar{u}) \geq \phi(\bar{x})$ e pelo Teorema (4.2.5) tem-se que $\phi(\bar{x}) \geq \psi(\bar{u})$. Assim $\psi(\bar{u}) = \phi(\bar{x})$. Pelo corolário (4.2.1), $\bar{u}$ é uma solução ótima do problema (D).

($\Leftarrow$)

Suponha que $\bar{u} \in \mathcal{L}_\infty[0, T]$ é uma solução ótima do problema (D). Já que por hipótese

$$\psi(\bar{u}) = \phi(\bar{x}) = p(0)$$

tem-se

$$\inf_{x \in X} \int_0^T [f(x(t), t) + \bar{u}(t)g(x(t), t)]dt = p(0).$$

Já que

$$\int_0^T \bar{u}(t)g(x(t), t)dt \leq \int_0^T \bar{u}(t)y(t)dt$$

para todo $x \in X$ e $y \in \Lambda_1^1[0, T]$ tal que $g(x(t), t) \leq y(t)$ q. s. em $[0, T]$, tem-se

$$\int_0^T [f(x(t), t) + \bar{u}(t)y(t)]dt \geq p(0)$$

para todo $x \in X$ e $y \in \Lambda_1^1[0, T]$ tal que $g(x(t), t) \leq y(t)$ q. s. em $[0, T]$. Disso pode-se concluir que

$$\inf_{x \in X} \left\{ \int_0^T f(x(t), t)dt : g(x(t), t) \leq y(t) \text{ q. s. em } [0, T] \right\} \geq p(0) + \int_0^T -\bar{u}(t)y(t)dt$$

para todo $y \in Y$. Assim

$$p(y) \geq p(0) + \int_0^T -\bar{u}(t)y(t)dt$$

para todo $y \in Y$. Portanto, $-\bar{u} \in \partial p(0)$. ■

Lema 4.2.5 Suponha que $p(0)$ é finito e $u \in \partial p(0)$, então $\bar{x} \in \mathcal{L}_\infty^n[0, T]$ é uma solução ótima do problema (P) se e somente se $(\bar{x}, -\bar{u})$ satisfaz as condições de otimalidade (a), (b) e (d) da definição (4.2.1).

Demonstração :

($\Rightarrow$)

Se $\bar{x}$ é uma solução ótima do problema (P) e $u \in \partial p(0)$, então pelo Lema (4.2.3), $-u$ é uma função multiplicador ótimo do problema (P), e pela definição (4.2.2), tem-se que $(\bar{x}, -\bar{u})$ satisfaz as condições de otimalidade do problema (P).

($\Leftarrow$)

Suponha que $(\bar{x}, -\bar{u})$ satisfaz as condições (a), (b) e (d) da definição (4.2.1). Vai-se mostrar que $(\bar{x}, -\bar{u})$ também satisfaz a condição (c) da definição (4.2.1).

Suponha que a condição (c) é violada, ou seja, que $-\bar{u}(t)$ é estritamente negativo em um subconjunto de $[0, T]$, com medida de Lebesgue positiva, assim

$$\int_0^T -\bar{u}(t)dt < 0.$$

Por hipótese tem-se

$$p(y) \geq p(0) + \int_0^T \bar{u}(t)y(t)dt \quad \text{para todo } y \in Y.$$

Se escolher $y(t) = 1$ e lembrando que a função p é não crescente, obtém-se

$$\int_0^T \bar{u}(t)dt \leq p(1) - p(0).$$

Multiplicando em ambos os lados por (-1) , obtém-se

$$\int_0^T -\bar{u}(t)dt \geq p(0) - p(1) \geq 0,$$

o que é uma contradição. Assim, $(\bar{x}, -\bar{u})$ satisfaz a condição (c) da definição (4.2.1). Como $(\bar{x}, -\bar{u})$ satisfaz as condições (a) - (d) da definição (4.2.1), tem-se que $\bar{x}$ é uma solução ótima do problema (P). ■

Teorema 4.2.6 (Teorema de dualidade forte)

Suponha que o problema (P) é estável. Então

- (1) *O problema (D) tem uma solução ótima;*
- (2) *Os valores ótimos dos problemas (P) e (D) são iguais;*
- (3) *$\bar{u} \in \mathcal{L}_\infty[0, T]$ é uma solução ótima do problema (D) se e somente se $-\bar{u} \in \partial p(0)$;*
- (4) *Toda solução ótima $\bar{u}$ do problema (D) caracteriza o conjunto de todas as soluções ótimas do problema (P) como os minimizadores de*

$$\int_0^T [f(x(t), t) + \bar{u}(t)g(x(t), t)]dt \quad \text{em } X,$$



e também satisfaz as restrições $g(x(t), t) \leq 0$ q. s. em $[0, T]$ e $\bar{u}(t)g(x(t), t) = 0$ q. s. em $[0, T]$.

Demonstração :

Como (P) é estável, então $p(0)$ é finito. Sabe-se que $p(0)$ é finito desde que o problema (P) tem uma solução ótima $\bar{x}$. Pelos lemas (4.2.1) e (4.2.2) garante-se a existência de um subgradiente de p em $y = 0$. Os itens (1), (2) e (3) seguem imediatamente do Lema (4.2.4) e o item (4) segue do item (3) e do Lema (4.2.5). ■

Conclusão

Estudou-se condições necessárias e suficientes de otimalidade de Fritz-John e Karush-Kuhn-Tucker para uma classe de problemas de programação não linear com tempo contínuo, no caso diferenciável.

Estas condições são análogas as conhecidas condições na área de programação não linear de dimensão finita e juntas assemelham-se ao critério Euler-Lagrange para o cálculo das variações.

Sabe-se bem que condições de otimalidade de Fritz-John e Karush-Kuhn-Tucker são de fundamental importância em muitos aspectos em programação não linear de dimensão finita.

Estudou-se a relação entre o critério de otimalidade de ponto estacionário e ponto de sela de Fritz-John e usando os conceito de perturbação e estabilidade, apresentou-se o critério de ponto estacionário e ponto de sela de Karush-Kuhn-Tucker e dualidade Lagrangiana para uma classe de problemas de programação convexa com tempo contínuo.

Estes resultados, encontrados em [23], são uma extensão dos resultados para problemas não lineares de dimensão finita dados em [7, 8].

Os estudos realizados neste trabalho podem ser utilizados como um estágio inicial para o estudo de problemas de programação fracional, função multiplicador de Lagrange generalizada, problemas de critério múltiplo, análise sensitiva, na construção de algoritmos computacionais e possibilita conhecer e pesquisar muitos outros aspectos de programação com tempo contínuo, inclusive, problemas mais complexos, tais como, problemas de controle ótimo e de cálculo das variações .



Referências Bibliográficas

- [1] ABRHAM, J. e BUIE, R. N. *Kuhn-Tucker conditions and duality in continuous programming*, Utilitas Math. 16, 15-37, 1979.
- [2] BAZARAA, M. S., SHERALI, H. D. e SHETTY, C. M. *Nonlinear Programming : Theory and Algorithms*, John Wiley & Sons, Inc., New York, 1993.
- [3] BELLMAN, R. *Bottleneck problems and dynamic programming*, Proc. Nat. Acad. Sci. U.S.A. 39, 947-951, 1953.
- [4] BELLMAN, R. *Dynamic Programming*, Princeton Univ. Press, Princeton, New Jersey, 1957.
- [5] DUNFORD, N. e SCHWARTZ, J. *Linear Operators, Part II*, Interscience, New York, 1958.
- [6] FARR, W. H. e HANSON, M. A. *Continuous time programming with nonlinear constraints*, J. Math. Anal. Appl. 45, 96-115, 1974.
- [7] GALE, D. *A geometric duality theorem with economic applications*, Rev. Econom. Stud. 34, 19-24, 1967.
- [8] GEOFFRION, A. M. *Duality in nonlinear programming : a simplified applications-oriented development*, SIAM Rev. 13, 1-37, 1971.
- [9] GEOFFRION, A. M. *Perspectives on optimization - A collection of expository articles*, University of California, Los Angeles, 1971.
- [10] GIRSANOV, I. V. *Lectures on Mathematical Theory of Extremum Problems*, Lecture Notes in Economics and Mathematical Systems 67, Springer-Verlag, New York, 1972.



- [11] GRINOLD, R. *Symetric duality for continuous linear programming*, SIAM J. Appl. Math. 18, 84-97, 1970.
- [12] HANSON, M. A. e MOND, B. *A class of continuous convex programming problems*, J. Math. Anal. Appl. 22, 427-437, 1968.
- [13] HOLMES, R. B. *Geometric Functional Analysis and its Applications*, Springer-Verlag, New York, 1975.
- [14] LEVINSON, N. *A class of continuous linear programming problems*, J. Math. Anal. Appl. 16, 73-83, 1966.
- [15] LUENBERGER, D. G. *Introduction to Linear and Nonlinear Programming*, Addison-Wesley, Reading, Mass., 1973.
- [16] MCLINDEN, L. *Duality theorems and theorems of the alternative*, Proc. Amer. Math. Soc. 53, 172-175, 1975.
- [17] PEREIRA, E. *Condições de Otimalidade e Dualidade em Programação não linear com tempo contínuo*. São José do Rio Preto, 1999. 80p. Dissertação (Mestrado em Matemática Aplicada) - Instituto de Biociências, Letras e Ciências Exatas, Universidade Estadual Paulista.
- [18] PONSTEIN, J. *Seven kinds of convexity*, SIAM Rev. 9, 115-119, 1967.
- [19] REILAND, T. W. e HANSON, M. A. *Generalized Kuhn-Tucker conditions and duality for continuous nonlinear programming problems*, J. Math. Anal. Appl. 74, 578-598, 1980.
- [20] REILAND, T. W. *Optimality and duality in continuous programming I. Convex programs and a theorem of the alternative*, J. Math. Anal. Appl. 77, 297-325, 1980.
- [21] Rockafellar, R. T. *Convex Analysis*, Princeton University Press, Princeton, New Jersey, 1970.



- [22] ROJAS-MEDAR, M. A., BRANDÃO, A. J. V. e SILVA, G. N. *Nonsmooth Continuous-Time Optimization Problems : Sufficient Conditions*, J. Math. Anal. Appl. 227, 305-318, 1998.
- [23] ZALMAI, G. J. *Optimality conditions and Lagrangian duality in continuous-time nonlinear programming*, Journal of mathematical analysis and Applications 109, 426-452, 1985.
- [24] ZALMAI, G. J. *A continuous-time generalization of Gordans's transposition theorem*, Journal of mathematical analysis and Applications 110, 130-140, 1985.
- [25] ZALMAI, G. J. *The Fritz-John and Kuhn-Tucker optimality conditions in continuous-time nonlinear programming*, Journal of mathematical analysis and Applications 110, 503-518, 1985.
- [26] ZALMAI, G. J. *Sufficient optimality conditions in continuous-time nonlinear programming*, Journal of mathematical analysis and Applications 111, 130-147, 1985.
- [27] ZANGWILL, W. I. *Nonlinear Programming : A Unified Approach*, Prentice-Hall, Englewood Cliffs, New Jersey, 1967.
- [28] WHEEDEN, R. L. e ZYGMUND, A. *Measure and Integral : An introduction to Real Analysis*, New York, 1977.



Apêndice A

O objetivo é demonstrar o Teorema de Gordan generalizado. Antes porém, vê-se alguns resultados que são necessários para a demonstração do Teorema de Gordan generalizado. Considere os seguintes problemas primal e dual de programação não linear de tempo contínuo:

$$\begin{aligned}
 (P) \quad & \min \quad \phi(x) = \int_0^T f(x(t), t) dt, \\
 & \text{s. a.} \quad g(x(t), t) \leq 0 \quad \text{q. s. em } [0, T], \\
 & \quad \quad h(x(t), t) = 0 \quad \text{q. s. em } [0, T], \\
 & \quad \quad x \in X,
 \end{aligned}$$

onde X é um subconjunto convexo não vazio de um espaço de Banach $\mathcal{L}_\infty^n[0, T]$, com $\|x\|_\infty$ definida como no Capítulo 2 e ϕ é uma função convexa a valor real definida em X . Sejam

$$g(x(t), t) = \Gamma(x)(t)$$

onde Γ é uma aplicação com domínio em X , tomando valores no espaço normado $\Lambda_1^m[0, T]$, com $\|y\|_1$ definida como no capítulo 2 e

$$h(x(t), t) = \eta(x)(t)$$

onde η é uma aplicação com domínio em X , tomando valores no espaço normado $\Lambda_1^k[0, T]$.

Supondo que a função g é convexa com relação a sua primeira variável por todo $[0, T]$ e a função h é linear com relação a sua primeira variável por todo $[0, T]$.

$$\begin{aligned}
 (D) \quad & \max \quad \psi(u, v), \\
 & \text{s. a.} \quad u(t) \geq 0 \quad \text{q. s. em } [0, T],
 \end{aligned}$$

onde

$$\psi(u, v) = \inf_{x \in X} \int_0^T [f(x(t), t) + u'(t)g(x(t), t) + v'(t)h(x(t), t)]dt, \quad (4.28)$$

$u \in \mathcal{L}_\infty^m[0, T]$ e $v \in \mathcal{L}_\infty^k[0, T]$.

Definição 4.2.8 A função perturbação $p : \Lambda_1^m[0, T] \times \Lambda_1^k[0, T] \rightarrow \mathbb{R}$ associado ao problema (P) é definida por:

$$p(y) = \inf_{x \in X} \left\{ \int_0^T f(x(t), t)dt : g(x(t), t) \leq y^1(t), h(x(t), t) = y^2(t) \text{ q. s. em } [0, T] \right\},$$

onde $y = (y^1, y^2)$.

Teorema 4.2.7 [24] Sejam X um subconjunto convexo não vazio de $\mathcal{L}_\infty^n[0, T]$ e $g : \mathbb{R}^n \times [0, T] \rightarrow \mathbb{R}$ definida por $g(x(t), t) = \Gamma(x)(t)$, onde Γ é uma aplicação de X em $\Lambda_1^1[0, T]$, sendo convexa com relação a sua primeira variável por todo $[0, T]$ e $h : \mathbb{R}^n \times [0, T] \rightarrow \mathbb{R}$ definida por $h(x(t), t) = \eta(x)(t)$, onde η é uma aplicação de X em $\Lambda_1^1[0, T]$, sendo linear com relação a sua primeira variável por todo $[0, T]$. Se o sistema

$$\begin{cases} g(x(t), t) < 0 \text{ q. s. em } [0, T], \\ h(x(t), t) = 0 \text{ q. s. em } [0, T], \end{cases}$$

não tem solução $x \in X$, então existe $u \in \mathcal{L}_\infty^m[0, T]$ e $v \in \mathcal{L}_\infty^k[0, T]$ tal que

$$\int_0^T [u'(t)g(x(t), t) + v'(t)h(x(t), t)]dt \geq 0 \text{ para todo } x \in X,$$

$$u(t) \geq 0 \text{ q. s. em } [0, T],$$

$$(u(t), v(t)) \neq 0 \text{ q. s. em } [0, T].$$

Demonstração :

Se o sistema não tem solução $x \in X$, então para qualquer solução factível, o problema convexo

$$\begin{aligned} \min \quad & \lambda, \\ \text{s. a.} \quad & g(x(t), t) - \lambda \leq 0 \text{ q. s. em } [0, T], \\ & h(x(t), t) - \lambda = 0 \text{ q. s. em } [0, T], \\ & \lambda \in \mathbb{R}, x \in X, \end{aligned}$$

tem solução $\lambda \geq 0$.

Este problema é estável.

De fato:

(a) Verificar que $p(0)$ é finito. Seja

$$p(y) = \inf_{\substack{x \in X, \\ \lambda \in \mathbb{R}}} \left\{ \lambda : g(x(t), t) - \lambda \leq y^1(t), h(x(t), t) - \lambda = y^2(t), \text{ q. s. em } [0, T] \right\},$$

onde $y = (y^1, y^2)$.

A função p é uma aplicação $p : \Lambda_1^m[0, T] \times \Lambda_1^k[0, T] \rightarrow \mathbb{R}$. Assim

$$p(0) = \inf_{\substack{x \in X, \\ \lambda \in \mathbb{R}}} \left\{ \lambda : g(x(t), t) - \lambda \leq 0, h(x(t), t) - \lambda = 0, \text{ q. s. em } [0, T] \right\}.$$

Como λ é solução do problema convexo, tem-se que $\lambda \geq 0$ e portanto finito.

(b) Verificar que existe $m > 0$ tal que

$$p(0) \leq p(y) + m\|y\|_1,$$

para todo $y = (y^1, y^2) \in \Lambda_1^m[0, T] \times \Lambda_1^k[0, T]$.

Já que $p(y)$ é uma função convexa, então por definição tem-se que o subgradiente de $p(0)$ é diferente de vazio. Assim, se $\partial p(0) \neq \emptyset$, pelo Lema (4.2.2), existe $m > 0$ tal que

$$p(0) \leq p(y) + m\|y\|_1,$$

para todo $y = (y^1, y^2) \in \Lambda_1^m[0, T] \times \Lambda_1^k[0, T]$.

De (a) e (b) tem-se que o problema convexo é estável, e assim, pelo Teorema (4.2.6), o problema dual

$$\max_{\substack{x \in X, \\ \lambda \in \mathbb{R}}} \left\{ \lambda + \int_0^T u'(t)[g(x(t), t) - \lambda]dt + \int_0^T v'(t)[h(x(t), t) - \lambda]dt \right\},$$

tem solução ótima (u, v) , com $(u, v) \in \mathcal{L}_\infty^m[0, T] \times \mathcal{L}_\infty^k[0, T]$ e $u(t) \geq 0$ q. s. em $[0, T]$, cujo valor ótimo é não negativo. Assim, fazendo algumas operações algébricas simples, pode-se concluir que

$$\inf_{\substack{x \in X, \\ \lambda \in \mathbb{R}}} \left\{ \lambda + \int_0^T u'(t)[g(x(t), t) - \lambda]dt + \int_0^T v'(t)[h(x(t), t) - \lambda]dt \right\}$$

$$= \inf_{x \in X} \int_0^T [u'(t)g(x(t), t) + v'(t)h(x(t), t)]dt + \inf_{\lambda \in \mathbb{R}} \lambda \left\{ 1 - \int_0^T [u'(t) + v'(t)]dt \right\} \geq 0.$$

Devido a natureza de λ , deve-se ter

$$1 - \int_0^T [u'(t) + v'(t)]dt = 0.$$

Assim

$$\int_0^T [u'(t) + v'(t)]dt = 1$$

e

$$\int_0^T [u'(t)g(x(t), t) + v'(t)h(x(t), t)]dt \geq 0$$

para todo $x \in X$. ■

Teorema 4.2.8 [24] (Teorema de Gordan generalizado)

Seja X um subconjunto convexo não vazio de $\mathcal{L}_\infty^n[0, T]$. Seja a aplicação $p : V \times [0, T] \longrightarrow \mathbb{R}^m$, definida por $p(x(t), t) = \pi(x)(t)$, onde π é uma aplicação de X em $\Lambda_1^m[0, T]$ e supondo que p é convexo com relação a sua primeira variável sobre V por todo $[0, T]$, então

(a) existe $x \in X$ tal que:

$$p(x(t), t) < 0 \text{ q. s. em } [0, T]$$

ou

(b) existe uma função m – vetorial não nula $u \in \mathcal{L}_\infty^m[0, T]$, $u(t) \geq 0$ q. s. em $[0, T]$ tal que:

$$\int_0^T u'(t)p(x(t), t)dt \geq 0 \text{ para todo } x \in X.$$

Mas nunca ambos.

Demonstração :

Suponha que (a) tem solução $\bar{x} \in X$. Então para qualquer

$$u \in \mathcal{L}_\infty^m[0, T], \quad u(t) \geq 0, \quad u(t) \neq 0 \text{ q. s. em } [0, T]$$

tem-se

$$\int_0^T u'(t)g(\bar{x}(t), t)dt < 0.$$

Assim (b) não pode ter solução .

Agora, suponha que (a) não tem solução $x \in X$. Então pelo Teorema (4.2.7), excluindo

$$h(x(t), t) = 0 \text{ q. s. em } [0, T]$$

têm-se que existe $u \in \mathcal{L}_\infty^m[0, T]$ tal que

$$\int_0^T u'(t)g(x(t), t)dt \geq 0 \text{ para todo } x \in X;$$

$$u(t) \geq 0 \text{ q. s. em } [0, T];$$

$$u(t) \neq 0 \text{ q. s. em } [0, T].$$

Assim (b) tem solução .

■

