

Célia Maria Souza Pires

Programação Linear Contínua





1910000761



PROGRAMAÇÃO LINEAR CONTÍNUA

Programação Linear
Contínua

Célia Maria Souza Pires

Dissertação de Mestrado
Pós-Graduação em Matemática Aplicada e
Computacional.
MAC - 025

Orientador: Prof. Dr. Geraldo Nunes Silva

São José do Rio Preto

setembro de 1997



761/48

PROGRAMAÇÃO LINEAR CONTÍNUA

Célia Maria Souza Pires

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 Ciências Exatas (Área de Concentração: Matemática Aplicada e Computacional).

Orientador: Prof. Dr. Geraldo Nunes Silva

São José do Rio Preto

setembro de 1997



761/98

14



“O saber é um presente *Divino*, agradeço a *Ele* esse privilégio.”

minha filha Meus pais!

e Dimitar.

Agradecimentos

Ao Prof. Dr. Geraldo por me aceitar como orientanda.

A Profa. Dra. Aldenice que pela sua sábia e firme mão me conduziu este trabalho.

Ao Prof. Dr. Borges pelo seu carinho.

A direção, professores e funcionários da minha escola "Cardenal Leme" pela compreensão,
em especial à Prof. Guiomar.

Para meus pais,
minha filha Fernanda

e Dimitar.

A todos os professores, amigos pós-graduandos, mesmo aqueles que já não estão mais aqui,
e funcionários, em especial ao Sr. Salvador, do DCCB, que de alguma forma contribuíram
para esta realização.

Aos meus amigos, Telcio, Marli, Adélia, Myrian e Adip, pela dedicação.

Resumo

Três métodos discretos para resolver aproximadamente alguns problemas de programação linear contínua são considerados. Os métodos para resolver o problema discreto correspondente são dados e exemplos numéricos são fornecidos. É provado que as soluções dos problemas discretos convergem, pelo menos para quase todo ponto, para as soluções dos problemas contínuos. Além disso, teoremas de dualidade forte e fraca são estabelecidos para alguns problemas lineares com tempo contínuo.

Agradecimentos

Ao Prof. Dr. Geraldo por me aceitar como orientanda.

À Profa.Dra. Aldenice que pela sua sábia e firme mão me conduziu este trabalho.

Ao Prof. Dr. Borges pelo seu carinho.

A direção , professores e funcionários da minha escola "Cardeal Leme" pela compreensão, em especial à Prof Guiomar.

A todos os professores, amigos pós-graduandos, mesmo aqueles que já' não estão mais aqui, e funcionários, em especial ao Sr. Salvador, do DCCE, que de alguma forma contribuíram para esta realização.

Aos meu amigos, Telcio, Marli, Adélia, Myrian e Adip, pela dedicação.

Resumo

Três métodos discretos para resolver aproximadamente alguns problemas de programação linear contínua são considerados. Os métodos para resolver o problema discreto correspondente são dados e exemplos numéricos são fornecidos. É provado que as soluções dos problemas discretos convergem, pelo menos para quase todo ponto, para as soluções dos problemas contínuos. Além disso, teoremas de dualidade forte e fraca são estabelecidos para alguns problemas lineares com tempo contínuo.



Abstract

Three discrete methods for solving approximately some problems of Linear Continuous Programming are considered. The ways of discretizing the continuous problems are discussed. The methods of solving the corresponding discret problem are given and numerical examples are provided. It is proved that the solutions of the discrete problems converge, at least almost everywhere, to the solutions of the continuous problems. Furthermore, weak and strong duality results are established for certain continuous time linear problems.

1	<i>Introdução</i>	8
2	<i>Preliminares</i>	13
2.1	<i>Definições e Resultados Básicos</i>	13
2.2	<i>Notações</i>	26
3	<i>Resolução da PPLC por aproximação via discretização</i>	28
3.1	<i>Introdução</i>	28
3.2	<i>Método de aproximação por integração numérica</i>	29
3.3	<i>Convergência e Factibilidade</i>	32
4	<i>Programação linear contínua separável</i>	38
4.1	<i>Introdução</i>	38
4.2	<i>O Problema Dual</i>	39
4.3	<i>Discretização</i>	41
4.4	<i>Construindo uma solução melhorada</i>	56
4.5	<i>Implementação</i>	61
5	<i>Resultados numéricos e considerações finais</i>	64
5.1	<i>Exemplo não separável</i>	64
5.2	<i>Exemplo separável</i>	67
5.3	<i>Conclusões</i>	68



Índice

Introdução

1	<i>Introdução</i>	8
2	<i>Preliminares</i>	13
2.1	<i>Definições e Resultados Básicos</i>	13
2.2	<i>Notações</i>	26
3	<i>Resolução do PPLC por aproximação via discretização</i>	28
3.1	<i>Introdução</i>	28
3.2	<i>Método de aproximação por integração numérica</i>	29
3.3	<i>Convergência e Factibilidade</i>	32
4	<i>Programação linear contínua separável</i>	38
4.1	<i>Introdução</i>	38
4.2	<i>O Problema Dual</i>	39
4.3	<i>Discretização</i>	41
4.4	<i>Construindo uma solução melhorada</i>	56
4.5	<i>Implementação</i>	61
5	<i>Resultados numéricos e considerações finais</i>	64
5.1	<i>Exemplo não separável</i>	64
5.2	<i>Exemplo separável</i>	67
5.3	<i>Conclusões</i>	68



Capítulo 1

Introdução

Os primeiros computadores digitais surgiram a partir de pesquisas de John Atanasov e John Von Newman, entre outros, nos anos 50. Simultaneamente ao desenvolvimento rápido desse novo tipo de máquina, foram feitas as primeiras tentativas de aplicar novos métodos para resolver problemas que surgem em situações reais e que requerem um número grande de cálculos. Alguns problemas reais deste tipo são: modelos econômicos, tais como investimento e planejamento; controle ótimo, como fixar horários em shopping; além dos problemas de rede [5] e [8]. Estes novos métodos têm também aplicações na engenharia [20].

Todas estas questões práticas podem ser modeladas matematicamente por problemas de programação linear contínuo (PPLC). Isto resultará em uma classe de problemas de otimização. Especificamente falando, estes problemas de programação linear contínuo têm sido estudados como generalização de programação linear para espaços de funções, e são formulados da seguinte maneira:

$$\begin{aligned} \text{(PLC)} \quad & \text{maximizar} \quad \int_0^T c^T(t)x(t)dt \\ & \text{sujeito à} \quad B(t)x(t) \leq a(t) + \int_0^t K(t,s)x(s)ds, \\ & \quad x(t) \geq 0, \quad t \in [0, T], \end{aligned}$$

onde $B(t)$, $K(t, s)$ são matrizes reais $M \times N$, $c(t)$ e $a(t)$ são, respectivamente, vetores de $\mathbb{R}^N$ e $\mathbb{R}^M$. Refere-se a este problema como problema primal.

Este problema se refere a problema primal e, foi introduzido por Bellmann [6], em 1953, onde foram tratados de “Problemas de gargalo” em processos de produção linear multi-



estágios.

Associa-se ao primal a seguinte formulação dual:

$$\begin{aligned} (\text{PLC}^*) \quad & \text{minimizar} \quad \int_0^T a^T(t)w(t)dt \\ & \text{sujeito à} \quad B'(t)w(t) \geq c(t) + \int_t^T K'(s,t)w(s)ds, \\ & \quad w(t) \geq 0, \quad t \in [0, T]. \end{aligned}$$

Bellman estabeleceu a dualidade fraca. Desde então, seus resultados foram estendidos em duas áreas: teorema da dualidade forte e métodos computacionais.

Um teorema da dualidade forte foi estabelecido por Tyndall [21] e Levinson [16], e subsequentemente estendido por outros autores, por exemplo, Grinold [11], Levine e Pomerol [15].

No lado computacional, existem duas abordagens principais:

1. Pura aproximação por discretização.
2. Generalização do método simplex de dimensão finita para o problema linear com tempo contínuo.

A primeira abordagem busca a solução do problema linear contínuo via aproximação por problemas discretos, cujas soluções ótimas convergem para solução do PLC (Buie e Abraham [7]). A discretização é feita particionando o intervalo $[0, T]$ em um número finito de subintervalos, normalmente equidistantes. Quando a partição é refinada, as soluções das discretizações convergem para a solução do problema original. Este método de discretização, embora funcione bem, deixa de levar em consideração a natureza de dimensão infinita do problema em questão.

A segunda abordagem que busca estender o método simplex finito dimensional para problemas de dimensão infinita, foi iniciado por Lehman [14] e continuada por Drews [9], Hartberger [12], e Segers [23], mas o maior progresso foi feito por Perold [18] e continuado por Anstreicher [3].

Considerando uma classe mais restrita de problemas PPLC, chamada de problemas de programação linear contínua separáveis, PPLCS, Anderson [1] apresentou um algoritmo primal-dual e resolveu alguns exemplos simples. Mas uma implementação global do algoritmo não foi possível.



Já Anderson e Philpott [2], em 1989, restringiram ainda mais a classe de problemas a problemas de redes e desenvolveram, pela primeira vez, um algoritmo implementável em computador capaz de resolver tais problemas e que levava em consideração o caráter de dimensão infinita destes.

Em 1993, Pullan[19], fazendo uma combinação dos dois caminhos descritos acima, desenvolveu um algoritmo, que, de alguma forma, contém os elementos da teoria de Tyndall[21], para resolver o PPLCS, e obteve um teorema de dualidade forte.

Este trabalho tem por objetivo discutir os métodos de discretização devido a Buie e Abrham para PLC geral e os ingredientes necessários para implementação de algoritmo híbridos (simplex + pura discretização) para a classe especial de problemas lineares com tempo contínuo separável PLCS (Pullan). Além disso, resultados numéricos de exemplos simples são apresentados

Definições e alguns teoremas necessários ao desenvolvimento desta dissertação são apresentados no capítulo 2.

No terceiro capítulo discute-se o método de discretização devido a Buie e Abrham, em que o problema primal transforma-se em problema discreto, obtendo-se a seguinte formulação :

$$\begin{aligned} (\mathbf{P}_n) \quad & \text{maximizar} \quad c_n^T x_n \\ & \text{sujeito à} \quad G_n x_n \leq a_n, \\ & \quad \quad \quad x_n \geq 0, \quad t \in [0, T], \end{aligned}$$

cujas soluções convergem para uma solução do problema primal, numa topologia adequada.

No quarto capítulo considera-se uma subclasse do PPLC, chamada “Programa Linear Contínuo Separável”, (**PPLCS**):

$$\begin{aligned} (\mathbf{PPLCS}) \quad & \text{minimizar} \quad \int_0^T c^T(t)x(t)dt \\ & \text{sujeito à} \quad \int_0^t Gx(s)ds + y(t) = a(t), \end{aligned} \tag{1.1}$$

$$Hx(t) + z(t) = b(t), \tag{1.2}$$

$$x(t), y(t), z(t) \geq 0, \quad t \in [0, T], \tag{1.3}$$



onde $x(t)$, $z(t)$, $b(t)$ e $c(t)$ são funções limitadas e mensuráveis e $y(t)$ e $a(t)$ são funções absolutamente contínuas. As dimensões de $x(t)$, $y(t)$ e $z(t)$ são n_1 , n_2 e n_3 , respectivamente, G é uma matriz $n_2 \times n_1$ e H é uma matriz $n_3 \times n_1$.

O significado da palavra separável refere-se ao fato de que as restrições se dão em dois conjuntos, a equação integral (1.1) e as restrições instantâneas (1.2) e (1.3).

O problema linear dual (**PLCS***), para PLCS, é formulado da seguinte maneira:

$$\begin{aligned}
 (\text{PLCS}^*) : \quad & \text{maximizar} \quad - \int_0^T d\pi'(t)a(t) - \int_0^T \eta'(t)b(t)dt \\
 & \text{sujeito à} \quad c(t) - G'\pi(t) + H'\eta(t) \geq 0, \\
 & \eta(t) \geq 0, \quad \text{qtp em } [0, T],
 \end{aligned}$$

onde as componentes de $\eta(t)$ pertencem a $L_1[0, T]$, e cada componente de $\pi(t)$ é monotônica crescente e contínua à direita em $[0, T]$ com $\pi(T) = 0$.

A solução para o problema separável, dá-se através da resolução do problema discreto correspondente ao *PPLCS*, nas quais estas transformações ocorrem naturalmente da discretização do intervalo $[0, T]$, isto é, particiona-se o intervalo, não necessariamente equidistante e usa-se a fórmula de quadratura do trapézio nas integrais. Duas discretizações são exibidas.

A primeira discretização ocorre naturalmente do PPLCS para funções $x(t)$, $y(t)$ e $z(t)$ tais que $x(t)$ e $z(t)$ sejam constantes por partes e $y(t)$ seja linear por partes. Para que o procedimento tenha sucesso é necessário que as funções $a(t)$ e $c(t)$ sejam lineares por partes, a função $b(t)$ seja constante por partes em $[0, T]$ e uma partição P fixa.

A segunda discretização é, na verdade, a mesma que a primeira, porém com os vetores $b(t)$ e $c(t)$ tendo o dobro da dimensão dos vetores da primeira discretização, obtém-se das restrições uma matriz $\hat{A}$ apropriada, dependente da partição escolhida. Neste mesmo capítulo mostra-se uma forma equivalente à segunda discretização, onde com algumas substituições nas funções $x(t)$, $y(t)$ e $z(t)$, obtém-se uma nova matriz das restrições, que não dependerá da partição, e sim das dimensões das matrizes dadas, G e H .

Ainda neste capítulo mostra-se como construir uma solução melhorada para PLCS, se existir uma solução inicial factível para PLCS. Isto formará a base de um algoritmo para resolver PLCS.

Finalmente, no capítulo 5, discutem-se os resultados computacionais baseados na dis-



cretização de Buie e Abrham e no algoritmo de Pullan, que foram implementados usando Mathematica e CPLEX e a conclusão do trabalho é feita tecendo-se as considerações finais.

Capítulo 2

Preliminares

Neste capítulo apresentam-se noções básicas da Teoria da Medida [13] e de programação linear clássica [4] necessárias às discussões posteriores desta dissertação.

2.1 Definições e Resultados Básicos

DEFINIÇÃO: 2.1.1 Partição e Norma

1. Diz-se que $P = \{t_0, \dots, t_m\}$ é uma partição de $[0, T]$ se

$$0 = t_0 < t_1 < \dots < t_m = T.$$

2. Uma partição Q de $[0, T]$ é um refinamento da partição P de $[0, T]$ se $P \subseteq Q$.

3. A norma, $\|P\|$, da partição $P = \{t_0, \dots, t_m\}$, define-se por

$$\|P\| = \max_i (t_i - t_{i-1})$$

DEFINIÇÃO: 2.1.2 Função constante por partes

Diz-se que uma função $f(t)$ é constante por partes (linear) com a partição $P = \{t_0, \dots, t_m\}$ se $f(t)$ é constante (linear) em $[t_{i-1}, t_i]$ para $i = 1, \dots, m$.

Dizemos que a função f é constante por partes (linear) em $[0, T]$ se f é constante por partes (linear) com alguma partição $[0, T]$.



1. Seja $P = \{t_0, \dots, t_m\}$ partição de $[0, T]$.

Supõe-se que haja um conjunto de $m + 1$ variáveis, $f(t_0), f(t_1), \dots, f(t_m)$, então a função $f(t)$ é definida por

Capítulo 2

Preliminares

Neste capítulo apresentam-se noções básicas da Teoria da Medida [13] e de programação linear clássica [4] necessárias às discussões posteriores desta dissertação.

2.1 Definições e Resultados Básicos

DEFINIÇÃO: 2.1.1 Partição e Norma

1. Diz-se que $P = \{t_0, \dots, t_m\}$ é uma partição de $[0, T]$ se

$$0 = t_0 < t_1 < \dots < t_m = T.$$

2. Uma partição Q de $[0, T]$ é um refinamento da partição P de $[0, T]$ se $P \subseteq Q$.

3. A norma, $\|P\|$, da partição $P = \{t_0, \dots, t_m\}$, define-se por

$$\|P\| = \max_i (t_i - t_{i-1})$$

DEFINIÇÃO: 2.1.2 Função constante por partes

Diz-se que uma função $f(t)$ é constante por partes(linear) com a partição $P = \{t_0, \dots, t_m\}$, se $f(t)$ é constante (linear) em $[t_{i-1}, t_i]$ para $i = 1, \dots, m$.

Dizemos que a função f é constante por partes (linear) em $[0, T]$ se f é constante por partes (linear) com alguma partição $[0, T]$.

DEFINIÇÃO: 2.1.3 Extensão por partes

1. Seja $P = \{t_0, \dots, t_m\}$ partição de $[0, T]$.

Supõe-se que haja um conjunto de $m + 1$ variáveis, $\hat{f}(t_0), \hat{f}(t_1), \dots, \hat{f}(t_m)$, então a função $f(t)$ é definida por

$$f(t) = \left(\frac{t_i - t}{t_i - t_{i-1}}\right)\hat{f}(t_{i-1}) + \left(\frac{t - t_{i-1}}{t_i - t_{i-1}}\right)\hat{f}(t_i), \text{ para } t \in [t_{i-1}, t_i], \text{ para } i = 1, \dots, m,$$

é chamada de extensão linear por partes de $\hat{f}$.

2. Similarmente, se houver um conjunto de $2m$ variáveis, $\hat{f}(t_0+), \hat{f}(t_1-), \hat{f}(t_1+), \dots, \hat{f}(t_m-)$, então a função $f(t)$ é definida por

$$f(t) = \begin{cases} \hat{f}(t_i+) & \text{se } t = t_0, \dots, t_{m-1}, \\ 0 & \text{se } t = T, \\ \left(\frac{t_i - t}{t_i - t_{i-1}}\right)\hat{f}(t_{i-1}+) + \left(\frac{t - t_{i-1}}{t_i - t_{i-1}}\right)\hat{f}(t_i-), & \text{para } t \in (t_{i-1}, t_i), \text{ para } i = 1, \dots, m, \end{cases}$$

é chamada de extensão linear por partes de $\hat{f}$.

3. Finalmente, se houver um conjunto de m variáveis, $\hat{f}(t_0+), \hat{f}(t_1+), \dots, \hat{f}(t_m+)$, então a função $f(t)$ definida por

$$f(t) = \begin{cases} \hat{f}(t_{m-1}+) & , \quad t = T \\ \hat{f}(t_{i-1}+) & , \quad t \in [t_{i-1}, t_i) \text{ para } i = 1, \dots, m, \end{cases}$$

é chamada de extensão constante por partes de $\hat{f}$.

DEFINIÇÃO: 2.1.4 Função de variação limitada

Uma função f definida num segmento se chama função de variação limitada quando existe uma constante C tal que, qualquer que seja a partição do segmento $[a, b]$ por pontos

$$a = x_0 < x_1 < \dots < x_n = b,$$

vale a desigualdade

$$\sum_{k=1}^n |f(x_k) - f(x_{k-1})| \leq C. \quad (2.1)$$

Toda função monótona é de variação limitada, já que a soma que aparece em (2.1) não depende da partição, e é sempre igual a $|f(b) - f(a)|$.

DEFINIÇÃO: 2.1.5 Variação total

Seja f uma função de variação limitada.

O supremo das somas (2.1) correspondente a todas as partições finitas do segmento $[a, b]$ se denomina variação total da função f no segmento $[a, b]$, denotado por $V_a^b[f]$.

De maneira que :

$$V_a^b[f] = \sup \sum_{k=1}^n |f(x_k) - f(x_{k-1})|.$$

Obs: O significado intuitivo de $V_a^b[f]$ é o comprimento total do caminho percorrido por $f(t)$ quando t cresce de a até b .

TEOREMA 2.1.1 (Teorema de Helly)

Seja $\{f_n\}$ uma sequência de funções limitadas em a , ($|f_n(a)| \leq M$) e de variação total uniformemente limitada em $[a, b]$, ($V_a^b(f_n) \leq M$). Então, existe uma subsequência de funções que convergem em $[a, b]$ para uma função de variação limitada.

DEFINIÇÃO: 2.1.6 Dualidade

Dado um programa linear(PL) na forma canônica e seu respectivo programa linear dual(PLD).

PL	PLD
minimizar cx	maximizar wb
sujeito a $Ax \leq b,$	sujeito a $wA \geq c,$
$x \geq 0.$	$w \geq 0.$

onde, A é uma matriz $m \times n$, dada; b um vetor m -coluna, dado; c um vetor n -linha, dado; x vetor n -coluna em PL e w vetor n -linha em PLD, desconhecidos.

Relações entre o par primal-dual.

A cada problema (primal) corresponde um outro (dual), formando o par de problema dual:

$$(\text{Problema Primal}) \longleftrightarrow (\text{Problema Dual})$$



Correspondência uma a uma, entre : m e N uma matriz $m \times (n - m)$. O ponto

$$\left(\begin{array}{l} \text{variável do primal} \leftrightarrow \text{restrição do dual} \\ \text{restrição do primal} \leftrightarrow \text{variável do dual} \end{array} \right)$$

Resultados completos podem ser vistos na Tabela 2.1 com mais detalhes.

Tabela 2.1: tabela geral

	problema de minimização	problema de maximização	
variável	≥ 0 ≤ 0 irrestrita	≤ 0 ≥ 0 =	restrição
restrição	≥ 0 ≤ 0 =	≥ 0 ≤ 0 irrestrita	variável

$$cx_0 \leq \text{val.}$$

DEFINIÇÃO: 2.1.7 Soluções factíveis

Considerando o programa linear (PL) na forma canônica:

DEFINIÇÃO: 2.1.9 Conjunto Convexo

Um conjunto $A \subseteq \mathbb{R}^n$ é chamado de conjunto convexo se, para quaisquer dois pontos $x, y \in A$, o segmento de reta que os conecta também está contido em A .

dados $x_1, x_2 \in A$ temos,

Cada vetor x que satisfaz as restrições

$$Ax_1 + (1 - \alpha)x_2 \in A, \quad \forall 0 \leq \alpha \leq 1$$

$$Ax \leq b, \quad x \geq 0$$

DEFINIÇÃO: 2.1.10 Extremidade convexa

é chamada de solução factível.

DEFINIÇÃO: 2.1.8 Soluções básicas factíveis

Considerando o sistema $Ax = b$ e $x \geq 0$, onde A é uma matriz $m \times n$ e b é um vetor m . Supor que $\text{rank}(A, b) = \text{rank}(A) = m$ Depois reagrupando as colunas de A , seja $A = [B, N]$,



onde B é uma matriz inversível $m \times m$ e N uma matriz $m \times (n - m)$. O ponto

$$\text{conv}S = \left\{ y : y = \sum_{i=1}^p \lambda_i x_i, \sum_{i=1}^p \lambda_i = 1, p \in \mathbb{N} \right\}$$

onde

$$x_B = B^{-1}b$$

$$x_N = 0$$

é chamado de solução básica do sistema.

Se $x_B \geq 0$, então x é chamado de solução básica factível do sistema. Aqui B é chamada de matriz básica (ou simplesmente de base) e N é chamada de matriz não básica. Os componentes de x_B são chamados de variáveis básicas, e os componentes de x_N são chamados de variáveis não básicas.

TEOREMA 2.1.2 Seja (PL) e (PLD) par de programas linear primal-dual, para cada par de solução factível x_0, w_0 de PL e PLD respectivamente, tem-se

$$cx_0 \leq w_0b.$$

COLORÁRIO 2.1.1 Seja x_0 solução factível para PL, e seja w_0 solução factível para PLD; Então se $cx_0 = w_0b$, x_0 e w_0 são soluções ótimas de PL e PLD respectivamente.

DEFINIÇÃO: 2.1.9 Conjunto Convexo

Um conjunto $A \in \mathbb{R}^n$ diz-se ser convexo se:
dados $x_1, x_2 \in A$ temos,

$$\alpha x_1 + (1 - \alpha)x_2 \in A, \quad \forall 0 \leq \alpha \leq 1$$

DEFINIÇÃO: 2.1.10 Envoltória convexa

A intersecção de todos os conjuntos convexos contendo um dado subconjunto $S \in \mathbb{R}^n$ é chamado de envoltória convexa de S denota-se por conv S , e é dada por:

$$\text{conv}S = \bigcap_{\alpha \in \Lambda} A_\alpha, \quad \alpha \in \Lambda, \quad S \subset A_\alpha$$

ou equivalentemente

$$\text{conv}S = \{y : y = \sum_{i=0}^p \lambda_i a_i; \lambda_i \geq 0; \sum_{i=0}^p \lambda_i = 1, p \in N\}.$$

Anéis e Álgebra de conjuntos

DEFINIÇÃO: 2.1.11 Sistema de conjunto

Chama-se sistemas de conjuntos a todo conjunto cujos elementos são, em si, "certos" conjuntos.

Denotaremos por letras góticas maiúsculas os sistemas de conjuntos.

DEFINIÇÃO: 2.1.12 Anel de conjuntos

Um sistema não vazio de conjuntos $\mathfrak{R}$ chama-se anel se $\forall A, B \in \mathfrak{R}$ tem-se

$$A \Delta B \in \mathfrak{R} \text{ e } A \cap B \in \mathfrak{R},$$

onde Δ indica a diferença simétrica de conjuntos.

A união e a diferença de conjuntos também pertencem a $\mathfrak{R}$, pois se $A \in \mathfrak{R}$ e $B \in \mathfrak{R}$ então

$$A \cup B = (A \Delta B) \Delta (A \cap B)$$

e

$$A \setminus B = A \Delta (A \cap B),$$

onde $\setminus$ é o símbolo que indica a diferença de conjuntos.

Todo anel contém o conjunto vazio $\emptyset$, pois $A \setminus A = \emptyset$.

DEFINIÇÃO: 2.1.13 Unidade do sistema de conjuntos

Um conjunto E chama-se unidade do sistema de conjuntos $\mathcal{Q}$ se $E \in \mathcal{Q}$ e se para todo $A \in \mathcal{Q}$ verifica-se a igualdade

$$A \cap E = A, \quad (A \subset E)$$

onde E nada mais é que o conjunto maximal deste sistema que contém todos os demais conjuntos que estão em $\mathcal{Q}$.

Um anel de conjuntos provido da unidade denomina-se Álgebra de conjuntos.

Álgebra de Borel

DEFINIÇÃO: 2.1.14 Semianel de conjuntos

Um sistema de conjuntos $\mathcal{Q}$ chama-se semianel se contém o conjunto vazio $\emptyset$, é fechado em relação a intersecção e satisfaz a seguinte propriedade:

Se A e $A_1 \subset A$ pertencem a $\mathcal{Q}$, pode-se representar o conjunto A na forma

$$A = \bigcup_{k=1}^n A_k,$$

onde A_k são conjuntos de $\mathcal{Q}$ disjuntos dois a dois e o primeiro deles é o conjunto dado A_1 .

DEFINIÇÃO: 2.1.15 Decomposição finita

Todo sistema de conjuntos disjuntos dois a dois $A_1, A_2, \dots, A_n$, cuja união é um conjunto dado A chama-se, decomposição finita do conjunto A .

Todo anel de conjuntos $\mathfrak{R}$ é um semianel porque se A é simétrico e pertence a $\mathfrak{R}$ podemos dizer que

$$A = A_1 \cup A_2, \quad \text{onde} \quad A_2 = A \setminus A_1 \in \mathfrak{R}.$$

Mas, nem todo semianel é anel de conjuntos.

Em vários problemas, na teoria de medida, em particular, é preciso considerar a união e a intersecção de uma quantidade enumerável, nem sempre de conjuntos finitos. Por isso introduzir-se-ão os seguintes conceitos:

DEFINIÇÃO: 2.1.16 σ -anel

Um anel de conjuntos chama-se σ -anel se junto com toda sucessão enumerável de conjuntos $A_1, A_2, \dots, A_n, \dots$ contém a união e a intersecção

$$S = \bigcup_{n=1}^{\infty} A_n,$$

e

$$D = \bigcap_{n=1}^{\infty} A_n.$$

Desde que cada anel de conjuntos contenha E , então $\forall A_1, A_2, \dots, A_n, \dots$, tem-se

$$\bigcup_{n=1}^{\infty} A_n = E \setminus \bigcap_{n=1}^{\infty} (E \setminus A_n), \quad E \setminus A_n \in S.$$

$$\bigcap_{n=1}^{\infty} A_n = E \setminus \bigcup_{n=1}^{\infty} (E \setminus A_n), \quad E \setminus A_n \in S.$$

Isto deduz uma relação de dualidade.

DEFINIÇÃO: 2.1.17 σ -álgebra

É natural chamar σ -álgebra a todo σ -anel com unidade. Assim uma σ -álgebra de conjuntos é um anel [álgebra] de conjuntos $\mathfrak{R}$ tal que se

$$\{A_n : n \in N\} \subset \mathfrak{R}, \text{ então } \bigcup_{n=1}^{\infty} A_n \in \mathfrak{R}.$$

Denomina-se as σ -álgebras de Álgebra de Borel ou simplesmente B-álgebra.

Medida de Borel-Lebesgue

A classe de conjuntos elementares não esgota todos os conjuntos que se consideram na Geometria elementar e nem na Análise clássica. Borel-Lebesgue, estenderam o conceito de medida, conservando as suas propriedades principais, a uma classe mais ampla que é composta por uniões enumeráveis de conjuntos e está definida em todos os conjuntos da menor σ -álgebra que contém os intervalos finitos.

Constrói-se a medida de Borel-Lebesgue nos passos da extensão dos intervalos finitos. Em cada passo estendem-se classes de conjuntos para se obter classes mais amplas até obter-se a mínima σ -álgebra. Simultaneamente, em cada passo estende-se a medida para todos os conjuntos da classe nova, considerando a medida inicial.

1. Sejam os intervalos P :

$$a < x < b = (a, b)$$

$$a \leq x < b = [a, b)$$

$$a < x \leq b = (a, b]$$

$$a \leq x \leq b = [a, b]$$



Designando-se a todo intervalo P um número $m(P)$, a medida deste intervalo, de maneira que:

$$m((a, b)) = m([a, b)) = m((a, b]) = m([a, b]) = b - a.$$

Com as seguintes condições :

(a) $m(P)$ são valores reais positivos;

(b) a medida $m(P)$ é aditiva, isto é, se $P = \bigcup_{j=1}^n P_j$ e $P_i \cap P_j = \emptyset$ para $i \neq j$, então,

$$m(P) = \sum_{j=1}^n m(P_j).$$

2. Estendendo-se a medida $m(P)$, definida para intervalos, a uma classe mais geral de conjuntos, conservando as condições (a) e (b).

O primeiro passo nesta direção consiste em estender o conceito de medida para os conjuntos elementares, que são assim chamados, pois podem ser representados na forma de união e intersecção, disjuntos dois a dois, ou seja,

$$A = \bigcup [a_i, b_i] \text{ e } [a_i, b_j] \cap [a_j, b_j] = \emptyset, \forall i \neq j, i = 1, \dots, m \text{ e } j = 1, \dots, n.$$

Definindo-se agora a medida $\mu(A)$ de conjuntos elementares do seguinte modo :

se

$$A = \bigcup_i P_i,$$

onde P_i são intervalos disjuntos dois a dois, então

$$\mu(A) = \sum_i^n (P_i),$$

Conclui-se, que $\mu(A)$ é um semianel de conjuntos.

Apresentando a teoria de medida de Borel-Lebesgue é necessário considerar não só as uniões finitas como também as infinitas. Para evitar que apareçam, neste caso, conjuntos de “medida infinita”, limita-se o conjunto.

Considerando o conjunto $E = \{0 \leq x \leq 1\}$. Nesta classe de conjuntos definem-se as funções $\mu^*(A)$ e $\mu_*(A)$.



DEFINIÇÃO: 2.1.18 Medida superior

Chama-se medida superior $\mu^*(A)$ do conjunto A ao

$$\inf_{ACUP_i} \sum \mu(P_i),$$

onde a cota inferior se refere a todos os recobrimentos do conjunto A por meio de sistemas finitos ou enumeráveis de intervalos.

DEFINIÇÃO: 2.1.19 Medida inferior

Chama-se medida inferior $\mu_*(A)$ do conjunto A o número

$$1 - \mu^*(E \setminus A).$$

DEFINIÇÃO: 2.1.20 Conjunto mensurável

Um conjunto A chama-se mensurável quando

$$\mu_*(A) = \mu^*(A).$$

O valor comum $\mu(A)$ das medidas superior e inferior de um conjunto mensurável A é a medida de Borel-Lebesgue.

É possível mostrar que quando

$$A \subset \bigcup_n A_n,$$

então

$$\mu^*(A) \leq \sum_n \mu^*(A_n).$$

Função Mensurável

Sejam X e Y dois conjuntos arbitrários escolhidos em dois sistemas de subconjuntos ϑ e ϑ' , respectivamente.

Uma função abstrata $y = f(x)$ com domínio em X e imagem em Y chama-se (ϑ, ϑ') -mensurável quando $A \in \vartheta'$ deduz-se que $f^{-1}(A) \in \vartheta$.

Por exemplo, se X e Y estão na reta numérica R^1 (diz-se que são funções reais de variável real) e se ϑ e ϑ' são o sistema de todos os subconjuntos abertos (ou fechados) de R^1 , a definição dada de função mensurável coincide com a de continuidade. Tomando-se para ϑ e ϑ' o sistema

de todos os conjuntos borelianos, obtém-se assim as chamadas funções B -mensuráveis (ou mensuráveis segundo Borel).

O conceito de função mensurável é fundamental para a teoria de integração. Corresponde ao principal conceito de μ -mensurabilidade das funções reais definidas no conjunto X , coincidindo ϑ com o sistema de todos os subconjuntos μ -mensuráveis do conjunto X e ϑ' com a coleção de todos os B -conjuntos da reta. Supondo que X é a unidade do domínio ϑ_μ de medida μ . Como toda medida σ -aditiva pode ser prolongada, na álgebra boreliana, é natural admitir desde o princípio que ϑ_μ é uma B -álgebra.

Portanto, para as funções reais dar-se-á a seguinte definição de mensurabilidade.

DEFINIÇÃO: 2.1.21 μ -mensurável

Uma função real $f(x)$ definida num conjunto X chama-se μ -mensurável quando

$$f^{-1}(A) \in \vartheta_\mu \quad (2.2)$$

qualquer que seja o conjunto boreliano A na reta numérica.

TEOREMA 2.1.3 Para que uma função $f(x)$ seja μ -mensurável, é necessário e suficiente que para $\forall c$ real o conjunto $\{x : f(x) < c\}$ seja μ -mensurável (isto é dizer que pertença a ϑ_μ).

TEOREMA 2.1.4 Seja (f_n) uma sequência μ -mensurável definida em X . Então as quatro funções $\inf_n f_n$, $\sup_n f_n$, $\lim_{n \rightarrow \infty} f_n$, e $\overline{\lim}_{n \rightarrow \infty} f_n$ são todas μ -mensuráveis.

DEFINIÇÃO: 2.1.22 Funções Simples

Uma função $f(x)$ chama-se simples quando é μ -mensurável e toma no máximo um número finito de valores. O conceito de função simples depende da escolha da medida μ .

A estrutura das funções simples é caracterizada pelo seguinte teorema.

TEOREMA 2.1.5 Uma função $f(x)$ que toma no máximo um número enumerável de valores distintos

$$y_1, y_2, \dots, y_n, \dots$$

é μ -mensurável se e somente se todos os conjuntos

$$A_n = \{x : f(x) = y_n\}$$

são μ -mensuráveis.

TEOREMA 2.1.6 Para que uma função $f(x)$ seja μ -mensurável é necessário e suficiente que seja representada como limite de uma sequência uniformemente convergente de funções mensuráveis simples.

DEFINIÇÃO: 2.1.23 Convergência em quase todo ponto (qtp)

Uma sequência de funções $f_n(x)$ definida em um espaço X chama-se convergente em quase todo ponto qtp para $F(x)$ quando

$$\lim_{n \rightarrow \infty} f_n(x) = F(x), \quad \text{qtp} \quad x \in X. \quad (2.2)$$

Isto significa dizer que o conjunto daqueles pontos x em que (2.2) não se verifica é de medida nula.

TEOREMA 2.1.7 Se uma sequência de funções μ -mensuráveis $f_n(x)$ converge para $F(x)$ em qtp do espaço X , $F(x)$ é também mensurável.

DEFINIÇÃO: 2.1.24 Convergência em Medida

Diz-se que uma sequência de funções mensuráveis $f_n(x)$ converge em medida para uma função $F(x)$ quando para qualquer $\sigma > 0$

$$\lim_{n \rightarrow \infty} \mu\{x : |f_n(x) - F(x)| \geq \sigma\} = 0.$$

TEOREMA 2.1.8 Se uma sequência de funções mensuráveis $f_n(x)$ converge em qtp para uma função $F(x)$, então $f_n(x)$ converge em medida para a mesma função limite $F(x)$.

TEOREMA 2.1.9 Supondo que uma sequência de funções mensuráveis $f_n(x)$ converge em medida para $f(x)$. Então, desta sequência obter-se-á uma subsequência $\{f_{n_k}(x)\}$ que converge em qtp para $f(x)$.

DEFINIÇÃO: 2.1.25 Integral de Lebesgue

O conceito da integral de Riemann, é aplicado somente para funções contínuas ou que não tenham “demasiados” pontos de descontinuidades. Para integral de funções mensuráveis que podem ser descontínuas em quase todos pontos onde estão definidas, (ou podem estar definidas num conjunto abstrato de maneira que o conceito de continuidade necessite de sentido para elas), a construção da integral de Riemann não é válida. Para estas funções existe um conceito perfeito e flexível da integral introduzido por Lebesgue.

A integral de Lebesgue define-se de um mesmo modo para determinadas funções em qualquer espaço provido de medida ([17]).

Seja μ uma medida, f uma função mensurável não negativa definida em um conjunto mensurável E .

A integral da função f sobre o conjunto E em relação a medida μ é definida a seguir:

$$\int_E f(x) d\mu(x) = \int_E f d\mu = \sup_{E=\bigcup E_i} \sum_{i=1}^{\infty} \mu(E_i),$$

onde os E_i , $i = 1, 2, \dots$ são mensuráveis e disjuntos. Se E_i é vazio então $\inf_{E_i} f = 0$.

Se a medida é medida de Lebesgue, $\mu = m_n$ então a integral é chamada de integral de Lebesgue da função f sobre o conjunto E e denota-se por

$$\int_E f(x) dx = \int_E f(x_1, \dots, x_n) dx_1 \dots dx_n \quad (x = (x_1, \dots, x_n)).$$

DEFINIÇÃO: 2.1.26 Fórmula da quadratura trapezoidal

Tem-se uma aproximação trapezoidal dada da seguinte forma

$$I(f) := \int_a^b f(x) dx \approx \frac{b-a}{2} (f(a) + f(b))$$

DEFINIÇÃO: 2.1.27 Ponto aderente e fecho

Seja A um subconjunto de um espaço métrico M . Um ponto $p \in M$ diz-se ponto aderente ao conjunto A se, para todo $\varepsilon > 0$, vale a relação

$$B(p, \varepsilon) \cap A \neq \emptyset$$

onde, $B(p, \varepsilon)$ é uma bola aberta de centro p e raio $\varepsilon > 0$

O conjunto dos pontos aderentes ao subconjunto A chama-se fecho de A e é indicado por $\overline{A}$.

É imediato que $A \subset \overline{A}$.

DEFINIÇÃO: 2.1.28 Conjuntos densos

Dado um espaço métrico (M, d) , um subconjunto $A \subset M$ diz-se denso em M se $\overline{A} = M$.

Em outros termos, esta definição significa que, para todo $p \in M$ e todo $\varepsilon > 0$, existe $a \in A$ de maneira que $d(p, a) < \varepsilon$. Intuitivamente, para cada ponto $p \in M$ existe, arbitrariamente “próximo” de p , um ponto $a \in A$.

Ex: $\mathbb{Q}$ é denso no corpo dos números reais $\mathbb{R}$.

Obs: $[A]$: aderência, isto é, total de pontos de aderência do conjunto M .

2.2 Notações

1. Quando se trabalha com vetores e matrizes consideram-se as seguintes definições de normas e variação total:

- (a) Se $x = (x_1, \dots, x_N)'$ é um N -vetor, então

$$\|x\| = \sum_{i=1}^N |x_i|.$$

- (b) Se $A = (a_{ij})$ é uma matriz $m \times n$, então

$$\|A\| = \sum_{i=1}^M \sum_{j=1}^N |a_{ij}|.$$

- (c) Se $A(t) = (a_{ij}(t))$ é uma matriz de funções definidas no intervalo $[a, b]$, então a *variação total* de $A(t)$ é definida por

$$V_a^b = V[a, b, A(t)] = \sup \sum_{i=1}^m \|A(t_i) - A(t_{i-1})\|,$$

onde o supremo é tomado de todas as possíveis partições do intervalo $[a, b]$

2. Para algum programa linear **PL** será escrito **V[PL]** para o seu valor ótimo.
3. Para a função $f(t)$ as notações serão:

$$f(t-) = \lim_{s \rightarrow t-} f(s) \quad e \quad f(t+) = \lim_{s \rightarrow t+} f(s),$$

quando existir o limite.

Resolução do PPLC por aproximação via discretização

Neste capítulo estuda-se um método para achar soluções aproximadas para problemas de programação linear contínua ([7]). Aproxima-se o problema original por um problema discreto, discretizando o intervalo e usando a fórmula de quadratura do trapézio, por um caminho natural. A solução do problema discreto converge em quase todo ponto para a solução ótima do problema original. Mostra-se que a solução ótima é uma função de variação limitada.

3.1 Introdução

O problema de programação linear contínuo (PPLC), é formulado da seguinte forma

$$(PPLC) \text{ maximizar } \int_0^T c^T(t)x(t)dt$$

$$\text{sujeito a } B(t)x(t) \leq a(t) + \int_0^t K(t,s)x(s)ds, \quad x(t) \geq 0$$

onde $B(t)$, $K(t,s)$ são matrizes $M \times N$, $a(t) \in \mathbb{R}^M$, $a(t) \in \mathbb{R}^M$ e $x(t) \in \mathbb{R}^N$, $x(t)$, $a(t)$, e os elementos de $B(t)$, $K(t,s)$ são funções limitadas e mensuráveis segundo Lebesgue.

Este problema refere-se ao problema primal.

Consideremos o problema dual,



Capítulo 3

Resolução do PPLC por aproximação via discretização

Neste capítulo estuda-se um método para achar soluções aproximadas para problemas de programação linear contínua ([7]). Aproxima-se o problema original por um problema discreto, discretizando o intervalo e usando a fórmula de quadratura do trapézio, por um caminho natural. A solução do problema discreto converge em quase todo ponto para a solução ótima do problema original. Mostra-se que a solução ótima é uma função de variação limitada.

3.1 Introdução

O problema de programação linear contínuo (PPLC), é formulado da seguinte forma

$$\begin{aligned} \text{(PPLC)} \quad & \text{maximizar} \quad \int_0^T c^T(t)x(t)dt \\ & \text{sujeito à} \quad B(t)x(t) \leq a(t) + \int_0^t K(t,s)x(s)ds, \quad x(t) \geq 0 \end{aligned}$$

onde $B(t)$, $K(t,s)$ são matrizes $M \times N$, $c(t) \in \mathbb{R}^N$, $a(t) \in \mathbb{R}^M$ e $x(t) \in \mathbb{R}^N$. $x(t)$, $a(t)$, e os elementos de $B(t)$, $K(t,s)$ são funções limitadas e mensuráveis segundo Lebesgue.

Este problema refere-se ao problema primal.

Consideremos o problema dual,



$$\begin{aligned}
 (\text{PPLCD}) \quad & \text{minimizar} \quad \int_0^T a^T(t)w(t)dt \\
 & \text{sujeito à} \quad B^T(t)x(t) \geq c(t) + \int_t^T K^T(s,t)w(s)ds, \quad w(t) \geq 0,
 \end{aligned}$$

onde $w(t) \in \mathbb{R}^M$, é a variável dual.

Este tipo de problema é resolvido por aproximação de problemas discretos, que se chama discretização do problema original. Isto é, constrói-se uma “boa” discretização, do problema original, a qual pode-se resolver efetivamente, com a propriedade, que soluções do problema discretizado convergem para uma solução do problema original, numa topologia adequada. Antes de prosseguir apresentar-se-á, como Proposição 3.1.1, resultados de dualidade necessários ao desenvolvimento deste capítulo.

PROPOSIÇÃO 3.1.1 (Tyndall [21], Levinson [16] e Grinold [11]) *Se $\hat{x}$ e $\hat{w}$ são factíveis para o primal e dual, respectivamente, então*

$$\int_0^T c^T(t)x(t)dt \leq \int_0^T a^T(t)w(t)dt.$$

Além disso, se $\int_0^T c^T(t)x(t)dt = \int_0^T a^T(t)w(t)dt$ então $\hat{x}$ e $\hat{w}$ são ótimos para os problemas primal e dual respectivamente.

3.2 Método de aproximação por integração numérica

Este método consiste em discretizar o intervalo $[0, T]$, dividindo-o em segmentos normalmente equidistantes, formando assim uma rede de pontos que na n -ésima partição σ_n são dadas por $t_i^n = \frac{iT}{l_n}$, $i = 0, \dots, l_n$, com l_n número de segmentos da discretização σ_n de $[0, T]$ e n o número de vezes que se discretiza o intervalo.

Todas as integrais que aparecem na função objetiva e nas restrições, são usadas na aproximação trapezoidal para se obter um problema de programação linear (PPL), para cada i entre 0 e l_n , faz-se $t = t_i^n$ e substitui-se na função objetiva e na restrição do problema original.

Discretizando o intervalo $[0, T]$ na forma acima e usando a fórmula de quadratura do trapézio obtém-se a seguinte discretização do problema, que é denotado por $\mathbf{P}_n$:

$$\begin{aligned}
 (\mathbf{P}_n) \quad & \text{maximizar} \quad c_n^T x_n \\
 & \text{sujeito à} \quad G_n x_n \leq a_n, \\
 & \quad \quad \quad x_n \geq 0, \quad t \in [0, T].
 \end{aligned}$$

onde

$$a_n^T = (a_{n0}^T, a_{n1}^T, \dots, a_{nl_n}^T), \text{ com}$$

$$a_{n0} = \frac{1}{2}a(0), a_{nl_n} = \frac{1}{2}a(T), a_{ni} = a(t_i^n) \text{ para } i = 1, 2, \dots, l_n - 1,$$

$$c_n^T = (c_{n0}^T, c_{n1}^T, \dots, c_{nl_n}^T), \text{ com}$$

$$c_{n0}^T = \frac{1}{2}c(0), c_{nl_n}^T = \frac{1}{2}c(T), c_{ni} = c(t_i^n) \text{ para } i = 1, 2, \dots, l_n - 1.$$

Denotando por

$$B_0^n = \frac{1}{2}B(0), B_{l_n}^n = \frac{1}{2}B(T), B_i^n = B(t_i^n) \text{ para } 1 \leq i \leq l_n - 1,$$

G_n é dado por:

$$G_n = \begin{pmatrix} \frac{1}{2}B_0^n & 0 & 0 & 0 & 0 \\ 0 & B_1^n & 0 & 0 & 0 \\ 0 & 0 & \ddots & 0 & 0 \\ 0 & 0 & 0 & B_{l_n-1}^n & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{2}B_{l_n}^n \end{pmatrix} - \frac{T}{l_n} \begin{pmatrix} \frac{K}{4} & 0 & 0 & 0 & 0 \\ \frac{K}{4} & \frac{K}{2} & 0 & 0 & 0 \\ \frac{K}{2} & K & \frac{K}{2} & 0 & 0 \\ \frac{K}{2} & K & \frac{K}{2} & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ \frac{K}{4} & K & K & \dots & \frac{K}{4} \\ \frac{K}{4} & \frac{K}{2} & \frac{K}{2} & \dots & \frac{K}{4} \end{pmatrix}.$$

A fórmula de quadratura trapezoidal é dada por:

$$\int_a^b f(x)dx \approx \frac{b-a}{2}(f(a) + f(b)).$$

Fazendo $a = 0$, $b = T$ e aplicando a regra do trapézio em cada segmento da partição σ_n , na função objetiva, nas restrições e somando obtém-se:

1. Função objetiva:

$$\begin{aligned}
 \int_0^T c^T(t)x(t)dt &= \sum_{i=1}^{l_n} \int_{t_{i-1}^n}^{t_i^n} c^T(t)x(t)dt \\
 &\approx \sum_{i=1}^{l_n} \frac{1}{2}(t_i^n - t_{i-1}^n)(c^T(t_{i-1}^n)x(t_{i-1}^n) + c^T(t_i^n)x(t_i^n))
 \end{aligned}$$



O problema discretizado, possui $N \times (l_n + 1)$ variáveis, $M \times (l_n + 1)$ restrições e as variáveis são:

$$x_j(t_k) \geq 0, 1 \leq j \leq N, 0 \leq k \leq l_n.$$

A mesma coisa é feita para o problema dual, aproximando-o para o problema

$$\begin{aligned} (\mathbf{D}_n) \quad & \text{minimizar} \quad a_n^T w_n \\ & \text{sujeito à} \quad G_n^T w_n \leq c_n \\ & w_n \geq 0, \quad t \in [0, T]. \end{aligned}$$

onde G_n^T é a transposta de G_n .

Com essas mudanças feitas nos problemas de aproximação, pode-se obter $\mathbf{P}_n$ e $\mathbf{D}_n$ como um par de programas lineares duais.

Em particular nota-se que os valores aproximados das funções objetivas do primal e do dual são os correspondentes valores das funções objetivas em $\mathbf{P}_n$ e $\mathbf{D}_n$ multiplicados por $\frac{T}{l_n}$.

3.3 Convergencia e Factibilidade

Aqui introduz-se algumas notações e hipóteses básicas para o estudo do problema PLC em questão.

1. Notações

(a) Seja,

$$x_n(t) = (x_{n,i+1} - x_{ni})(t - t_i^n)^{l_n-1} + x_{ni} \text{ para } t_i^n \leq t \leq t_{i+1}^n, i = 0, 1, \dots, l_n-1, w_n(t) = (u$$

$$\text{para } t_i^n \leq t \leq t_{i+1}^n, i = 0, 1, \dots, l_n-1, V[0, T, x_n(t)] \leq R \text{ para todo } n.$$

$$K(B, d) = \{x : Bx \leq d, x \geq 0\},$$

$$J(B, h) = \{y : yB \geq h, y \geq 0\};$$

$$\overline{K}(B, d) = \text{envoltória convexa dos pontos extremos de } K(B, d)$$

$$\overline{J}(B, h) = \text{envoltória convexa dos pontos extremos de } J(B, h)$$

(b) $(A : B)$ = matriz obtida de A adicionando em A todas as colunas da matriz B (ambas tem o mesmo número de linhas),

(c) *abs x* : Se x é um vetor de componentes $x_1, x_2, \dots, x_r$ então *abs x* é o vetor de componentes $|x_1|, |x_2|, \dots, |x_r|$.

2. Hipóteses básicas

Hipótese I: $a(t)$, $c(t)$ e $B(t)$, são funções de variação limitada e K é uma matriz constante.

Hipótese II: Para cada $n \in \mathbb{N}$, existe soluções factíveis para P_n e D_n .

Hipótese III: Existe um $n_0 \in \mathbb{N}$ tal que, para todo $n \geq n_0$ obtém-se:

(a) $x_n \in \overline{K}(G_n, a_n)$, onde x_n é solução ótima de P_n ,

(b) $w_n \in \overline{J}(G_n^T, c_n)$, onde w_n é solução ótima de D_n ,

3. Propriedades das seqüências $\{x_n\}$ e $\{w_n\}$

LEMA 3.3.1 *Seja A uma matriz constante. Então existe um número $\rho > 0$ tal que para algum vetor d , d^* vale as implicações a seguir:*

(a) $x \in \overline{K}(A, d) \Rightarrow \|x\| \leq \rho \|d\|$,

(b) $y \in \overline{J}(A', d^*) \Rightarrow \|y\| \leq \rho \|d^*\|$.

LEMA 3.3.2 *Sejam as hipóteses I, II, III, satisfeitas. Então existe um número real positivo R tal que :*

(a) $\|x_{nk}\| \leq R$ para todo $n, k (0 \leq k \leq l_n)$ e $V[0, T, x_n(t)] \leq R$ para todo n .

(b) $\|w_{nk}\| \leq R$ para todo $n, k (0 \leq k \leq l_n)$ e $V[0, T, w_n(t)] \leq R$ para todo n .

TEOREMA 3.3.1 *Suponha que as hipóteses I, II, III estejam satisfeitas e que, para cada n , $\{x_n(t)\}$ e $\{w_n(t)\}$, são soluções ótimas dos problemas P_n e D_n , respectivamente. Então existe uma seqüência de índices $\{k\} = \{n_i\}$ tal que as subsequências $\{x_k(t)\}, \{w_k(t)\}$ convergem qtp em $[0, T]$ para funções de variação limitada, que denota-se por $\hat{x}(t)$ e $\hat{w}(t)$ respectivamente.*

Prova:

Pelo Lema 3.3.2 as seqüências de funções $\{x_n(t)\}, \{w_n(t)\}$ são uniformemente limitadas e de variação total uniformemente limitada em $[0, T]$. O resultado segue do Teorema 2.1.1. $\square$

4. Factibilidade

LEMA 3.3.3 *Sejam as hipóteses I, II, III satisfeitas e sejam $\{x_k(t)\}$ e $\{w_k(t)\}$ subsequências como no Teorema 3.3.1. Então*

$$(a) \lim_{k \rightarrow \infty} \int_0^T c^T(t) x_k(t) dt = \int_0^T c^T(t) \hat{x}(t) dt,$$

$$(b) \lim_{k \rightarrow \infty} \int_0^T a^T(t) w_k(t) dt = \int_0^T a^T(t) \hat{w}(t) dt,$$

$$(c) \lim_{k \rightarrow \infty} [(B(t) : I) x_k(t) - (K : 0) \int_0^t x_k(s) ds] = (B(t) : I) \hat{x}(t) - (K : 0) \int_0^t \hat{x}(s) ds,$$

$$(d) \lim_{k \rightarrow \infty} [(B^T(t) : I) w_k(t) - (K^T : 0) \int_0^t w_k(s) ds] = (B^T(t) : I) \hat{w}(t) - (K^T : 0) \int_0^t \hat{w}(s) ds.$$

Prova: É consequência imediata dos teoremas de Helly e da convergência dominada de Lebesgue. $\square$

TEOREMA 3.3.2 *Seja σ_k um refinamento de σ_{k-1} para cada k . Sejam as hipóteses I, II, III satisfeitas e $\{x_k(t)\}, \{w_k(t)\}$ subsequências obtidas de acordo com o Teorema 3.3.1. Suponha que*

$$\hat{x}(t) = \lim_{k \rightarrow \infty} x_k(t) \quad \text{qtp,} \quad e \quad \hat{w}(t) = \lim_{k \rightarrow \infty} w_k(t) \quad \text{qtp.}$$

Então $\hat{x}(t)$ e $\hat{w}(t)$ são soluções factíveis para os problemas primal e o dual do PPLC respectivamente.

Prova:

Para cada k , seja $\varepsilon_k(t)$ definido por

$$(B(t) : I) x_k(t) - (K : 0) \int_0^t x_k(s) ds = a(t) + \varepsilon_k(t), \quad 0 \leq t \leq T,$$

onde I é a matriz indetidade.

Provando-se que $\lim_{k \rightarrow \infty} \varepsilon_k(t) = 0$ qtp em $[0, T]$, então a factibilidade de $\hat{x}(t)$ segue da seguinte afirmação

$$\lim_{k \rightarrow \infty} [(B(t) : I)x_k(t) - (K : 0) \int_0^t x_k(s) ds] = (B(t) : I)\hat{x}(t) - (K : 0) \int_0^t \hat{x}(s) ds.$$

Como $B(t)$, $x_k(t)$ e $a(t)$ sejam contínuos qtp em $[0, T]$, $\varepsilon_k(t)$ também é contínua qtp para cada k .

O limite $\lim_{k \rightarrow \infty} \varepsilon_k(t)$ existe qtp; e será denotado por $\hat{\varepsilon}(t)$.

Por construção, $\varepsilon_k(t) = 0$ para $t \in \sigma_k$.

Se $\hat{t} \in \bigcup_{k=1}^{\infty} \sigma_k$, então existe um k_0 tal que $\hat{t} \in \sigma_k$ para $k \geq k_0$, isto é $\varepsilon_k(\hat{t}) = 0$ para $k \geq k_0$, isto é $\lim_{k \rightarrow \infty} \varepsilon_k(\hat{t}) = 0 = \hat{\varepsilon}(\hat{t})$.

Pode-se deduzir facilmente as afirmações abaixo, que são essenciais para completar a demonstração.

1. O conjunto $\sigma = \bigcup_{k=1}^{\infty} \sigma_k$ é denso em $[0, T]$.
2. A função $\hat{\varepsilon}(t)$ é contínua qtp em $[0, T]$, pois $\lim_{k \rightarrow \infty} \varepsilon_k(t) = \hat{\varepsilon}(t)$.

Seja t_0 algum ponto de continuidade de $\hat{\varepsilon}(t)$. Pela afirmação (1) t_0 é um ponto aderente ao conjunto σ , isto é, existe uma sequência $\{\tau_r\}_{r=1}^{\infty}$ de pontos de σ distintos de t_0 tal que $\lim_{r \rightarrow \infty} \tau_r = t_0$.

Da afirmação (2) podemos concluir que $0 = \lim_{r \rightarrow \infty} \hat{\varepsilon}(\tau_r) = \hat{\varepsilon}(t_0)$, o qual completa a prova de factibilidade de $\hat{x}(t)$.

A factibilidade de $\hat{w}(t)$ pode ser provada de forma análoga. □

$\{x_n\}$ é soluções factível do problema primal P_n , que é uma sequência de soluções aproximadas do correspondente problema de programação linear primal contínuo, dessa sequência conhecida existe uma subsequência de soluções que converge qtp para uma função $\hat{x}(t)$. Essa função satisfaz as restrições (segundo Lebesgue) qtp.

Uma afirmação similar se aplica ao problema Dual D_n , razoável esperar que as funções $\hat{x}(t)$ e $\hat{w}(t)$ sejam soluções ótimas para o primal e dual do PPLC, respectivamente. Isto será visto com mais detalhes no próximo teorema.



TEOREMA 3.3.3 *Sejam as hipóteses I, II, III satisfeitas e $\sigma_{n+1} \subset \sigma_n$ para cada n . Então existe uma subsequência $\{k\} = \{n_i\}$ tal que*

- a) $x_k(t) \rightarrow \hat{x}(t)$ qtp com $k \rightarrow \infty$ e $\hat{x}(t)$ é uma função de variação limitada.*
- b) $w_k(t) \rightarrow \hat{w}(t)$ qtp com $k \rightarrow \infty$ e $\hat{w}(t)$ é uma função de variação limitada.*
- c) $(T/l_k)c_k^T x_k \rightarrow \int_0^T c^T(t)\hat{x}(t)dt$ quando $k \rightarrow \infty$.*
- d) $(T/l_k)a_k^T w_k \rightarrow \int_0^T a^T(t)\hat{w}(t)dt$ quando $k \rightarrow \infty$.*
- e) $\int_0^T c^T(t)\hat{x}(t)dt = \int_0^T a^T(t)\hat{w}(t)dt$ (otimilidade)*

Prova:

1. Está provado no Teorema 3.3.1

2. Está provado no Teorema 3.3.1

Usam-se as seguintes notações adicionais para provar este item apenas.

$$\hat{x}_{k0} = \frac{1}{2}\hat{x}(0), \quad \hat{x}_{ki} = \hat{x}(t_i^k), \quad i = 1, 2, \dots, l_k-1, \quad \hat{x}_{kl_k} = \frac{1}{2}\hat{x}(T)$$

$$\hat{x}_k^T = (\hat{x}_{k0}^T, \hat{x}_{k1}^T, \dots, \hat{x}_{kl_k-1}^T, \hat{x}_{kl_k}^T)$$

$$\hat{x}_k(t) = \hat{x}_{ki} + [\hat{x}_{ki+1} - \hat{x}_{ki}] \cdot (t - t_i^k)/l_k \quad \text{para} \quad t_i^k \leq t \leq t_{i+1}^k, i = 0, 1, \dots, l_k - 1.$$

Observe que $\lim_{k \rightarrow \infty} \hat{x}_k(t) = \hat{x}(t)$ qtp,

E = vetor cujos elementos são iguais a um. (o número de elementos do vetor E ficará claro no contexto).

Usando a formula de quadratura do trapézio tem-se:

$$\begin{aligned} \int_0^T c_k^T(t)x_k(t)dt &\approx \frac{T}{2l_k} [c^T(t_0^k)x(t_0^k) + 2 \sum_{i=1}^{k-1} c^T(t_i^k)x(t_i^k) + c^T(t_k^k)x(t_k^k)] \\ &= \frac{T}{l_k} \sum_{i=0}^k c_{ki}^T x_{ki} = \frac{T}{l_k} c_k^T \hat{x}_k. \end{aligned}$$

Aplicando o limite quando $k \rightarrow \infty$, pelo Teorema 3.3.1, obtem-se

$$\lim_{k \rightarrow \infty} \left(\frac{T}{l_k} \right) c_k^T x_k = \int_0^T c^T(t)\hat{x}(t)dt.$$

Agora mostrar-se-á que

$$\lim_{k \rightarrow \infty} \left[\left(\frac{T}{l_k} \right) a_k^T (\hat{x}_k - x_k) \right] = 0.$$

De fato

$$\left| \frac{T}{l_k} c_k^T (\hat{x}_k - x_k) \right| \leq \frac{T}{l_k} (abs \ c_k^T) (abs \ (\hat{x}_k - x_k)).$$

Como $c(t)$ é limitada, existe uma constante $\lambda > 0$ tal que $abs a_k \leq \lambda E$ para algum k .

Substituindo, obtem-se

$$\begin{aligned} \left| \frac{T}{l_k} c_k^T (\hat{x}_k - x_k) \right| &\leq \frac{\lambda T}{l_k} E \ abs \ (\hat{x}_k - x_k) \\ &= \lambda \int_0^T E \ abs \ (\hat{x}_k - x_k) dt \rightarrow 0 \text{ para } k \rightarrow \infty. \end{aligned}$$

A afirmação (c) do teorema é consequência imediata do Teorema de convergência dominada de *Lebesgue*.

3. análogo a c)

4. A igualdade $\int_0^T c^T(t) \hat{x}(t) dt = \int_0^T a^T(t) \hat{w}(t) dt$ e otimalidade de $\hat{x}$ e $\hat{w}$ seguem da Proposição 3.1.1.

Assim, permitindo que as soluções factíveis do PLC sejam funções de variação limitada, de acordo com os resultados apresentados neste capítulo, é possível aproximar o PLC por problemas discretos cujas soluções convergem qtp para a solução ótima do problema original. Um exemplo simples é resolvido numericamente no quinto capítulo a título de ilustração.

Capítulo 4

Programação linear contínua separável

Neste capítulo discute-se uma classe especial de problemas de programação linear contínua, chamada de programa linear contínua separável (PLCS) ([19]). Um programa linear dual e uma aproximação discreta correspondente são introduzidos e em seguida é feita uma discussão de suas propriedades. A aproximação discreta dá origem à solução melhorada, construída a partir de alguma solução factível para PLCS e disso resulta a dualidade forte. Existe uma variedade de implementações possíveis de um algoritmo para resolver problemas PLCS usando a solução melhorada.

4.1 Introdução

Considere o problema de programação linear contínua,

$$\begin{aligned} \text{(PPLC)} : \quad & \text{minimizar} \quad \int_0^T c^T(t)x(t)dt \\ & \text{sujeito à} \quad B(t)x(t) \leq a(t) + \int_0^t K(t,s)x(s)ds, \\ & \quad \quad \quad x(t) \geq 0, \quad t \in [0, T], \end{aligned}$$

com $x(t)$, $c(t)$, e os elementos de $B(t)$ e $K(t, s)$ funções limitadas e mensuráveis.

Neste capítulo considera-se uma subclasse do PPLC, Programa Linear Contínua Se-



parável (PLCS), que se defini da seguinte forma:

$$\begin{aligned} \text{(PLCS)} : \quad & \text{minimizar} \quad \int_0^T c^T(t)x(t)dt \\ & \text{sujeito à} \quad \int_0^t Gx(s)ds + y(t) = a(t), \end{aligned} \quad (4.1)$$

$$Hx(t) + z(t) = b(t), \quad (4.2)$$

$$x(t), y(t), z(t) \geq 0, \quad t \in [0, T], \quad (4.3)$$

onde $x(t)$, $z(t)$, $b(t)$ e $c(t)$ são funções limitadas e mensuráveis e $y(t)$ e $a(t)$ são funções absolutamente contínuas. As dimensões de $x(t)$, $y(t)$ e $z(t)$ são n_1 , n_2 e n_3 , respectivamente, G é uma matriz $n_2 \times n_1$ e H é uma matriz $n_3 \times n_1$.

Seja $n = n_1 + n_2 + n_3$ e $\omega^T(t) = (x(t), y(t), z(t))$, onde $\omega(t)$ é o conjunto das variáveis do PLCS.

4.2 O Problema Dual

O problema linear dual (PLCS*), para PLCS, é dado por:

$$\begin{aligned} \text{(PLCS*)} : \quad & \text{maximizar} \quad - \int_0^T d\pi^T(t)a(t) - \int_0^T \eta^T(t)b(t)dt \\ & \text{sujeito à} \quad c(t) - G^T\pi(t) + H^T\eta(t) \geq 0, \\ & \eta(t) \geq 0, \quad \text{qtp em } [0, T], \end{aligned} \quad (4.4)$$

onde os componentes de $\eta(t)$ pertencem a $L_1[0, T]$, e cada componente de $\pi(t)$ é monotônico crescente e contínuo à direita em $[0, T]$ com $\pi(T) = 0$.

Será conveniente representar as variáveis $(\pi(t), \eta(t))$ por $\theta(t)$ e a primeira restrição por $\psi(t)$, ou seja

$$\theta^T(t) = (\pi^T(t), \eta^T(t))$$

e

$$\psi(t) = c(t) - G^T\pi(t) + H^T\eta(t)$$

LEMA 4.2.1 Dualidade Fraca

$$V[PLCS^*] \leq V[PLCS].$$

Prova: Suponha que $\omega(t)$ seja factível para PLCS e que $\theta(t)$ seja factível para $PLCS^*$. Como $a(t) = \int_0^t Gx(s)ds + y(t)$ e $b(t) = Hx(t) + z(t)$, então

$$\begin{aligned} V[PLCS^*] &\leq - \int_0^T d\pi^T(t)a(t) - \int_0^T \eta^T(t)b(t)dt \\ &= - \int_0^T d\pi^T(t)\left(\int_0^t Gx(s)ds + y(t)\right) - \int_0^T \eta^T(t)(Hx(t) + z(t))dt. \end{aligned}$$

Mas, $\int_0^t Gx(s)ds = Gx(t) - Gx(0) = Gx(t)$, pois $x(0) = 0$, logo,

$$\begin{aligned} V[PLCS^*] &\leq - \int_0^T d\pi^T(t)Gx(t) - \int_0^T d\pi^T(t)y(t) - \int_0^T \eta^T(t)Hx(t)dt - \int_0^T \eta^T(t)z(t)dt \\ &\leq - \left(- \int_0^T G^T x(t)\pi(t)dt\right) - \int_0^T \eta^T(t)Hx(t)dt - \int_0^T d\pi^T(t)y(t) - \int_0^T \eta^T(t)z(t)dt \\ &= \int_0^T (G^T \pi(t) - H^T \eta(t))^T x(t)dt - \int_0^T d\pi^T(t)y(t) - \int_0^T \eta^T(t)z(t)dt. \end{aligned}$$

Sob as condições de não negatividade das funções $x(t)$, $y(t)$ e $z(t)$ e monotonicidade de $\pi(t)$, usando (4.4), podemos concluir que

$$\int_0^T (G^T \pi(t) - H^T \eta(t))^T x(t)dt - \int_0^T d\pi^T(t)y(t) - \int_0^T \eta^T(t)z(t)dt \leq \int_0^T c^T x(t)dt.$$

Portanto

$$V[PLCS^*] \leq V[PLCS].$$

□

Outro programa linear dual, para PLCS, formulado por Anderson e Nash é definido da seguinte forma

$$\begin{aligned} (PLCS^{**}): \quad &\text{maximizar} \quad - \int_0^T a^T(t)u(t)dt - \int_0^T b^T(t)v(t)dt \\ &\text{sujeito à} \quad c(t) + \int_t^T G^T u(s)ds + H^T v(t) \geq 0, \\ &\quad u(t), v(t) \geq 0, \quad t \in [0, T], \end{aligned}$$

onde $u(t)$ e $v(t)$ são funções mensuráveis e limitadas.

O problema dual $PLCS^*$ é uma generalização do $PLCS^{**T}$.

Note que se $\theta(t)$ é factível para $PLCS^*$ e $\pi(t)$ é absolutamente contínua, então

$$u(t) = \dot{\pi}(t) \quad e \quad v(t) = \eta(t),$$



são factíveis para $PLCS^{**}$ e os valores da função objetivo são os mesmos nos seus respectivos problemas.

Reciprocamente, se $u(t)$ e $v(t)$ é factível para $PLCS^{**}$ então

$$\pi(t) = - \int_t^T u(s) ds \quad e \quad \eta(t) = v(t),$$

é factível para $PLCS^*$ e novamente as duas soluções tem o mesmo valor na função objetivo em seus respectivos programas lineares.

4.3 Discretização

Nesta seção são introduzidas duas discretizações para PLCS.

Para que o procedimento tenha sucesso fazem-se necessárias algumas hipóteses sobre a natureza das funções $a(t)$, $b(t)$ e $c(t)$.

Hipótese 1: As funções $a(t)$ e $c(t)$ são lineares por partes e a função $b(t)$ é constante por partes em $[0, T]$.

Hipótese 2: A região factível para PLCS é limitada.

Fixe uma partição arbitrária $P = \{t_0, \dots, t_m\}$ de $[0, T]$ de modo que $a(t)$ e $c(t)$ são lineares por partes e $b(t)$ é constante por partes com partição P , isto é, $a(t)$ e $c(t)$ são lineares e $b(t)$ é constante em cada subintervalo $[t_{i-1}, t_i]$, $i = 1, \dots, m$. Supõe-se também, que $a(t)$ é contínua em $[0, T]$.

Dado P , seja

$$Q = \{t_0, \frac{t_0 + t_1}{2}, t_1, \dots, t_{m-1}, \frac{t_{m-1} + t_m}{2}, t_m\}.$$

Isto é, cada intervalo $[t_i, t_{i+1}]$ da partição P é dividido ao meio. Sejam

$$u_i = \frac{t_{i-1} + t_i}{2}, \quad i = 1, \dots, m,$$

$$\tau_i = \frac{t_i - t_{i-1}}{2}, \quad i = 1, \dots, m.$$

Então,

$$Q = \{t_0, u_1, t_1, u_2, \dots, t_{m-1}, u_m, t_m\}$$



e τ_i , $i = 1, \dots, m$ são os comprimentos dos intervalos da partição P . Com essas notações e hipóteses pode-se definir as duas discretizações.

1. Discretização Padrão

Formula-se a discretização padrão $PD(P)$, em relação a partição P e mostra-se algumas de suas propriedades.

Seja

$PD(P)$:

minimizar

$$\sum_{i=1}^m (t_i - t_{i-1}) \hat{x}^T(t_{i-1}+) c(u_i)$$

sujeito à

$$(t_1 - t_0) G \hat{x}(t_0+) + \hat{y}(t_1) = a(t_1) \quad (4.5)$$

$$(t_i - t_{i-1}) G \hat{x}(t_{i-1}+) + \hat{y}(t_i) - \hat{y}(t_{i-1}) = a(t_i) - a(t_{i-1}), i = 2, \dots, m \quad (4.6)$$

$$H \hat{x}(t_{i-1}+) + \hat{z}(t_{i-1}+) = b(t_{i-1}+), i = 1, \dots, m \quad (4.7)$$

$$\hat{x}(t_{i-1}+), \hat{y}(t_{i-1}+), \hat{z}(t_{i-1}+) \geq 0, i = 1, \dots, m \quad (4.8)$$

O problema $PD(P)$ ocorre naturalmente como problema discreto correspondente ao $PLCS$ para as funções $x(t)$, $y(t)$ e $z(t)$, tais que $x(t)$ e $z(t)$ são constantes por partes e $y(t)$ linear por partes.

Escrevendo o problema discreto de forma abreviada, tem-se

$$PD(P) : \quad \text{minimizar} \quad \hat{c}_D^T \hat{\omega}_D$$

$$\text{sujeito à} \quad A_D \hat{\omega}_D = \hat{b}_D,$$

$$\hat{\omega}_D \geq 0,$$

para alguma matriz apropriada A_D e vetores apropriados $\hat{c}_D$ e $\hat{b}_D$. Chamam-se as variáveis, por conveniência, de $\hat{x}_{t_i}$ e $\hat{y}_{t_i}$.

Para facilitar a redação vamos incorporar a restrição (4.5) em (4.6) entendendo que quando $i = 1$ tem a restrição (4.5).

LEMA 4.3.1 As restrições

$$\int_0^{t_i} Gx(s)ds + y(t_i) = a(t_i), i = 1, \dots, m,$$

do problema PLCS para $t = t_1, t = t_2, \dots, t = t_m$ são equivalentes às restrições

$$(t_1 - t_0)G\hat{x}(t_0+) + \hat{y}(t_1) = a(t_1)$$

$$\sum_{i=1}^{m-1} (a(t_{i+1}) - a(t_i) - y(t_{i+1}) + y(t_i)) + e = (t_m - t_{m-1})G\hat{x}(t_{m-1}+) + \hat{y}(t_m) = a(t_m) \Rightarrow$$

$$(t_i - t_{i-1})G\hat{x}(t_{i-1}+) + \hat{y}(t_i) - \hat{y}(t_{i-1}) = a(t_i) - a(t_{i-1}), \text{ para } i = 2, \dots, m,$$

do problema PD(P).

Prova: Usaremos indução sobre m .

O resultado é óbvio para as primeiras duas restrições, quando $t = t_1, t = t_2$ que são equivalentes à primeira restrição (4.1). Isto é,

$$\int_0^{t_1} Gx(s)ds + y(t_1) = a(t_1) \Leftrightarrow (t_1 - t_0)Gx(t_0+) + \hat{y}(t_1) = a(t_1),$$

$$e$$

$$\int_0^{t_2} Gx(s)ds + y(t_2) = a(t_2) \Leftrightarrow$$

$$\int_0^{t_1} Gx(s)ds + \int_{t_1}^{t_2} Gx(s)ds + y(t_2) = a(t_2) \Leftrightarrow$$

$$a(t_1) - y(t_1) + (t_2 - t_1)Gx(t_1+) + y(t_2) = a(t_2) \Leftrightarrow$$

$$(t_2 - t_1)Gx(t_1+) + y(t_2) - y(t_1) = a(t_2) - a(t_1).$$

Suponhamos que as restrições (4.6) sejam válidas para $t = t_{i-1}, i = 2, \dots, m$, isto é,

$$(t_{i-1} - t_{i-2})Gx(t_{i-1}+) + y(t_{i-1}) - y(t_{i-2}) = a(t_{i-1}) - a(t_{i-2}), i = 2, \dots, m.$$

Mostremos que as restrições sejam válidas para $t = t_m$. De (4.1) sabemos que

$$\int_0^{t_m} Gx(s)ds + y(t_m) = a(t_m)$$

o que é equivalente a

$$\sum_{i=0}^{m-2} \int_{t_i}^{t_{i+1}} Gx(s)ds + \int_{t_{m-1}}^{t_m} Gx(s)ds + y(t_m) = a(t_m) \Leftrightarrow$$

$$\sum_{i=0}^{m-2} (t_{i+1} - t_i)Gx(t_i+) + (t_m - t_{m-1})Gx(t_i+) + y(t_m) = a(t_m).$$

Pela hipótese de indução , conclui-se que

$$\begin{aligned} \sum_{i=0}^{m-2} (a(t_{i+1}) - a(t_i) - y(t_{i+1}) + y(t_i)) + (t_m - t_{m-1})Gx(t_m+) + y(t_m) &= a(t_m) \Leftrightarrow \\ a(t_{m-1}) - y(t_{m-1}) + (t_m - t_{m-1})Gx(t_{m-1}+) + y(t_m) &= a(t_m) \Leftrightarrow \\ (t_m - t_{m-1})Gx(t_{m-1}+) + y(t_m) - y(t_{m-1}) &= a(t_m) - a(t_{m-1}). \end{aligned}$$

□

LEMA 4.3.2 *Supõe-se que $\hat{\omega}_D$ é factível para $PD(P)$.*

Se $x(t)$ e $z(t)$ são extensões constantes por partes de $\hat{x}$ e $\hat{z}$, respectivamente, e $y(t)$ é extensão linear por partes de $\hat{y}$, então $\omega^T(t) = (x^T(t), y^T(t), z^T(t))$ é factível para $PLCS$ e

$$\int_0^T c^T(t)x(t)dt = \hat{c}_D^T \hat{\omega}_D. \quad (4.9)$$

Por outro lado se $\omega(t)$ é factível para $PLCS$ com $x(t)$ constante por partes na partição P , então $x(t_{i-1}+), y(t_i), z(t_{i-1}+), i = 1, \dots, m$, forma uma solução factível $\hat{\omega}_D$ para $PD(P)$ e a igualdade (4.9) é válida.

Prova:

Por hipótese tem-se que $x(t)$ e $z(t)$ são extensões constantes por partes de $\hat{x}$ e $\hat{z}$ respectivamente. Assim, por definição tem-se

$$x(t) = \begin{cases} \hat{x}(t_{m-1}+) & , \quad t = T \\ \hat{x}(t_{i-1}+) & , \quad t \in [t_{i-1}, t_i) \quad \text{para } i = 1, \dots, m, \end{cases}$$

$$z(t) = \begin{cases} \hat{z}(t_{m-1}+) & , \quad t = T \\ \hat{z}(t_{i-1}+) & , \quad t \in [t_{i-1}, t_i) \quad \text{para } i = 1, \dots, m \end{cases}$$

e que $y(t)$ é extensão linear por partes de $\hat{y}$, e

$$y(t) = \begin{cases} \hat{y}(t_i+), & \text{se } t = t_0, \dots, t_{m-1}, \\ 0, & t = T, \\ (\frac{t_i - t}{t_i - t_{i-1}})\hat{y}(t_{i-1}+) + (\frac{t - t_{i-1}}{t_i - t_{i-1}})\hat{y}(t_i-), & \text{para } t \in (t_{i-1}, t_i), i = 1, \dots, m. \end{cases}$$

Então

$$\begin{aligned}
 \int_0^T c^T(t)x(t)dt &= \sum_{i=1}^m \int_{t_{i-1}}^{t_i} c^T(t)x(t)dt \\
 &= \sum_{i=1}^m (t_i - t_{i-1}) c^T\left(\frac{t_{i-1} + t_i}{2}\right) \hat{x}(t_{i-1}+) \\
 &= \sum_{i=0}^m (t_i - t_{i-1}) \hat{x}^T(t_{i-1}+) c(u_i) \\
 &= \hat{c}_D^T \hat{\omega}_D.
 \end{aligned}$$

A prova da recíproca é similar. □

2. Nova Discretização

A nova discretização, **MD(P)**, em relação à partição P é dado por:

$$\begin{aligned}
 MD(P) : \quad & \text{minimizar} \quad \sum_{i=1}^m \tau_i (c^T(t_{i-1}+) \hat{x}(t_{i-1}+) + c^T(t_i-) \hat{x}(t_i-)) \\
 & \text{sujeito à} \quad \tau_1 G \hat{x}(t_0+) + \hat{y}(u_1) = a(u_1), \\
 & \quad \tau_i G \hat{x}(t_i-) + \hat{y}(t_i) - \hat{y}(u_i) = a(t_i) - a(u_i), \quad i = 1, \dots, m \\
 & \quad \tau_i G \hat{x}(t_{i-1}+) + \hat{y}(u_i) - \hat{y}(t_{i-1}) = a(u_i) - a(t_{i-1}), \quad i = 2, \dots, m \\
 & \quad H \hat{x}(t_{i-1}+) + \hat{z}(t_{i-1}+) = b(t_{i-1}+), \quad i = 1, \dots, m \\
 & \quad H \hat{x}(t_i-) + \hat{z}(t_i-) = b(t_i-), \quad i = 1, \dots, m \\
 & \quad \hat{x}(t_{i-1}+), \hat{x}(t_i-), \hat{y}(t_i), \hat{y}(u_i), \hat{z}(t_{i-1}+), \hat{z}(t_i-) \geq 0, \quad i = 1, \dots, m,
 \end{aligned}$$

Na forma abreviada

$$\begin{aligned}
 MD(P) : \quad & \text{minimizar} \quad \hat{c}^T \hat{\omega} \\
 & \text{sujeito à} \quad \hat{A} \hat{\omega} = \hat{b}, \\
 & \quad \hat{\omega} \geq 0,
 \end{aligned}$$

para alguma matriz $\hat{A}$ apropriada e vetores $\hat{c}$ e $\hat{b}$ apropriados.

Define-se por conveniência $\hat{y}(0) = a(0)$.

Note que as restrições de MD(P) são exatamente as mesmas de PD(Q) na partição Q ,

logo os vetores $\hat{b}$ e $\hat{c}$ possuem o dobro da dimensão de $\hat{b}_D$ e $\hat{c}_D$, de fato:

Seja a partição

$$Q = \{t_0, \frac{t_0 + t_1}{2}, t_1, \dots, t_{m-1}, \frac{t_{m-1} + t_m}{2}, t_m\}.$$

Fazendo

$$(\frac{t_0 + t_1}{2} - t_0) = (\frac{t_1 - t_0}{2}) = \tau_1$$

e

$$\frac{t_0 + t_1}{2} = u_1$$

e substituindo em (4.5) obtém-se,

$$\tau_1 G\hat{x}(t_0+) + \hat{y}(u_1) = a(u_1).$$

Substituindo similarmente,

$$(\frac{t_{i-1} + t_i}{2} - t_{i-1}) = (\frac{t_i - t_{i-1}}{2}) = \tau_i$$

e

$$\frac{t_{i-1} + t_i}{2} = u_i, \quad i = 1, \dots, m.$$

em (4.6) obtém-se, $PD(Q) = MD(P)$.

A diferença entre MD e PD está na função custo, que em vez de se tomar o custo médio em cada intervalo na partição, pegam-se os custos nas extremidades.

LEMA 4.3.3 *Supõe-se $\hat{\omega}_D$ factível para $PD(P)$, aumentando $\hat{\omega}_D$, usando*

$$\begin{aligned} \hat{x}(t_{i-}) &= \hat{x}(t_{i-1}+), \\ \hat{y}(\frac{t_i + t_{i-1}}{2}) &= \frac{\hat{y}(t_i) + \hat{y}(t_{i-1})}{2}, \\ \hat{z}(t_{i-}) &= \hat{z}(t_{i-1}+), \quad i = 1, \dots, m, \end{aligned}$$

obtém-se uma solução factível para $MD(P)$ tal que

$$\hat{c}^T \hat{\omega} = \hat{c}_D^T \hat{\omega}_D. \quad (4.10)$$

Prova:

Por hipótese $\hat{\omega}_D$ é factível para PD(P). Assim,

$$\begin{aligned}
 \hat{c}^T \hat{\omega} &= \sum_{i=1}^m \left(\frac{t_i - t_{i-1}}{2} \right) (c^T(t_{i-1}+) \hat{x}(t_{i-1}+) + (c^T(t_i-) \hat{x}(t_i-))) \\
 &= \sum_{i=1}^m \left(\frac{t_i - t_{i-1}}{2} \right) (c^T(t_{i-1}+) \hat{x}(t_{i-1}+) + c^T(t_i-) \hat{x}(t_{i-1}+)) \\
 &= \sum_{i=1}^m (t_i - t_{i-1}) \left(\frac{c^T(t_{i-1}+) + c^T(t_i-)}{2} \right) \hat{x}(t_{i-1}+) \\
 &= \sum_{i=1}^m (t_i - t_{i-1}) c^T \left(\frac{t_{i-1} + t_i}{2} \right) \hat{x}^T(t_{i-1}+) \\
 &= \hat{c}_D^T \hat{\omega}_D.
 \end{aligned}$$

Usou-se o fato que $\hat{x}(t_i-) = \hat{x}(t_{i-1}+)$. □

No problema MD(P) a matriz das restrições $\hat{A}$, depende de uma partição escolhida. Pode-se reescrever MD(P) de forma equivalente onde a nova matriz das restrições não dependa da partição. Fazendo as substituições abaixo,

$$\begin{cases}
 \tau_i \hat{x}(t_{i-1}+) = \tilde{x}(t_{i-1}+), \\
 \tau_i \hat{x}(t_i-) = \tilde{x}(t_i-), \\
 \hat{y}(t_i) = \tilde{y}(t_i), \\
 \hat{y}(u_i) = \tilde{y}(u_i), \\
 \tau_i \hat{z}(t_{i-1}+) = \tilde{z}(t_{i-1}+), \\
 \tau_i \hat{z}(t_i-) = \tilde{z}(t_i-), \quad i = 1, \dots, m.
 \end{cases}$$

em MD(P) pode-se reescrever MD(P) de forma equivalente, onde a nova matriz das restrições não dependa da partição, isto é,

$$\begin{aligned}
 MD(P) : \quad & \text{minimizar} \quad \sum_{i=1}^m (c^T(t_{i-1}+) \tilde{x}(t_{i-1}+) + c^T(t_i-) \tilde{x}(t_i-)) \\
 & \text{sujeito à} \quad G\tilde{x}(t_0+) + \tilde{y}(u_1) = a(u_1), \\
 & \quad G\tilde{x}(t_i-) + \tilde{y}(t_i) - \tilde{y}(u_i) = a(t_i) - a(u_i), \quad i = 1, \dots, m \\
 & \quad G\tilde{x}(t_{i-1}+) + \tilde{y}(u_i) - \tilde{y}(t_{i-1}) = a(u_i) - a(t_{i-1}), \quad i = 2, \dots, m \\
 & \quad H\tilde{x}(t_{i-1}+) + \tilde{z}(t_{i-1}+) = \tau_i b(t_{i-1}+), \quad i = 1, \dots, m \\
 & \quad H\tilde{x}(t_i-) + \tilde{z}(t_i-) = \tau_i b(t_i-), \quad i = 1, \dots, m \\
 & \quad \tilde{x}(t_{i-1}+), \tilde{x}(t_i-), \tilde{y}(t_i), \tilde{y}(u_i), \tilde{z}(t_{i-1}+), \tilde{z}(t_i-) \geq 0, \quad i = 1, \dots, m,
 \end{aligned}$$

ou na forma matricial,

$$MD(P) : \quad \text{minimizar} \quad \hat{c}^T \tilde{\omega}$$

$$\text{s. a.} \quad A\tilde{\omega} = \hat{b},$$

$$\tilde{\omega} \geq 0,$$

onde

$$\hat{c}^T = (c(t_0+)^T, 0^T, c(t_1-)^T, 0^T, c(t_1+)^T, \dots, c(t_{m-1}+)^T, 0^T, c(t_m-)^T, 0^T),$$

$$\hat{b}^T = (a(u_1)^T, \tau_1 b(t_0+)^T, (a(t_1) - a(u_1))^T, \tau_1 b(t_1-)^T, \dots, (a(t_m) - a(u_m))^T, \tau_m b(t_m-)^T),$$

$$\tilde{\omega}^T = (\tilde{x}(t_0)^T, \tilde{y}(u_1+)^T, \tilde{z}(t_0+)^T, \tilde{x}(t_1-)^T, \tilde{y}(t_1)^T, \tilde{z}(t_1-)^T, \dots, \tilde{x}(t_m-)^T, \tilde{y}(t_m)^T, \tilde{z}(t_m-)^T),$$

e

$$A = \begin{pmatrix} G & I & 0 & 0 & 0 & 0 & \dots \\ H & 0 & I & 0 & 0 & 0 & \dots \\ 0 & -I & 0 & G & I & 0 & \dots \\ 0 & 0 & 0 & H & 0 & I & \dots \\ 0 & 0 & 0 & 0 & -I & 0 & G & \dots \\ 0 & 0 & 0 & 0 & 0 & 0 & H & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \\ 0 & \dots & \dots & \dots & -I & 0 & G & I & 0 \\ 0 & \dots & \dots & \dots & 0 & 0 & H & 0 & I \end{pmatrix}.$$

A matriz A não depende da partição P escolhida e sim do tamanho de P .

Considerando agora o dual de $MD(P)$

$$MD^*(P) : \quad \text{maximizar} \quad \hat{b}^T \hat{\theta} \quad (4.12)$$

$$\text{s. a.} \quad \hat{\theta}^T A \leq \hat{c}.$$

$$\hat{\theta} \text{ irrestrita,}$$

Fazendo

$$b(t_{i-1}+) = b(t_i-)$$

e

$$a(u_i) - a(t_{i-1}) = a(t_i) - a(u_i),$$

e algumas transformações pode-se escrever $MD^*(P)$ na forma abaixo:

$MD^*(P)$: maximizar $\hat{\pi}^T(t_0+)a(t_0)$

$$+ \sum_{i=1}^m (\hat{\pi}(t_{i-1}+) + \hat{\pi}(t_i-))^T (a(t_i) - a(u_i))$$

$$- \sum_{i=1}^m (\tau_i)(\hat{\eta}(t_{i-1}+) + \hat{\eta}(t_i-))^T b(t_i-)$$

sujeito à $c(t_i-) - G^T \hat{\pi}(t_i-) + H^T \hat{\eta}(t_i-) \geq 0, \quad i = 1, \dots, m$ (4.11)

$$c(t_{i-1}+) - G^T \hat{\pi}(t_{i-1}+) + H^T \hat{\eta}(t_{i-1}+) \geq 0, \quad i = 1, \dots, m,$$

$$\hat{\eta}(t_i-), \hat{\eta}(t_{i-1}+) \geq 0, \quad i = 1, \dots, m,$$

$$\hat{\pi}(t_i-) - \hat{\pi}(t_{i-1}+) \geq 0, \quad i = 1, \dots, m,$$

$$\hat{\pi}(t_i+) - \hat{\pi}(t_i-) \geq 0, \quad i = 1, \dots, m-1,$$

$$\hat{\pi}(t_m-) \leq 0.$$

A seguir é apresentada uma série de resultados referentes a: MD , MD^* , $PLCS$ e $PLCS^*$.

O primeiro resultado mostra que se pode construir uma solução factível para $PLCS^*$ de forma natural a partir de solução factível para MD^* com o mesmos custo.

TEOREMA 4.3.1 *Suponha que $\hat{\theta}$ é factível para $MD^*(P)$. Então $\theta^T(t) = (\pi^T(t), \eta^T(t))$, onde $\pi(t)$ e $\eta(t)$ são extensões lineares por partes de $\hat{\pi}(t)$ e $\hat{\eta}(t)$, respectivamente, é factível para $PLCS^*$.*

Além disso,

$$\hat{b}^T \hat{\theta} = - \int_0^T d\pi^T(t)a(t) - \int_0^T \eta^T(t)b(t)dt \quad (4.12)$$

Por outro lado, se $\theta(t)$ é factível para $PLCS^*$ com $\hat{\pi}(t)$ e $\hat{\eta}(t)$ lineares por partes com a partição P , então $\eta(t_{i-1}+), \eta(t_i-), \pi(t_{i-1}+), \pi(t_i-), i = 1, \dots, m$, forma uma solução factível $\hat{\theta}$ para $MD^*(P)$ e a igualdade (4.12) se confirma.

Prova:

Suponha que $\hat{\theta}$ factível para $MD^*(P)$.

Analisando a inequação $\hat{\theta}^T A \leq \hat{c}$ como $\pi(t)$ e $\eta(t)$ são lineares por partes com a partição P ,

tem-se que $\psi(t), \eta(t) \geq 0$ para todo $t \in [0, T]$ e que $\pi(t)$ é crescente monotônico em $[0, T]$ com $\pi(T) = 0$, isto é, $\theta(t)$ é factível para $PLCS^*$.

Pela linearidade por partes de $a(t)$ e $\pi(t)$ e fazendo $\pi(t_m+) = \pi(T)$, tem-se

$$\begin{aligned}\int_{(t_{i-1}, t_i)} d\pi^T(t)a(t) &= \frac{1}{2}(\pi(t_i-) - \pi(t_{i-1}+))^T(a(t_i) + a(t_{i-1})), \quad i = 1, \dots, m, \\ \int_{\{t_i\}} d\pi^T(t)a(t) &= (\pi(t_i+) - \pi(t_i-))^T a(t_i), \quad i = 1, \dots, m, \\ \int_{\{t_0\}} d\pi^T(t)a(t) &= \pi^T(t_0+)a(t_0) = 0.\end{aligned}$$

Como $\eta(t)$ é linear por partes e $b(t)$ é constante por partes com a partição P , tem-se

$$\int_{t_{i-1}}^{t_i} \eta^T(t)b(t)dt = (\tau_i)(\hat{\eta}(t_{i-1}+) + \hat{\eta}(t_i-))^T b(t_{i-}), \quad i = 1, \dots, m.$$

Agora provar-se-á que a igualdade (4.12) é válida.

De fato

$$-\int_0^T d\pi^T(t)a(t) - \int_0^T \eta^T(t)b(t)dt$$

$$= -\left[\int_{\{t_0\}} d\pi^T(t)a(t) + \int_{(t_{i-1}, t_i)} d\pi^T(t)a(t) + \int_{\{t_i\}} d\pi^T(t)a(t) + \int_{\{t_m\}} d\pi^T(t)a(t)\right] \quad (4.13)$$

$$- \sum_{i=1}^m (\tau_i)(\hat{\eta}(t_{i-1}+) + \hat{\eta}(t_i-))^T b(t_{i-})$$

$$= -[\pi(t_0+)a(t_0) + \frac{1}{2} \sum_{i=1}^m (\pi(t_i-) - \pi(t_{i-1}+))^T(a(t_i) + a(t_{i-1})) + \sum_{i=1}^{m-1} (\pi(t_i+) - \pi(t_i-))a(t_i)$$

$$+ (\pi(t_m+) - \pi(t_m-))a(t_m)] - \sum_{i=1}^m (\tau_i)(\hat{\eta}(t_{i-1}+) + \hat{\eta}(t_i-))^T b(t_{i-})$$

Por hipótese temos que $\pi(t_m+) = \pi(T) = 0$, logo

$$= -\left[\frac{1}{2} \sum_{i=1}^m (\pi(t_{i-1}+) + \pi(t_i-))a(t_{i-1}) - \frac{1}{2} \sum_{i=1}^m (\pi(t_{i-1}+) + \pi(t_i-))a(t_i)\right]$$



$$\begin{aligned}
& - \sum_{i=1}^m (\tau_i) (\hat{\eta}(t_{i-1}+) + \hat{\eta}(t_i-)) b(t_i-) \\
& = - \frac{1}{2} \sum_{i=1}^m (\pi(t_{i-1}+) + \pi(t_i-)) (a(t_i) - a(t_{i-1})) - \sum_{i=1}^m (\tau_i) (\hat{\eta}(t_{i-1}+) + \hat{\eta}(t_i-)) b(t_i-) \\
& = \sum_{i=1}^m (\pi(t_{i-1}+) + \pi(t_i-)) \left(a\left(\frac{t_i + t_{i-1}}{2}\right) - a(t_{i-1}) \right) - \sum_{i=1}^m (\tau_i) (\hat{\eta}(t_{i-1}+) + \hat{\eta}(t_i-)) b(t_i-) \\
& = \sum_{i=1}^m (\pi(t_{i-1}+) + \pi(t_i-)) (a(t_i) - a(t_{i-1})) - \sum_{i=1}^m (\tau_i) (\hat{\eta}(t_{i-1}+) + \hat{\eta}(t_i-)) b(t_i-) \\
& = \hat{b}^T \hat{\theta}.
\end{aligned}$$

Na terceira igualdade usa-se o fato que $\pi(t_m+) = \pi(T) = 0$ e na última igualdade que $\pi(t_0+)a(t_0) = 0$. A prova da recíproca é similar. $\square$

O próximo resultado é análogo para MD(P) e PLCS, mas este resultado não garante que as soluções factíveis para PLCS são factíveis para MD(P) com o mesmo custo.

Nem toda solução factível para MD(P) corresponde a uma solução factível para PLCS.

TEOREMA 4.3.2 *Supor que $\omega(t)$ é factível para PLCS com $x(t)$ constante por partes com partição P .*

Então $x(t_{i-1}+)$, $x(t_i-)$, $y(t_i)$, $z(t_{i-1}+)$, $z(t_i-)$, $i = 1, \dots, m$, forma uma solução factível $\hat{\omega}$ para MD(P). Além disso,

$$\hat{c}^T \hat{\omega} = \int_0^T c^T x(t) dt. \quad (4.13)$$

Por outro lado, se $\hat{\omega}$ é factível para MD(P) e $\hat{x}(t_{i-1}+) = \hat{x}(t_i-)$ para $i = 1, \dots, m$ então $\omega(t) = (x(t), y(t), z(t))$ onde $x(t)$ e $z(t)$ são extensões contantes por partes de $\hat{x}$ e $\hat{z}$ respectivamente e $y(t)$ extensão linear por partes de $\hat{y}$. Além disso, (4.13) é válido.

Prova:

O resultado é imediato pelo Lema 4.3.2 e Lema 4.3.3. $\square$

Usando o Teorema 4.3.1, pode-se mostrar, usando o Teorema de Dualidade Forte para PL de dimensão-finita (ver Dantzig), que o valor ótimo do problema MD(P) construído de uma solução factível de PLCS com $x(t)$ constante por partes com a partição P , estabelece um limite inferior para o valor ótimo do PLCS.



TEOREMA 4.3.3 *Suponha que $\omega(t)$ é factível para PLCS com $x(t)$ constante por partes com partição P . Então*

$$V[MD(P)] \leq V[PLCS^*] \leq V[PLCS]$$

Prova:

Usando o Teorema 4.3.1 e o Teorema de Dualidade forte para PL de dimensão-finita, tem-se que $\hat{b}^T \hat{\theta} = - \int_0^T d\pi^T a(t) - \int_0^T \eta^T(t) b(t) dt$ e $V[MD(P)] = V[MD^*(P)]$.

Assim

$$V[MD(P)] = V[MD^*(P)] \leq V[PLCS^*].$$

Pelo Lema 4.2.1 tem-se que $V[PLCS^*] \leq V[PLCS]$. Logo

$$V[MD(P)] \leq V[PLCS^*] \leq V[PLCS].$$

□

COLORÁRIO 4.3.1 *Supondo-se que $\omega(t)$ é factível para PLCS e que o correspondente $\hat{\omega}$ é ótimo para MD(P), então $\omega(t)$ é ótimo para PECS.*

Prova: Sejam $\omega(t)$ factíveis para PLCS e o correspondente $\hat{\omega}$ ótimo para MD(P). Do Teorema 4.3.2 sabe-se que $\hat{c}^T \hat{\omega} = \int_0^T c^T(t) x(t) dt$. Assim

$$V[PLCS] \leq \int_0^T c^T(t) x(t) dt = \hat{c}^T \hat{\omega} = V[MD(P)].$$

Pelo Teorema 4.3.3 conclui-se que

$$V[PLCS] = V[MD(P)] = V[PLCS^*].$$

Logo $\omega(t)$ é ótimo para PLCS. □

O próximo passo é construir uma solução factível para PLCS, que de alguma forma corresponde a uma solução para MD(P). Isto pode ser feito da seguinte maneira:

Suponha que $\hat{\omega}$ é factível para MD(P) e defina

$$x_Q(t) = \begin{cases} \hat{x}(t_{i-1}+), & t \in [t_{i-1}, \frac{t_i + t_{i-1}}{2}) \text{ para } i = 1, \dots, m \\ \hat{x}(t_i-), & t \in [\frac{t_i + t_{i-1}}{2}, t_i) \text{ para } i = 1, \dots, m \\ \hat{x}(t_m-), & t = T, \end{cases} \quad (4.14)$$

$$y_Q(t) = \begin{cases} \left(\frac{t_i + t_{i-1} - 2t}{t_i - t_{i-1}}\right)\hat{y}(t_{i-1}+) + 2\left(\frac{t - t_{i-1}}{t_i - t_{i-1}}\right)\hat{y}(u_i-), & t \in (t_{i-1}, \frac{t_i + t_{i-1}}{2}) \quad \text{para } i = 1, \dots, m, \\ 2\left(\frac{t_i - t}{t_i - t_{i-1}}\right)\hat{y}(u_i+) + \left(\frac{2t - t_i - t_{i-1}}{t_i - t_{i-1}}\right)\hat{y}(t_i-), & t \in (\frac{t_i + t_{i-1}}{2}, t_i) \quad \text{para } i = 1, \dots, m, \\ \hat{y}(t_i+), & t = t_{i-1}, \quad t \in (\frac{t_i + t_{i-1}}{2}, t_i) \quad \text{para } i = 1, \dots, m - 1, \\ 0, & \text{para } t = T, \end{cases}$$

$$z_Q(t) = \begin{cases} \hat{z}(t_{i-1}+), & t \in [t_{i-1}, \frac{t_i + t_{i-1}}{2}) \quad \text{para } i = 1, \dots, m \\ \hat{z}(t_i-), & t \in [\frac{t_i + t_{i-1}}{2}, t_i) \quad \text{para } i = 1, \dots, m \\ \hat{z}(t_m-), & \text{para } t = T. \end{cases}$$

Sejam

$$\omega^T(t) = (x_Q^T(t), y_Q^T(t), z_Q^T(t))$$

e

$$\alpha(\omega) = \hat{c}^T \hat{\omega} - \int_0^T c^T(t) x_Q(t) dt. \quad (4.15)$$

O próximo teorema mostra uma propriedade para $\omega(t)$.

TEOREMA 4.3.4 $\omega(t)$ é factível para PLCS. Além disso,

$$\alpha(\omega) = \sum_{i=1}^m \left(\frac{(t_i - t_{i-1})^2}{8} \right) (x(t_i) - x(t_{i-1}))^T \dot{c}(t_{i-1})$$

Prova: Pelas definições dadas é claro que $\omega(t)$ é factível para PLCS.

Sabe-se que $c(t)$ é linear por partes e $x_Q(t)$ é constante por partes, assim

$$\begin{aligned} \alpha(\omega) &= \hat{c}^T \hat{\omega} - \int_0^T c^T(t) x_Q(t) dt \\ &= \sum_{i=1}^m \left(\frac{t_i - t_{i-1}}{2} \right) (c^T(t_{i-1}+) (\hat{x}(t_{i-1}+) + c^T(t_i-) (\hat{x}(t_i-) \\ &\quad - (\sum_{i=1}^m [\int_{t_{i-1}}^{u_i} c^T(t) \hat{x}(t) dt + \int_{u_i}^{t_i} c^T(t) \hat{x}(t) dt]) \\ &= \sum_{i=1}^m \left\{ \left(\frac{t_i - t_{i-1}}{2} \right) (c^T(t_{i-1}+) \hat{x}(t_{i-1}+) + c^T(t_i-) \hat{x}(t_i-) \right. \\ &\quad - (\int_{t_{i-1}}^{u_i} c^T(t_{i-1}+) \hat{x}(t_{i-1}+) + (c^T(t_{i-1}+) \hat{x}(t_{i-1}+))' (t - t_{i-1}) dt \\ &\quad \left. + \int_{u_i}^{t_i} c^T(t_i-) \hat{x}(t_i-) + (c^T(t_i-) \hat{x}(t_i-))' (t - t_i) dt) \right\} \end{aligned}$$

$$\begin{aligned}
&= \sum_{i=1}^m \left\{ \left(\frac{t_i - t_{i-1}}{2} \right) (c^T(t_{i-1}+) \hat{x}(t_{i-1}+) + c^T(t_i-) \hat{x}(t_i-)) \right. \\
&\quad - \left[\left(\frac{t_i - t_{i-1}}{2} \right) c^T(t_{i-1}+) \hat{x}(t_{i-1}+) + \dot{c}^T(t_{i-1}+) \hat{x}(t_{i-1}+) \int_{t_{i-1}}^{u_i} (t - t_{i-1}) dt \right. \\
&\quad \left. \left. + \left(\frac{t_i - t_{i-1}}{2} \right) c^T(t_i-) \hat{x}(t_i-) + \dot{c}^T(t_i-) \hat{x}(t_i-) \int_{u_i}^{t_i} (t - t_i) dt \right] \right\}
\end{aligned}$$

Cancelando termo a termo, temos

$$\begin{aligned}
\alpha(\omega) &= - \sum_{i=1}^m \left(\dot{c}^T(t_{i-1}+) \hat{x}(t_{i-1}+) \frac{(t_i - t_{i-1})^2}{2} / u_{i-1} + \dot{c}^T(t_i-) \hat{x}(t_i-) \frac{(t_i - t_i)^2}{2} / u_i \right) \\
&= - \sum_{i=1}^m \left[\dot{c}^T(t_{i-1}+) \hat{x}(t_{i-1}+) \frac{(t_i - t_{i-1})^2}{8} + \dot{c}^T(t_i-) \hat{x}(t_i-) \left(- \frac{(t_{i-1} - t_i)^2}{8} \right) \right]
\end{aligned}$$

Por hipótese $c(t)$ é linear, logo $\dot{c}^T(t_{i-1}+) = \dot{c}^T(t_i-)$, então

$$\alpha(\omega) = - \sum_{i=1}^m \left(\frac{(t_i - t_{i-1})^2}{8} (\hat{x}(t_i-) - \hat{x}(t_{i-1}+)) \dot{c}(t_i-) \right).$$

□

Considerando-se o que acontece quando a partição é refinada progressivamente, do Teorema 4.3.4 e pela hipótese 2, pode-se deduzir que, como a norma da partição tende para zero, os valores do PLCS e MD tendem para o mesmo valor. Disto resulta que

$$V[PLCS] = V[PLCS^*]$$

como uma consequência imediata.

Será necessário o seguinte resultado técnico sobre sequências, o qual pode ser provado facilmente.

LEMA 4.3.4 Supor que $\{a_n\}_{n=1}^\infty$ e $\{b_n\}_{n=1}^\infty$, são sequências tal que

$$b_n - a_m \geq 0, \quad m, n = 1, 2, \dots,$$

e

$$\lim_{n \rightarrow \infty} (b_n - a_n) = 0.$$

Então $\lim_{n \rightarrow \infty} a_n$ e $\lim_{n \rightarrow \infty} b_n$ existe e

$$\lim_{n \rightarrow \infty} a_n = \lim_{n \rightarrow \infty} b_n.$$

Prova: Seja

$$\epsilon = \inf_n b_n - \sup_n a_n,$$

então $\epsilon \geq 0$. Desde que $b_n - a_n \geq \epsilon$ para cada n e $\lim(b_n - a_n) = 0$ temos $\epsilon = 0$.

Seja

$$L = \inf_n b_n = \sup_n a_n.$$

Desde que $a_m \leq L \leq b_n$ para todo m , temos

$$(L - a_n) \leq (b_n - a_n) \quad e \quad (b_n - L) \leq (b_n - a_n).$$

Aplicando o limite quando $n \rightarrow \infty$ nas duas desigualdades acima temos:

$$\lim_{n \rightarrow \infty} (L - a_n) \leq \lim_{n \rightarrow \infty} (b_n - a_n) \quad e \quad \lim_{n \rightarrow \infty} (b_n - L) \leq \lim_{n \rightarrow \infty} (b_n - a_n)$$

com $\lim_{n \rightarrow \infty} (b_n - a_n) = 0$, segue que

$$\lim_{n \rightarrow \infty} L - \lim_{n \rightarrow \infty} a_n \leq 0$$

e

$$\lim_{n \rightarrow \infty} b_n - \lim_{n \rightarrow \infty} L \leq 0.$$

Substituindo L por $\inf_n b_n$ na primeira desigualdade e por $\sup_n a_n$ na segunda desigualdade obtém-se:

$$\lim_{n \rightarrow \infty} (\inf_n b_n) \leq \lim_{n \rightarrow \infty} a_n$$

e

$$\lim_{n \rightarrow \infty} b_n \leq \lim_{n \rightarrow \infty} (\sup_n a_n)$$

Logo

$$\lim_{n \rightarrow \infty} b_n \leq \lim_{n \rightarrow \infty} a_n.$$

Como $a_m \leq b_n$ por hipótese segue que

$$\lim_{n \rightarrow \infty} b_n = \lim_{n \rightarrow \infty} a_n$$

□

COLORÁRIO 4.3.2 Supondo que a Hipótese 2 esteja satisfeita, e que $\omega(t)$ seja factível para PLCS com $x(t)$ constante por partes com partição P , então

$$\lim_{\substack{\|Q\| \rightarrow 0 \\ P \subseteq Q}} V[MD(Q)] = \lim_{\substack{\|Q\| \rightarrow 0 \\ P \subseteq Q}} \int_0^T c^T(t) x_Q(t) dt, \quad (4.16)$$

onde $x_Q(t)$ é obtido da solução ótima para $MD(Q)$ por (4.14), e assim

$$\lim_{\substack{\|Q\| \rightarrow 0 \\ P \subseteq Q}} \int_0^T c^T(t) x_Q(t) dt = V[PLCS] = V[PLCS^*]. \quad (4.17)$$

Prova:

Suponha que para algum $x(t)$ factível, $\|x(t)\| \leq M$ e que $\|\dot{c}(t)\| \leq C$ para $t \in [0, T] - P$.

Seja $Q = \{s_1, \dots, s_k\} \supseteq P$. Então

$$|\alpha(\omega_Q)| \leq \frac{MC}{4} \sum_{i=1}^k (s_i - s_{i-1})^2 \leq \frac{MC\|Q\|T}{4}$$

$\alpha(\omega_Q) \rightarrow 0$ quando $\|Q\| \rightarrow 0$.

Seja $\{Q_n\}_{n=1}^\infty$ uma sequência de partições com

$$\lim_{n \rightarrow \infty} \|Q_n\| = 0, \quad P \subseteq Q_n$$

e

$$a_n = V[MD(Q_n)], \quad b_n = \int_0^T c^T(t) x_{Q_n}(t) dt,$$

Assim $\lim_{n \rightarrow \infty} (b_n - a_n) = 0$, e pelo Teorema 4.3.3 tem-se $b_n - a_m \geq 0$ para todo m e n . Do Lema 4.3.4 segue (4.16). Com uma nova aplicação do Teorema 4.3.3 conclui-se (4.17). $\square$

4.4 Construindo uma solução melhorada

Seja $\omega^T(t) = (x^T(t), y^T(t), z^T(t))$ factível para PLCS com $x(t)$ constante por partes com partição P . Usando a notação da seção anterior, seja $\hat{\omega}^T(t) = (\hat{x}^T(t), \hat{y}^T(t), \hat{z}^T(t))$ solução de MD construída a partir da partição P . Suponha que $\hat{\omega}$ não é ótimo para MD. Então existe $\hat{\hat{\omega}}$, factível para MD, com o valor objetivo rigorosamente melhorado, isto é,

$$\delta \stackrel{\text{def}}{=} c^T \hat{\hat{\omega}} - c^T \hat{\omega} < 0.$$

Seja $\tilde{\omega}^T(t) = (\tilde{x}^T(t), \tilde{y}^T(t), \tilde{z}^T(t))$ uma solução factível para PLCS generalizada por $\hat{\omega}$ como definida na seção anterior. Então $\tilde{\omega}$ é constante em $[t_{i-1}, (t_{i-1} + t_i)/2)$ e em $[(t_{i-1} + t_i)/2, t_i]$ para $i = 1, \dots, m$.

Agora “juntam-se” as duas soluções $\omega(t)$ e $\tilde{\omega}(t)$ como segue.

	t_0	Time
$\frac{\tilde{x}(t)}{x(t)}$	$t_0 + \epsilon_1$	$\downarrow$
	$t_1 - \epsilon_1$	
$\tilde{x}(t)$	t_1	
	$t_1 + \epsilon_2$	
$\vdots$	$\vdots$	
	$t_{m-1} - \epsilon_{m-1}$	
$\tilde{x}(t)$	t_{m-1}	(4.18)
$x(t)$	$t_{m-1} + \epsilon_m$	(4.19)
$\tilde{x}(t)$	$t_m - \epsilon_m$	(4.20)
	t_m	

Prova: Estabelecemos estas igualdades por indução sobre i .
 Seja
 Este caso que para $i = 0$, $\tilde{y}_i(0) = \tilde{y}(0) = y(0) = y(t_0) = y(0)$, (4.20) é verdadeira. Supõe-se que (4.20) seja verdadeira para $i - 1$, isto é,

e $\epsilon \in [0, \tau]$. Define-se

$$\epsilon_i = \frac{(t_i - t_{i-1})\epsilon}{2\tau}.$$

Note que se $\epsilon = \tau$ então $\epsilon_i = (t_i - t_{i-1})/2$. Define-se

$$\bar{x}_\epsilon(t) = \begin{cases} \tilde{x}(t), & t \in [t_{i-1}, t_{i-1} + \epsilon) \cup [t_i - \epsilon, t_i) \ i = 1, \dots, m, \\ x(t), & \text{caso contrário.} \end{cases}$$

Sejam $\bar{y}_\epsilon(t)$ e $\bar{z}_\epsilon(t)$ dados pelas restrições de PLCS, isto é,

$$\int_0^t G\bar{x}_\epsilon(s)ds + \bar{y}_\epsilon(t) = a(t),$$

$$H\bar{x}_\epsilon(t) + \bar{z}_\epsilon(t) = b(t), \quad t \in [0, T].$$

Afirmção : $\bar{\omega}_\epsilon^T(t) = (\bar{x}_\epsilon^T(t), \bar{y}_\epsilon^T(t), \bar{z}_\epsilon^T(t))$ é factível para PLCS.

De fato, como $x(t), \tilde{x}(t) \geq 0$ para todo $t \in [0, T]$ pela factibilidade de $\omega(t)$ e $\tilde{\omega}(t)$, é claro que $\bar{x}_\epsilon(t) \geq 0$ para todo $t \in [0, T]$. Para $t \in [t_{i-1} + \epsilon_i, t_i - \epsilon_i]$, $\bar{x}_\epsilon(t) = x(t)$, assim $\bar{z}_\epsilon(t) = z(t)$ para $t \in [t_{i-1} + \epsilon_i, t_i - \epsilon_i]$ e $\bar{x}_\epsilon(t) = \tilde{x}(t)$ e $\bar{z}_\epsilon(t) = \tilde{z}(t)$, caso contrário. Pela factibilidade de $\omega(t)$ e $\tilde{\omega}(t)$ novamente, tem-se que $\bar{z}_\epsilon(t) \geq 0$, para $t \in [0, T]$.

Assim, para estabelecer a factibilidade de $\bar{\omega}_\epsilon(t)$ resta mostrar que $\bar{y}_\epsilon(t) \geq 0$.

O fato de $\bar{x}_\epsilon(t)$ ser constante no intervalo $[t_{i-1}, t_{i-1} + \epsilon_i]$, $[t_{i-1} + \epsilon_i, t_i - \epsilon_i]$ e $[t_i - \epsilon_i, t_i]$ para $i = 1, \dots, m$, implica que $\bar{y}_\epsilon(t)$ é linear nesses intervalos. Assim, é suficiente mostrar que $\bar{y}_\epsilon(t_{i-1}), \bar{y}_\epsilon(t_{i-1} + \epsilon_i), \bar{y}_\epsilon(t_i - \epsilon_i)$ e $\bar{y}_\epsilon(t_i) \geq 0$ para $i = 1, \dots, m$. Isto será feito no próximo lema.

LEMA 4.4.1

$$\bar{y}_\epsilon(t_{i-1} + \epsilon_i) = (1 - \frac{\epsilon}{\tau})y(t_{i-1}) + \frac{\epsilon}{\tau}\tilde{y}(\frac{t_i + t_{i-1}}{2}), \quad i = 1, \dots, m, \quad (4.18)$$

$$\bar{y}_\epsilon(t_i - \epsilon_i) = (1 - \frac{\epsilon}{\tau})y(t_i) + \frac{\epsilon}{\tau}\tilde{y}(\frac{t_i + t_{i-1}}{2}), \quad i = 1, \dots, m, \quad (4.19)$$

$$\bar{y}_\epsilon(t_i) = (1 - \frac{\epsilon}{\tau})y(t_i) + \frac{\epsilon}{\tau}\tilde{y}(t_i), \quad i = 0, \dots, m, \quad (4.20)$$

Prova: Estabelecemos estas igualdades por indução sobre i .

Está claro que para $i = 0$, $\bar{y}_\epsilon(0) = \tilde{y}(0) = y(0) = a(0)$, (4.20) é verdadeira. Supõe-se que (4.20) seja verdadeira para $i - 1$, isto é,

$$\bar{y}_\epsilon(t_{i-1}) = (1 - \frac{\epsilon}{\tau})y(t_{i-1}) + \frac{\epsilon}{\tau}\tilde{y}(t_{i-1}), \quad i = 0, \dots, m.$$

Mostra-se que (4.18) - (4.20) é verdadeira para i .

Observar que $\bar{x}_\epsilon(t) = \tilde{x}(t)$, para $t \in [t_{i-1}, t_{i-1} + \epsilon_i]$, logo

$$\begin{aligned} \bar{y}_\epsilon(t_{i-1} + \epsilon_i) - \bar{y}_\epsilon(t_{i-1}) &= \tilde{y}(t_{i-1} + \epsilon_i) - \tilde{y}(t_{i-1}) \\ &= \frac{2\epsilon_i}{t_i - t_{i-1}}[\tilde{y}(\frac{t_i + t_{i-1}}{2}) - \tilde{y}(t_{i-1})], \end{aligned} \quad (4.21)$$

pois, $\tilde{y}$ é linear em $[t_{i-1}, \frac{t_i + t_{i-1}}{2}]$. Assim,

$$\bar{y}_\epsilon(t_{i-1} + \epsilon_i) - \bar{y}_\epsilon(t_{i-1}) = \frac{\epsilon}{\tau}[\tilde{y}(\frac{t_i + t_{i-1}}{2}) - \tilde{y}(t_{i-1})]. \quad (4.22)$$

Usando a hipótese de indução e substituindo em (4.22), obtém-se

$$\bar{y}_\epsilon(t_{i-1} + \epsilon_i) - (1 - \frac{\epsilon}{\tau})y(t_{i-1}) - \frac{\epsilon}{\tau}\tilde{y}(t_i) = \frac{\epsilon}{\tau}\tilde{y}(\frac{t_i + t_{i-1}}{2}) - \tilde{y}(t_i).$$

Fazendo os cálculos chega-se a

$$\bar{y}_\epsilon(t_{i-1} + \epsilon_i) = (1 - \frac{\epsilon}{\tau})y(t_{i-1}) + \frac{\epsilon}{\tau}\tilde{y}(\frac{t_i + t_{i-1}}{2}).$$

Isto é, (4.18) é satisfeita para i .

Analogamente, $\bar{x}_\epsilon(t) = x(t)$ em $[t_{i-1} + \epsilon_i, t_i - \epsilon_i]$, então

$$\begin{aligned}\bar{y}_\epsilon(t_i - \epsilon_i) - \bar{y}_\epsilon(t_{i-1} + \epsilon_i) &= y(t_i - \epsilon_i) - y(t_{i-1} + \epsilon_i) \\ &= (1 - \frac{2\epsilon_i}{t_i - t_{i-1}})(y(t_i) - y(t_{i-1})).\end{aligned}$$

Novamente usa-se o fato que, y é linear em $[t_{i-1}, t_i]$.

Assim como (4.20) é válido para i (4.19) também é.

Argumentos semelhantes mostram que (4.19) é verdadeira para i e segue pelo princípio da indução finita. $\square$

COLORÁRIO 4.4.1 $\bar{\omega}_\epsilon$ é factível para todo $\epsilon \in [0, \tau]$.

Prova: Basta mostrar que $\bar{y}_\epsilon(t) \geq 0$ para todo $t \in [0, T]$. Sabemos que $\bar{y}_\epsilon(t_i)$ é uma combinação convexa de $y(t_i)$ e $\tilde{y}(t_i)$ e como $\omega(t)$ e $\tilde{\omega}(t)$ são factíveis para PLCS, então $y(t_i), \tilde{y}(t_i) \geq 0$. Logo, por (4.18) temos $\bar{y}_\epsilon(t_i) \geq 0$ para $i = 1, \dots, m$.

Similarmente, $\bar{y}_\epsilon(t_{i-1} + \epsilon)$ e $\bar{y}_\epsilon(t_i - \epsilon) \geq 0$ para $i = 1, \dots, m$. O resultado segue da linearidade por partes de $\bar{y}_\epsilon$. $\square$

Tendo estabelecido a factibilidade de $\bar{\omega}_\epsilon$, o próximo passo é mostrar que para ϵ suficientemente pequeno é possível melhorar o valor da função objetivo. Isto é feito nos próximos resultados.

LEMA 4.4.2

$$\int_0^T c^T(t)\bar{x}_\epsilon(t)dt - \int_0^T c^T(t)x(t)dt = \frac{\epsilon}{\tau}(\delta - \frac{\epsilon\alpha}{\tau}),$$

onde $\alpha = \alpha(\tilde{\omega})$, como dado em (4.11).



Prova: Como $c(t)$ é linear em (t_{i-1}, t_i) para $i = 1, \dots, m$, temos

$$\begin{aligned}\int_{t_{i-1}}^{t_{i-1}+\epsilon_i} c(t)dt &= \frac{\epsilon_i}{2}(c(t_{i-1}+) + c(t_{i-1} + \epsilon_i)) \\ &= \epsilon_i c(t_{i-1}+) + \frac{\epsilon^2(t_i - t_{i-1})^2}{8\tau^2} \dot{c}(t_{i-1}).\end{aligned}$$

Analogamente, tem-se

$$\int_{t_i-\epsilon_i}^{t_i} c(t)dt = \epsilon_i c(t_i-) - \frac{\epsilon^2(t_i - t_{i-1})^2}{8\tau^2} \dot{c}(t_i-).$$

Agora para $t \in [0, T)$,

$$\bar{x}_\epsilon(t) - x(t) = \begin{cases} \tilde{x}(t_{i-1}+) - x(t_{i-1}+), & t \in [t_{i-1}, t_{i-1} + \epsilon_i), i = 1, \dots, m, \\ \tilde{x}(t_i-) - x(t_i-), & t \in [t_i + \epsilon_i, t_i), i = 1, \dots, m, \\ 0, & \text{caso contrário.} \end{cases}$$

Então

$$\begin{aligned}& \int_0^T c^T(t) \bar{x}_\epsilon(t) dt - \int_0^T c^T(t) x(t) dt \\ &= \sum_{i=1}^m \int_{t_{i-1}}^{t_{i-1}+\epsilon_i} c^T(t) (\tilde{x}(t) - x(t)) dt + \sum_{i=1}^m \int_{t_i-\epsilon_i}^{t_i} c^T(t) (\tilde{x}(t) - x(t)) dt \\ &= \sum_{i=1}^m (\tilde{x}(t_{i-1}+) - x(t_{i-1}+))^T + (\epsilon_i c(t_{i-1}+) + \frac{\epsilon^2(t_i - t_{i-1})^2}{8\tau^2} \dot{c}(t_{i-1})) \\ & \quad + \sum_{i=1}^m (\tilde{x}(t_i-) - x(t_i-))^T + (\epsilon_i c(t_i-) + \frac{\epsilon^2(t_i - t_{i-1})^2}{8\tau^2} \dot{c}(t_i-)).\end{aligned}$$

O resultado agora segue de cálculos algébricos e pelo fato de que

$$\begin{aligned}\delta &= \hat{c}^T \hat{\omega} - \hat{c}^T \hat{\omega} \\ &= \sum_{i=1}^m \left(\frac{t_i - t_{i-1}}{2} \right) (\tilde{x}(t_{i-1}+) - x(t_{i-1}+))^T + c(t_{i-1}+) \\ & \quad + \sum_{i=1}^m \left(\frac{t_i - t_{i-1}}{2} \right) (\tilde{x}(t_i-) - x(t_i-))^T + c(t_i-).\end{aligned}$$

□

COLORÁRIO 4.4.2 Para ϵ suficientemente pequeno, $\int_0^T c^T(t)\bar{x}_\epsilon(t)dt < \int_0^T c^T(t)x(t)dt$ e

$$\min_{\epsilon} \int_0^T c^T(t)\bar{x}_\epsilon(t)dt - \int_0^T c^T(t)x(t)dt = \begin{cases} \frac{\delta^2}{4\alpha}, & \alpha < 0 \text{ e } \frac{\delta}{2\alpha} < 1, \\ \delta - \alpha, & \text{caso contrário,} \end{cases}$$

e o mínimo ocorre em

$$\epsilon^* = \begin{cases} \frac{\delta\tau}{2\alpha}, & \alpha < 0 \text{ e } \frac{\delta}{2\alpha} < 1, \\ \tau, & \text{caso contrário.} \end{cases}$$

Temos agora um Teorema de dualidade forte.

TEOREMA 4.4.1 (Dualidade Forte) Supõe que a hipótese 2 seja satisfeita. Se $\omega(t)$ é ótima para PLCS, então existe θ ótimo para PLCS* com

$$\int_0^T c^T(t)x(t)dt = - \int_0^T d\pi^T(t)a(t) - \int_0^T \eta^T(t)b(t)dt.$$

Prova: Se $\omega(t)$ é uma solução ótima de PLCS então podemos supor que $x(t)$ é constante por partes. A solução correspondente $\hat{\omega}$ é então ótima para MD(P), pois, caso contrário, poder-se-ia contruir uma solução melhor para PLCS. Agora o resultado segue do Colorário 4.3.1. $\square$

4.5 Implementação

Temos todos os ingredientes para implementar um algoritmo para resolver problemas de programação linear contínua separável. Nesta seção esquematizam-se possíveis caminhos diferentes, juntando esses ingredientes. A discussão dos resultados computacionais de tal implementação é objeto do próximo capítulo.

1. Discretização pura.

Resolve-se MD(P) para alguma partição P e constrói-se uma solução factível para PLCS (veja Teorema 4.3.2). Esta discretização tem duas vantagens sobre a discretização padrão, esquematizada na seção §1. Primeiro, pelo Teorema 4.3.3, a função custo da solução de MD fornece uma cota inferior para a função custo da solução ótima



do PLCS.

Segundo, se a Hipótese 2 estiver satisfeita, então, usando o Teorema 4.3.4 e o Corolário 4.3.2 pode-se construir uma partição de modo que a solução do PLCS esteja tão próxima da solução ótima quanto se deseja. Solução ótima.

2. Tempo contínuo

Dada uma solução factível para PLCS, constrói-se uma solução melhorada usando o procedimento detalhado na seção 4.4. Repete-se o procedimento refinando as partições até que a solução fique próxima da solução ótima, dentro de um limite pré-estabelecido. Há várias maneiras de se fazer isso. Algumas delas são resumidas abaixo:

- (a) Pelo Lema 4.2.1, uma solução do PLCS gera uma solução para $PD(P)$, para alguma partição P , com o mesmo custo. $PD(P)$ pode ser resolvido, ou pelo menos melhorado e sua solução usada como solução factível inicial para PLCS, da qual uma solução melhorada para a PLCS é obtida pelo procedimento esquematizado na §4.4. Resolvendo-se $PD(P)$ completamente, produzir-se-á a melhor solução para PLCS com a partição P .
- (b) Pelo Teorema 4.3.2, uma solução para PLCS gera uma solução factível para $MD(P)$ para alguma partição P , com o mesmo custo. Resolvendo $MD(P)$ até a otimalidade, usando isso como uma solução melhorada para $MD(P)$ e “juntando” as duas soluções produz um limite inferior para a função objetiva.
- (c) Dada uma solução do PLCS, constrói-se uma solução do $MD(P)$, purifique-a tornando-a uma solução básica factível e executa-se (ou apenas alguns) pivoteamentos simplex. Isto produz uma solução melhorada de MD , que, depois de “juntar” com a solução velha, produzirá uma solução melhorada para PLCS.

Ainda não é claro qual dos procedimentos acima produzirá os melhores resultados. Entretanto, uma mudança simples que aplicada em qualquer implementação sugerida acima parece melhorar o tempo de solução consideravelmente e a convergência dos pontos da partição, é a remoção de pontos desnecessários à partição. Depois de completada cada interação do algoritmo verifica-se se a nova solução do PLCS (tanto quanto $\dot{a}$, $b(t)$, e $\dot{c}$) é constante além



de dois ou mais intervalos da nova partição . Se for verdade, então removem-se os pontos da partição de modo que existirá somente um no tal intervalo. Isto produzirá uma nova partição com um menor número de pontos, e com cada ponto correspondendo a uma descontinuidade em um ou mais componentes de $x(t)$, $\dot{a}$, $b(t)$, $\dot{c}(t)$.

Resultados computacionais são apresentados no próximo capítulo.

Resultados numéricos e considerações finais

Neste capítulo serão apresentados a solução numérica de um exemplo simples não separável utilizando a discretização estudada no capítulo 3 e um exemplo separável utilizando ingredientes básicos de um algoritmo estudados no capítulo 4.

1 Exemplo não separável

Considera-se o PPLC, ver ([7]), com restrições, funções e matrizes dadas no intervalo $0 \leq t \leq 1$:

$$x^p(t) = (3 - 3t, 4 - 4t, 2 - 2t)^T$$

$$a(t) = (12t, 2t, 9t)^T$$

$$B(t) = \begin{pmatrix} 1 & 5 & 0 & 1 & 0 & 0 \\ -1 & 2 & 0 & 0 & 1 & 0 \\ 3 & 0 & 1 & 0 & 0 & 1 \end{pmatrix}, \quad -K(t) = \begin{pmatrix} 1 & 10 & 0 & 0 & 0 & 0 \\ -1 & 4 & 0 & 0 & 0 & 0 \\ 3 & 0 & 3 & 0 & 0 & 0 \end{pmatrix}.$$

A solução analítica deste problema é dado por:

$$x_1(t) = \begin{cases} 2 - 2e^{-t} & , \quad 0 \leq t < T_0 \\ 0 & , \quad T_0 \leq t \leq 1 \end{cases}$$

$$x_2(t) = \begin{cases} 1 - e^{-2t} & , \quad 0 \leq t < T_0 \\ e^{T_0-2t} - \frac{1}{2}e^{2T_0-2t} - e^{-2t} + \frac{1}{2} & , \quad T_0 \leq t \leq 1 \end{cases}$$



Capítulo 5

Resultados numéricos e considerações finais

Neste capítulo serão apresentados a solução numérica de um exemplo simples não separável utilizando a discretização estudada no capítulo 3 e um exemplo separável utilizando os ingredientes básicos de um algoritmo estudados no capítulo 4.

5.1 Exemplo não separável

Considera-se o PPLC, ver ([7]), com restrições, funções e matrizes dadas no intervalo $0 \leq t \leq 1$:

$$\begin{aligned}c^T(t) &= (3 - 3t, 4 - 4t, 2 - 2t) \\a(t) &= (12t, 2t, 9t)^T \\B(t) &= \begin{pmatrix} 1 & 5 & 0 & 1 & 0 & 0 \\ -1 & 2 & 0 & 0 & 1 & 0 \\ 3 & 0 & 1 & 0 & 0 & 1 \end{pmatrix}, \quad -K(t) = \begin{pmatrix} 1 & 10 & 0 & 0 & 0 & 0 \\ -1 & 4 & 0 & 0 & 0 & 0 \\ 3 & 0 & 3 & 0 & 0 & 0 \end{pmatrix}.\end{aligned}$$

A solução analítica deste problema é dado por:

$$\begin{aligned}x_1(t) &= \begin{cases} 2 - 2e^{-t} & , \quad 0 \leq t < T_0 \\ 0 & , \quad T_0 \leq t \leq 1 \end{cases} \\x_2(t) &= \begin{cases} 1 - e^{-2t} & , \quad 0 \leq t < T_0 \\ e^{T_0-2t} - \frac{1}{2}e^{2T_0-2t} - e^{-2t} + \frac{1}{2} & , \quad T_0 \leq t \leq 1 \end{cases}\end{aligned}$$

$$x_2(t) = \begin{cases} 1 - e^{-2t} & , \quad 0 \leq t < T_0 \\ e^{T_0-2t} - \frac{1}{2}e^{2T_0-2t} - e^{-2t} + \frac{1}{2} & , \quad T_0 \leq t \leq 1 \end{cases}$$

$$x_3(t) = \begin{cases} 1 - e^{-3t} & , \quad 0 \leq t < T_0 \\ 6e^{2T_0-3t} + 4e^{3T_0-3t} - e^{-3t} + 3 & , \quad T_0 \leq t \leq 1 \end{cases}$$

Constrói-se um programa usando o software Mathematica para obter a solução do problema discreto (P_n), correspondente á solução do problema primal (PLC). Neste exemplo usa-se $l_n = 30$, l_n número de segmentos da discretização do intervalo $[0, 1]$, estes segmentos são normalmente equidistantes, t_i , $i = 0, \dots, 30$. As matrizes dada são de ordem 3, e assim o problema discreto tem 93 variáveis e 93 restrições .

Os passos básicos do programa para construir o problema discreto são os seguintes:

1. Função objetiva

Estendemos a função objetiva do problema PLC, discretizando o intervalo na forma acima e usando a fórmula de quadratura do trapézio, (veja na seção 3.3 §1).

2. Restrições

Estendemos as restrições usando a mesma discretização e usando a fórmula de quadratura do trapézio, (veja na seção 3.3 §2).

3. Solução do problema discreto

Resolve-se o problema PLC com a função objetiva obtida pelo item 1 e com as restrições obtidas pelo item 2, usando o comando CosntrainedMin do software Mathematica.

A solução obtida é:



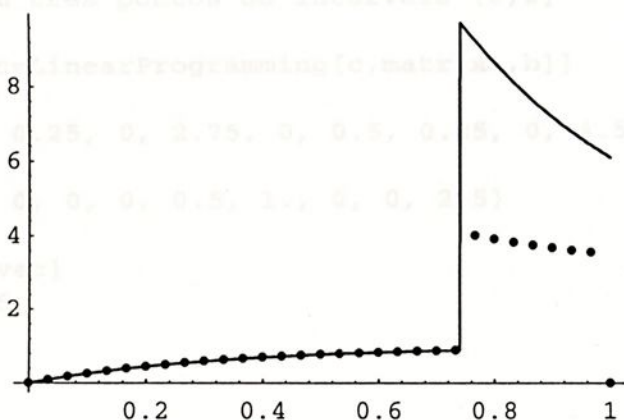
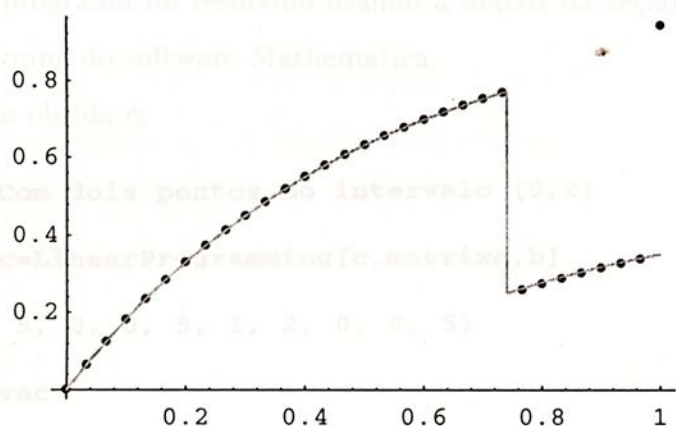
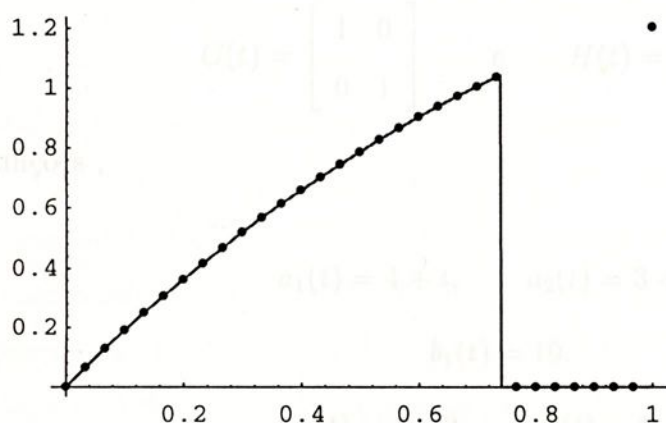
N[ConstrainedMax[a,lista1,lista]]

```
{2.21094, {x[0] -> 0, x[1.] -> 0.0655738, x[2.] -> 0.128998,
x[3.] -> 0.190342, x[4.] -> 0.249675, x[5.] -> 0.307063,
x[6.] -> 0.362569, x[7.] -> 0.416255, x[8.] -> 0.468181,
x[9.] -> 0.518405, x[10.] -> 0.566982, x[11.] -> 0.613966,
x[12.] -> 0.65941, x[13.] -> 0.703363, x[14.] -> 0.745876,
x[15.] -> 0.786995, x[16.] -> 0.826766, x[17.] -> 0.865232,
x[18.] -> 0.902438, x[19.] -> 0.938423, x[20.] -> 0.973229,
x[21.] -> 1.00689, x[22.] -> 1.03945, x[23.] -> 0, x[24.] -> 0,
x[25.] -> 0, x[26.] -> 0, x[27.] -> 0, x[28.] -> 0, x[29.] -> 0,
x[30.] -> 1.20548, x[31.] -> 0, x[32.] -> 0.0645161,
x[33.] -> 0.12487, x[34.] -> 0.18133, x[35.] -> 0.234147,
x[36.] -> 0.283557, x[37.] -> 0.329779, x[38.] -> 0.373019,
x[39.] -> 0.41347, x[40.] -> 0.45131, x[41.] -> 0.48671,
x[42.] -> 0.519825, x[43.] -> 0.550804, x[44.] -> 0.579785,
x[45.] -> 0.606895, x[46.] -> 0.632257, x[47.] -> 0.655982,
x[48.] -> 0.678177, x[49.] -> 0.69894, x[50.] -> 0.718363,
x[51.] -> 0.736533, x[52.] -> 0.753531, x[53.] -> 0.769432,
x[54.] -> 0.25747, x[55.] -> 0.273117, x[56.] -> 0.287755,
x[57.] -> 0.301448, x[58.] -> 0.314258, x[59.] -> 0.326241,
x[60.] -> 0.337452, x[61.] -> 0.940959, x[62.] -> 0,
x[63.] -> 0.0952381, x[64.] -> 0.181406, x[65.] -> 0.259367,
x[66.] -> 0.329904, x[67.] -> 0.393722, x[68.] -> 0.451463,
x[69.] -> 0.503705, x[70.] -> 0.550971, x[71.] -> 0.593736,
x[72.] -> 0.632427, x[73.] -> 0.667434, x[74.] -> 0.699107,
x[75.] -> 0.727764, x[76.] -> 0.753691, x[77.] -> 0.777149,
x[78.] -> 0.798373, x[79.] -> 0.817575, x[80.] -> 0.834949,
x[81.] -> 0.850668, x[82.] -> 0.86489, x[83.] -> 0.877758,
x[84.] -> 0.8894, x[85.] -> 4.01078, x[86.] -> 3.91452,
x[87.] -> 3.82742, x[88.] -> 3.74862, x[89.] -> 3.67732,
x[90.] -> 3.61281, x[91.] -> 3.55445, x[92.] -> 0}}
```



Para efeito de visualização plotamos abaixo as soluções analíticas (linha contínua) e a solução numérica (ponto).

A solução numérica tem alguns pontos que estão longe da solução analítica, mas acreditamos que seja um problema do sistema Mathematica.



5.2 Exemplo separável

Considera o problema de programação linear contínua separável, ver ([19]), no intervalo $[0, 2]$, cujas dimensões das variáveis $x(t)$, $y(t)$ e $z(t)$ são $n_1 = 2$, $n_2 = 2$ e $n_3 = 1$, respectivamente.

Dadas as matrizes,

$$G(t) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad e \quad H(t) = \begin{bmatrix} 1 & 2 \end{bmatrix},$$

e as funções ,

$$a_1(t) = 4 + t, \quad a_2(t) = 3 + 2t.$$

$$b_1(t) = 10.$$

$$c_1(t) = t - 2, \quad c_2(t) = t - 2.$$

Este programa foi resolvido usando a matriz da seção 4.3 e usando o comando LinearProgramming do software Mathematica.

A solução obtida é:

1. Com dois pontos do intervalo $[0, 2]$

```
xvec=LinearProgramming[c,matrixa,b]
```

```
{5, 5, 0, 0, 5, 1, 2, 0, 0, 5}
```

```
c.xvec
```

```
-20
```

2. Com três pontos do intervalo $[0, 2]$

```
N[xvec=LinearProgramming[c,matrixa,b]]
```

```
{4.5, 0.25, 0, 2.75, 0, 0.5, 0.25, 0, 4.5, 4., 0.5,
```

```
5.5, 0, 0, 0, 0.5, 1., 0, 0, 2.5}
```

```
N[c.xvec]
```

```
-16.25
```


5.3 Conclusões

Neste trabalho discutem-se métodos de solução de problemas de programação linear contínuo. Um dos caminhos seguidos foi a pura discretização onde o problema contínuo é aproximado por problemas discretos em que as soluções ótimas convergem para a solução do problema contínuo. Outro caminho bastante estudado na literatura é a tentativa de generalização do método simplex para espaços funcionais. Este método foi objeto de estudo desta dissertação, porém estudou-se uma classe de problemas lineares contínuos separáveis, no qual, foi possível combinar os elementos do método simplex com os ingredientes dos métodos de aproximação via discretização. Isto foi possível devido à natureza dos problemas separáveis que, permitem buscar soluções ótimas $\hat{\omega}(t) = (\hat{x}(t), \hat{y}(t), \hat{z}(t))$ em que $\hat{x}(t)$ e $\hat{z}(t)$ são constantes por partes e $\hat{y}(t)$ linear por partes.

No futuro pretende-se testar as diversas implementações sugeridas no quarto capítulo para deduzir qual delas possui o menor tempo de CPU e possivelmente a melhor taxa de convergência.

Além disso, pode-se tentar estender a teoria de dualidade contida no capítulo 4, para problemas quadráticos com restrições separáveis.



Bibliografia

- [1] Anderson, E. J., *A continuous model for job-shop scheduling*, Unpublished, PhD. thesis, University of Cambridge, 1978.
- [2] Anderson, E. J. and Philpott, A. B., *A continuous-time network simplex algorithm*, Networks, v.19, pp. 395-425, 1989.
- [3] Anstreicher, K. M., *Generation of feasible descent directions in continuous-times linear programming*, Tech. Report, SOL 83-18, Department of Operations Research, Stanford University, CA, 1983.
- [4] Bazaraa, M. S. and Jarvis, J. J., *Linear programming and network flows*, School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, Georgia, John Wiley & Sons, 1977.
- [5] Bellmann, R.E., *Dynamic programming*, Princeton, 1972.
- [6] Bellmann, R.E., *Bottleneck problems and dynamic programming*, Proc. Nat. Acad. Sci., v.39, p. 947-951, 1953.
- [7] Buie, R. N. and Abrham, J., *Numerical solutions to continuous linear programming problems*, Z. Oper. Res., v.17, p. 107-117, 1973.
- [8] Dantzig, G.B., *Large scale systems optimization with applications to energy*, Technical Report SOL 77-3, Department of Operations Research, Stanford Unirvesity, Stanford, California, april 1977.



- [9] Drews, W. P., *A simplex-like algorithm for continuous-time linear optimal control problems*, in Optimization Methods for Resource Allocation, R. W. Cottle and J. Krarup, eds., Crane Russak and Co. Inc., New York, p.309-322, 1974.
- [10] Drews, W. P., Hartberger, R. J. and Segers, R. B., *On continuous mathematical programming*, Optimization Methods in Resource Allocation, edited by R.W. Cottle and J.Kraup, Crane, Russak and Co.Inc., N.Y., 1974.
- [11] Grinold, R. C., *Continuous programming*, Doctoral Thesis, University of California Berkeley, 1968.
- [12] Hartberger, R. J., *Representation extended to continuous time*, in Optimization Methods for Resource Allocation, R. W. Cottle and J. Krarup, eds., Crane Russak and Co. Inc., New York, p.297-307, 1974.
- [13] Kohmogorov, A. N., and Fomin, S. V., *Elementos de la Teoria de Funciones y del Análisis Funcional*. Tradução por Carlos Vega. Editora Mir, 1972. Original em russo.
- [14] Lehman, R.S., *On the continuous simplex method*, RAND Reserch memorandum RM-1386, Santa Monica, California, 1954.
- [15] Levine, P. and Pomerol, J.C., *C-closed mappings and Kuhn-Tucher vectors in convex programming*, Discussion paper 7620, Center for Operations Research and Economics, Universite Catholique de Louvain, Heverlee, Belgium 1976.
- [16] Levinson, N., *A class of continuous linear programming problems*, J. Math. Anal. Appl. v.16, pp. 78-83.
- [17] Lojasiewicz S., *An introduction to the theory of real functions*. Tradução por G. H. Lawden (University of Sussex). Editado por A. V. Ferreira (University of Bologna), John Wiley & Sons, 1988. Original em polones.
- [18] Perold, A. F., *Fundamentals of a continuous times simplex method*, Tech.Report, SOL 78-26, Department of Operations Research, Stanford University, CA, 1978.



- [19] Pullan, M. C., *An algorithm for a class of continuous linear programs*, SIAM J. Control Optim., v.31 n.6, p. 1558-1577, november 1993.
- [20] Teren, F., *Minimun time acceleration of aircraft turbofan engines by using an algorithm based on nonlinear programming*, NASA Technical Memorandum TM-73741 Lewis Research Center, Cleveland, Ohio, september 1977.
- [21] Tyndall, W.F., *A duality theorem for a class of continuous linear programming problems*, SIAM J. Applied Mathematics, v.13, n.3 p. 644-666, september 1965.
- [22] Tyndall, W. F., *An extended duality theorem for continuous linear programming problems*, SIAM J. Applied Math., pp. 1294-1298, 1967.
- [23] Segers, R. G., *A generalised function setting for dynamic optimal control problems*, in Optimization Methods for Resource Allocation, R. W. Cottle and J. Krarup, eds., Crane Russak and Co. Inc., New York, p.270-296, 1974.



