

Ivaldiney Garcia de Andrade

**Localização Elipsoidal na Análise de Erro
Devido à Agregação de Variáveis em
Programação Linear**



1910000941





**Localização Elipsoidal na Análise
de Erro Devido à Agregação de Variáveis
em Programação Linear**

Ivaldiney Garcia de Andrade

**Dissertação de Mestrado
Pós-Graduação em Matemática Aplicada
MAP-045**



Localização Elipsoidal na Análise de Erro Devido à Agregação de Variáveis em Programação Linear

Ivaldiney Garcia de Andrade

Orientador: Prof. Dr. Roberto Andreani

Co-orientador: Prof. Dr. Igor S. Litvinchev

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

Observação: Informamos que esta tese, em princípio, foi orientada pelo Prof. Dr. Igor S. Litvinchev. Por razões do seu afastamento, a tese passou a ser orientada pelo Prof. Dr. Roberto Andreani, na sua fase final de preparação.



941/00

São José do Rio Preto (SP)

Agosto de 1999



”... mas nada vai conseguir
mudar o que ficou”.

Renato Russo



A minha avó

Agradecimentos

A Deus, pela força e iluminação.

Ao Prof. Igor Litvinchev, que foi um grande incentivador durante a realização desse trabalho.

Ao Prof. Roberto Andreani, que muito contribuiu com a resolução de diversos problemas deste texto.

À Profa. Socorro Rangel, pelas sugestões e críticas construtivas.

À FAPESP, pelo apoio financeiro.

A todas as demais pessoas, e, em especial, ao amigo Fernando, que, de uma forma ou de outra, contribuíram para a realização deste trabalho.

RESUMO

Para estimar o erro causado pela utilização de um modelo agregado aproximado ao invés do modelo original, várias técnicas foram propostas para o caso de agregação de variáveis em Programação Linear. A maioria delas tenta melhorar o limite clássico de Zipkin[1], obtido sob a suposição da disponibilidade de uma localização tipo simplex do conjunto das soluções ótimas.

Neste texto, o limite de Zipkin é alterado para tratar localizações convexas, com especial atenção para as localizações elipsoidais.

A partir da abordagem de Choi-Goldfarb[10], são propostas várias técnicas para obtenção de elipsóides de tamanho reduzido, e as localizações elipsoidais resultantes são aplicados na análise do erro agregacional.

Por fim, um estudo numérico é apresentado para comparar as diferentes abordagens de cálculo de erro e são discutidas as direções para pesquisas futuras.

ABSTRACT

To estimate the error caused by using an approximate aggregated model instead of the original one, various techniques were proposed for variable-aggregation in linear programming. Most of them try to improve the classical Zipkin's bound [1], obtained under the assumption that a simplex-like localization of the optimal solution set is available.

In this work Zipkin's bound is revised to treat convex localizations. The attention is focused on the solution-containing ellipsoids.

Starting with Choi-Goldfarb approach [10], several techniques to reduce the size of the solution-containing ellipsoids are proposed. The resulting ellipsoids are applied to aggregation error analysis.

A numerical study is presented to compare the different aggregation bounds. The directions for future research are discussed.

Índice

1	Introdução	7
2	Agregação em Programação Linear	11
2.1	Construção do Problema Agregado	12
2.2	Desagregação	16
2.3	Análise de Erro	18
3	Limites <i>a Posteriori</i> e Localizações Lineares	23
3.1	Localizações Genéricas	24
3.2	Refinamento do Limite Superior	26
3.3	Localizações Lineares	28
4	Localizações Elipsoidais	33
4.1	Elipsóides Primais	35
4.2	Elipsóides Duais	51
5	Limites <i>a Posteriori</i> e Localizações Elipsoidais	59
5.1	Análise de Erro sob Localização Elipsoidal	59
5.2	Localizações tipo caixa	63
5.2.1	Localização primal tipo caixa	63
5.2.2	Localização dual tipo caixa	69
5.3	Exemplo	74
5.4	Estudo Numérico	82
6	Conclusão	90



Capítulo 1

Introdução

Apesar do grande avanço que se tem observado na tecnologia computacional, surgem situações, entretanto, em que se depara com recursos computacionais insuficientes para resolução de problemas de otimização de grande porte, tão comuns na prática.

Surge, então, a necessidade de se desenvolver ferramentas para tratar essas classes de problemas.

Com esse intuito, têm surgido técnicas como a decomposição e a agregação/desagregação, que procuram reduzir o tamanho dos problemas, e, portanto, o custo da obtenção da solução.

Abordaremos, neste texto, as técnicas de agregação/desagregação, que consistem, basicamente, de métodos de resolução aproximada de problemas de otimização através da combinação dos dados originais em um modelo auxiliar, de tamanho reduzido, cujos resultados são analisados com o objetivo de se derivar uma estimativa para o valor ótimo da função objetivo do problema original.

Tem-se, com isso, uma ferramenta para dosar o custo computacional da resolução do problema em função da precisão requerida para a sua solução.

A abordagem agregacional tem sido empregada em uma grande variedade de problemas, tais como planejamento da produção e gerenciamento da indústria florestal (vide [4] e suas referências para maiores detalhes).

Além de fornecer soluções aproximadas rápidas para problemas computacionalmente intratáveis, pode-se utilizar a técnica da agregação na análise dos vários níveis de detalhes dos modelos sob otimização, de forma a se obter informações para os diferentes níveis de tomada de decisão[4].

Essa abordagem pode ser empregada, ainda, para testar formulações alternativas para o modelo, principalmente nos casos em que se tem dados iniciais aproximados [4].



A metodologia da agregação desempenha, assim, papel de ferramenta auxiliar na modelagem de problemas.

Pode-se, ainda, utilizar a metodologia agregacional para fornecer uma boa solução inicial para muitos tipos de problemas da Programação Matemática, o que constitui ponto crucial em muitos esquemas de otimização.

A redução do tamanho do problema original é feita agrupando-se variáveis e/ou restrições em quantidades únicas, o que, conseqüentemente reduz a ordem da matriz de restrições do problema.

Considere o problema original escrito como:

$$\begin{aligned} \max \quad & z = c^t x \\ \text{s.a.} \quad & Ax \leq b \\ & x \geq 0 \end{aligned} \tag{1.1}$$

onde $A \in \mathcal{R}^{m \times n}$, $b \in \mathcal{R}^m$, e $x, c \in \mathcal{R}^n$.

O agrupamento de restrições é feito multiplicando-se a matriz A e o vetor b por uma matriz $T \in \mathcal{R}^{l \times m}$, com $l < m$. O problema resultante é:

$$\begin{aligned} \max \quad & z = c^t x \\ \text{s.a.} \quad & TAx \leq Tb \\ & x \geq 0, \quad x \in \mathcal{R}^n \end{aligned} \tag{1.2}$$

Observe que a matriz de restrições de (1.2) tem ordem $l \times n$, apresentando um número menor de linhas (restrições), e, portanto, uma dimensão menor que a dimensão da matriz de restrições original. Não houve alteração na dimensão do vetor x .

Analogamente, podemos agrupar variáveis, utilizando uma matriz $S \in \mathcal{R}^{n \times p}$, $p < n$, que, multiplicada pela matriz A e pelo vetor c , dá origem ao problema:

$$\begin{aligned} \max \quad & z = c^t Sx \\ \text{s.a.} \quad & ASx \leq b \\ & x \geq 0, \quad x \in \mathcal{R}^p \end{aligned} \tag{1.3}$$

Note que, agora, o vetor de variáveis x é tal que $x \in \mathcal{R}^p$ e a matriz de restrições de (1.3) possui ordem $m \times p$, e, portanto, apresenta um número menor de colunas (variáveis) que a matriz original.

A agregação simultânea de variáveis e restrições considera ambas as matrizes S e T , de modo que o problema agregado resultante tem o formato:



$$\begin{aligned}
\max \quad & z = c^t Sx \\
\text{s.a.} \quad & TAsx \leq Tb \\
& x \geq 0, x \in \mathcal{R}^p
\end{aligned} \tag{1.4}$$

onde a matriz de restrições TAS tem ordem $l \times p$.

As matrizes S e T são genericamente chamadas de *matrizes peso*. O capítulo seguinte analisa a obtenção dessas matrizes no caso de agregação de variáveis em Programação Linear.

Constitui ponto chave da metodologia da agregação/desagregação o fato de que o modelo agregado deve ser aproximadamente consistente com o modelo original, isto é, se a solução obtida a partir do modelo reduzido não for ótima para o problema original, ela deve ser, no mínimo, uma boa aproximação para a sua solução ótima.

Dessa forma, uma medida para o erro introduzido pela agregação/desagregação deve ser calculada para avaliar a consistência dos modelos, de modo a se estimar o quanto se afastou da solução ótima do problema inicial.

Para estimar o erro causado pela agregação, várias abordagens têm sido propostas (vide [4, 5] e suas referências). A maioria delas tentam aperfeiçoar os limites introduzidos por Zipkin [1], obtidos para Programação Linear.

Um dos pontos mais importantes no método de Zipkin é a suposição da disponibilidade de certa localização da solução ótima do problema original.

No trabalho original de Zipkin, foi suposto que uma restrição linear tipo ‘mochila’ (ou um número independente de restrições desse tipo) se verifica para a solução ótima, e esta localização pode ser obtida *antes* de o problema agregado ter sido construído.

Por outro lado, se o problema agregado já foi solucionado, isto significa que obtivemos informação adicional sobre o problema original. Então, é natural tentar construir localizações melhores *depois* de resolver o problema agregado.

Tem este texto o objetivo de mostrar que, na computação de limites para a função objetivo utilizando agregação, podemos usar diferentes tipos de localizações, e mais ainda, localizações elipsoidais para a solução ótima podem ser obtidas para o caso de Programação Linear. Usando estas localizações não-lineares, novos limites serão construídos.

O texto se acha estruturado como segue: o Capítulo 2 discorre sobre a estrutura da técnica da agregação aplicada à Programação Linear, enquanto o Capítulo 3 discute a extensão do cálculo de limites a localizações genéricas, e utiliza alguns tipos de localizações lineares na obtenção de expressões analíticas para limites para o valor ótimo da função objetivo do



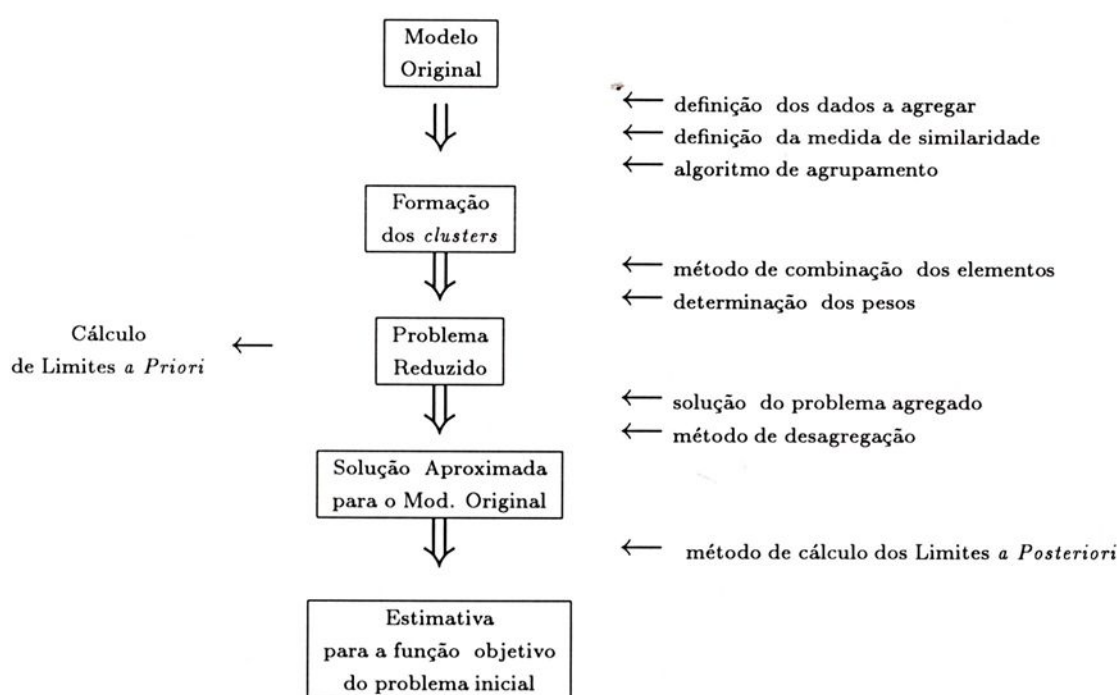
problema inicial. O Capítulo 4 investiga as localizações não-lineares tipo elipsoidais primais e duais, enquanto o Capítulo 5 investiga as consequências de sua aplicação na análise do erro devido à agregação de variáveis, e apresenta um estudo numérico para avaliar a qualidade dos novos limites obtidos. Por fim, o Capítulo 6 tece as considerações finais a respeito do trabalho.



Capítulo 2

Agregação em Programação Linear

A aplicação do esquema agregacional a problemas de programação matemática tem estrutura esquematizada no diagrama abaixo.



As seções seguintes discutem cada um dos elementos da estrutura dessa metodologia aplicada à Programação Linear. A seção 2.1 aborda a construção do problema agregado a partir do problema linear inicial, enquanto a seção 2.2 descreve as técnicas de desagregação, e a seção 2.3 caracteriza as principais metodologias de análise do erro devido à agregação.

2.1 Construção do Problema Agregado

Considere o par de problemas lineares, denominados *problema primal* e *problema dual*, respectivamente:

$$\begin{aligned} \max \quad & c^t x \\ \text{s.a.} \quad & Ax \leq b \\ & x \geq 0 \end{aligned} \tag{2.1}$$

e

$$\begin{aligned} \min \quad & b^t u \\ \text{s.a.} \quad & A^t u \geq c \\ & u \geq 0 \end{aligned} \tag{2.2}$$

onde $c \in \mathcal{R}^n$, $A \in \mathcal{R}^{m \times n}$, $b \in \mathcal{R}^m$, e os vetores $x \in \mathcal{R}^n$ e $u \in \mathcal{R}^m$ são os vetores de variáveis dos problemas (2.1) e (2.2), respectivamente.

Uma abordagem completa sobre o relacionamento dos problemas acima pode ser encontrada em [18, 16].

A primeira consideração que deve ser feita para aplicar os métodos agregacionais à resolução de (2.1) é sobre quais dados devemos agregar: variáveis, restrições, ou variáveis e restrições, simultaneamente.

A resposta a esta pergunta depende do problema sob resolução e da metodologia de solução empregada, de modo que considerações genéricas não são possíveis.

Considere, por exemplo, que se esteja utilizando o Método Simplex na resolução de (2.1). Esse método possui complexidade algorítmica exponencial no pior caso (vide [21] para uma discussão a respeito da complexidade desse método). Entretanto, na prática, a sua complexidade algorítmica média é de $O(m^2n)$ (vide [4] e suas referências). Tem-se, portanto, um crescimento quadrático do esforço computacional em função do número de restrições do problema, o que sugere que a agregação de restrições é a melhor opção quando se está utilizando alguma implementação desse método. No entanto, a agregação de restrições pode gerar soluções infactíveis para o modelo original, conforme discutido em [9].

Neste texto, abordaremos somente a agregação de variáveis (ou colunas), embora muito do que for considerado possa ser estendido, sem dificuldade, para a agregação de restrições. Para maiores esclarecimentos acerca de agregação de restrições, vide [9].

Após a opção pelos dados (variáveis e/ou restrições) a agregar, é necessário considerar como os dados originais serão classificados como similares, ou seja, necessita-se de uma



medida de similaridade para identificar os dados candidatos a serem agregados em um mesmo grupo (ou *cluster*).

No caso da agregação de variáveis (colunas), podemos definir uma distância métrica para medir a similaridade entre duas colunas de (2.1) da forma descrita a seguir, conforme proposto em [11].

Seja um vetor coluna definido por $A^j = (c_j, a_{1j}, \dots, a_{mj})^t$, onde c_j é o coeficiente da função objetivo referente à variável x_j e $a_{1j}, \dots, a_{mj}$ são elementos da coluna j da matriz de restrições.

Então, uma medida da distância entre duas colunas A^j e A^k da matriz A é dada pela métrica:

$$d(A^j, A^k) = 1 - \cos(A^j, A^k) \quad (2.3)$$

onde $\cos(A^j, A^k)$ é o cosseno do ângulo formado pelos vetores A^j e A^k , que pode ser calculado como

$$\cos(A^j, A^k) = \frac{\langle A^j, A^k \rangle}{\|A^j\| \|A^k\|},$$

em que $\langle, \rangle$ indica o produto escalar e $\|.\|$ indica a norma euclidiana.

Observe que, se a distância (2.3) entre cada par de elementos do *cluster* é zero em todos os *clusters*, então a solução obtida da desagregação da solução ótima do problema agregado é ótima para o problema original.

Estabelecida uma medida de similaridade, deve-se, em seguida, decidir sobre o *nível de clustering*, ou seja, o número de *clusters* que será empregado no processo de agrupamento, bem como o procedimento de construção desses *clusters*.

A determinação do número de *clusters* envolve uma comparação entre o esforço computacional requerido para criar e resolver o modelo reduzido e a qualidade da solução obtida para o problema original.

Em geral, o nível ideal de agrupamento é dependente do contexto de resolução do problema.

Por exemplo, se o tamanho do problema original viola uma restrição do software a ser utilizado para resolvê-lo, então a melhor estratégia pode ser agrupar o suficiente para satisfazer esta restrição.

Passamos a descrever um procedimento para construção dos *clusters*, sugerido em [11], para o caso de agregação de colunas, onde os *clusters* serão formados por colunas similares, assim classificadas de acordo com a métrica definida em (2.3).

A primeira coluna selecionada em um *cluster* será chamada de *núcleo do cluster*.

Inicialização - Tome a coluna $j = 1$ como núcleo do *cluster* número 1.

Defina a tolerância ε aceitável para a distância entre duas colunas de um mesmo *cluster*, isto é, se $d(A^j, A^k) \leq \varepsilon$, então as colunas A^j e A^k são candidatas a pertencer ao mesmo *cluster*.

Passo 1 - Tome a coluna A^j e compare-a com os núcleos de todos os *clusters* já definidos.

Se existem *clusters* k cujos núcleos A^k satisfazem $d(A^j, A^k) \leq \varepsilon$, então vá para o passo 2. Caso contrário, vá ao passo 3.

Passo 2 - Escolha o *cluster* com o núcleo mais próximo de A^j , e adicione A^j ao *cluster*.

Faça $j = j + 1$ e, se $j \leq n$, volte ao Passo 1.

Passo 3 - Defina um novo *cluster* com núcleo A^j . Faça $j = j + 1$ e, se $j \leq n$, volte ao Passo 1.

Observe que o *nível de clustering* é definido como resultado dessa construção e dependerá, exclusivamente, da tolerância ε utilizada, de forma que quanto menor o valor de ε , maior o *nível de clustering*.

O algoritmo proposto fornece como resultado final uma partição do conjunto de índices de variáveis, de forma que os *clusters* formados são disjuntos, ou seja, têm intersecção vazia, e sua união recupera o conjunto de índices das variáveis do problema original.

Depois da construção dos *clusters*, deve-se determinar o método de combinação dos seus elementos em quantias únicas, de modo que a cada *cluster* corresponderá uma variável agregada.

O problema de combinação desses elementos consiste na determinação de pesos que refletem a importância relativa dos dados do problema original na criação do problema agregado.

Essa é a característica da agregação que a difere dos algoritmos de decomposição, uma vez que esses últimos preservam os dados originais do problema na construção de subproblemas de ordem menor que o problema inicial, enquanto a agregação combina os dados originais de forma a obter novos dados para a formação do problema reduzido. Um tratamento completo acerca dos algoritmos de decomposição pode ser encontrado em [7].

O procedimento da *combinação de peso fixo* [1, 4] propõe uma combinação linear convexa dos elementos dos *clusters*, por meio da utilização de um vetor de dimensão igual à cardinalidade do *cluster*, com elementos não-negativos e de soma unitária. Estes vetores, aqui



chamado *vetores peso* e denotados g^k , para cada cluster k , dispostos numa matriz com formato bloco-diagonal, constituem a *matriz peso*, que aqui denotaremos simplesmente por g , e que é utilizada para multiplicar a matriz de dados original, de modo a derivar os dados correspondentes para o problema reduzido.

Um procedimento simples para calcular os vetores peso g^k para cada cluster k é o seguinte [4]:

Seja $s^k = \sum_{j \in S_k} c_j^2$, onde S_k denota o cluster k . Então, define-se:

$$g_j^k = \frac{c_j^2}{s^k}, \quad \forall j \in S_k \quad (2.4)$$

Segundo essa abordagem, os dados são combinados de acordo com sua importância na função objetivo, de modo que quanto maior o valor de c_j , maior a influência da variável x_j na construção das restrições do problema agregado.

Outro método para determinar os coeficientes do problema reduzido é chamado *combinação por dominância*[4].

Este método requer pouco esforço computacional e consiste, basicamente, em se tomar como coeficiente da variável agregada em cada restrição do problema reduzido o maior coeficiente das variáveis do cluster na correspondente restrição original, ou seja,

$$\bar{a}_{ik} = \max_{j \in S_k} a_{ij}, \quad (2.5)$$

onde S_k denota o cluster k e $\bar{a}_{ik}$ é o coeficiente da variável agregada correspondente ao cluster k , na linha i do problema agregado, $i = 1, \dots, m$.

Note que a seleção desses coeficientes garante a factibilidade do problema agregado resultante.

Baseados nessas considerações acerca da construção do problema agregado, descrevemos, a seguir, um procedimento para a agregação de variáveis em Programação Linear. Será utilizada notação introduzida por Zipkin em [1].

Suponha que o problema (2.1) tenha solução ótima finita, e seja $\sigma = \{S_k, k = 1, \dots, K\}$ uma partição arbitrária de $\{1, \dots, n\}$, tal que $|S_k| = n_k$.

Defina $\mathcal{A}^k$ a submatriz de A formada das colunas de A cujos índices estão em S_k . Similamente, defina os subvetores c^k e x^k , com elementos ordenados de acordo com a submatriz $\mathcal{A}_k$.

Agora, seja $g^k \in R^{n_k}$, vetor peso com componentes não negativas e de soma unitária, e defina $\bar{A}^k = \mathcal{A}^k g^k$ e $\bar{c}^k = c^{kt} g^k$, para $k = 1, \dots, K$. Note que $\bar{A}^k \in \mathcal{R}^m$ e $\bar{c}^k \in \mathcal{R}$.

Defina, agora, a matriz $\bar{A} = (\bar{A}^1, \dots, \bar{A}^K) = (\bar{a}_{ik})$ e o vetor $\bar{c} = (\bar{c}^1, \dots, \bar{c}^K)$, e seja $X \in \mathcal{R}^K$ um vetor com K variáveis.

Nestas condições, o problema agregado é:

$$\begin{aligned} \max \quad & \bar{c}^t X \\ \text{s.a.} \quad & \bar{A}X \leq b \\ & X \geq 0 \end{aligned} \tag{2.6}$$

e seu dual, que desempenha papel importante em nossa teoria, é escrito como:

$$\begin{aligned} \min \quad & b^t U \\ \text{s.a.} \quad & \bar{A}^t U \geq \bar{c} \\ & U \geq 0 \end{aligned} \tag{2.7}$$

A partição σ e o vetor peso, que podem ser obtidos conforme descrito anteriormente, serão, daqui por diante, considerados fixos.

Ao longo do texto, os vetores x^* e u^* denotarão as soluções ótimas dos problemas originais (2.1) e (2.2), enquanto X^* e U^* denotarão as soluções ótimas dos problemas agregados (2.6) e (2.7), respectivamente. Além disso, z^* indicará o valor ótimo da função objetivo original, enquanto Z^* indicará o valor ótimo da função objetivo do problema agregado.

Após a construção de (2.6), deve-se considerar, em seguida, técnicas de obtenção de uma solução aproximada para o problema original a partir da solução de (2.6).

Esse processo é chamado *desagregação* e será abordado na próxima seção.

2.2 Desagregação

Chamamos *desagregação* ao processo de obtenção de uma solução factível para o problema original (2.1) a partir de uma solução factível do problema agregado (2.6).

Trata-se, portanto, de processo inverso ao da construção do problema reduzido.

O método mais comum de desagregação é análogo ao método de combinação de peso fixo e consiste em se tomar o mesmo vetor utilizado na agregação dos dados para desagregar a solução do problema reduzido, com base na seguinte definição:

Definição 2.1 (*Desagregação de Peso Fixo*)[1] - Seja X factível para (2.6) e defina $x^k \in \mathcal{R}^{n_k}$ tal que $x^k = g^k X_k$, $k = 1, \dots, K$. Seja $\bar{x}$ a coleção dos vetores x^k , arranjados de acordo com a ordem das variáveis de (2.1). Nestas condições, $\bar{x}$ é chamado de solução de peso fixo obtida de X .



A proposição a seguir mostra que a desagregação de peso fixo dá uma solução viável para (2.1).

Proposição 2.1 [1] - Se X é factível para (2.6) e $\bar{x}$ é a solução de peso fixo obtida de X , então $\bar{x}$ é factível para (2.1) e $c^t \bar{x} = \bar{c}^t X$.

Demonstração - Considerando a construção do problema agregado e a definição de desagregação de peso fixo, pode-se escrever:

$$\begin{aligned} A\bar{x} &= \sum_{k=1}^K \mathcal{A}^k(g^k X_k) = \sum_{k=1}^K (\mathcal{A}^k g^k) X_k \\ &= \sum_{k=1}^K \bar{A}_k X_k = \bar{A}X \leq b \end{aligned} \quad (2.8)$$

Além disso,

$$c^t \bar{x} = \sum_{k=1}^K c^{kt}(g^k X_k) = \sum_{k=1}^K (c^{kt} g^k) X_k = \sum_{k=1}^K \bar{c}^k X_k = \bar{c}^t X. \quad \bullet$$

Como consequência da Proposição (2.1), tem-se o seguinte corolário:

Corolário 2.1 [1] - Se X^* é ótimo para (2.6), e $\bar{x}$ é a solução de peso fixo obtida de X^* , então $c^t \bar{x} = Z^* \leq z^*$.

Demonstração - É imediata, se considerarmos que $c^t x \leq c^t x^*$ para todo x factível para (2.1), e se observarmos que a solução desagregada de peso fixo é factível para o problema original, conforme a proposição (2.1). $\bullet$

Dessa forma, o valor ótimo da função objetivo de (2.6) fornece um limite inferior para o valor ótimo da função objetivo de (2.1).

Uma solução desagregada melhor que a solução desagregada de peso fixo pode ser obtida através da *desagregação ótima* de X^* [5], que consiste na resolução dos K subproblemas dados por:

$$\begin{aligned} \max \quad & c^{kt} x^k \\ \text{s.a.} \quad & \mathcal{A}^k x^k \leq \bar{A}^k X_k^* \\ & x^k \geq 0 \end{aligned} \quad (2.9)$$

A solução x^{k*} desses subproblemas dá uma solução desagregada que é, pelo menos, tão boa quanto a solução desagregada de peso fixo, uma vez que a solução desagregada de peso fixo $\bar{x} = (x^1, \dots, x^K)$, obtida da solução ótima X^* de (2.6), satisfaz $\mathcal{A}^k x^k = \mathcal{A}^k g^k X_k^* = \bar{A}_k X_k^*$, para $k = 1, \dots, K$, e, portanto, define uma solução factível para (2.9). Daí, podemos



dizer que $Z^* = c^t \bar{x} = \sum_{k=1}^K c^{kt} x^k \leq \sum_{k=1}^K c^{kt} x^{k*} = \sum_{k=1}^K z_k(X_k^*)$, onde $z_k(X_k^*)$ é o valor da função objetivo ótima do k -ésimo subproblema (2.9).

Assim como a solução desagregada de peso fixo, a solução desagregada ótima é também factível para o problema (2.1).

Dessa forma, tanto o valor ótimo do problema agregado, quanto a soma dos valores ótimos das funções objetivo dos subproblemas (2.9), determinam limites inferiores para o valor ótimo da função objetivo de (2.1).

De fato, podemos escrever:

$$\begin{array}{ccccc} Z^* & \leq & \sum_{k=1}^K z_k(X_k^*) & \leq & z^* \\ \text{desagregação de peso fixo} & & \text{desagregação ótima} & & \text{problema original} \end{array} \quad (2.10)$$

Tendo os limites inferiores, conforme (2.10), faz-se necessária a derivação de limites superiores para z^* . Na seção seguinte, focalizamos as principais abordagens de obtenção desses limites.

2.3 Análise de Erro

O maior problema da metodologia da agregação refere-se ao estudo da informação perdida devido à agregação/desagregação das variáveis, uma vez que, quando um problema agregado é solucionado, geralmente a solução desagregada não resulta em solução ótima para o problema original.

Tem-se, então, que introduzir ferramentas para estimar o erro cometido, de modo a se obter uma medida aproximada da perda de precisão havida em função dos processos agregacionais.

Uma maneira de medir a perda de precisão é a diferença entre os valores das funções objetivo dos problemas original e agregado.

Desde que z^* não é conhecido de antemão, podemos somente estimar $z^* - Z^*$, através da derivação de limites para esta diferença.

Não consideraremos mais a obtenção de limites inferiores para z^* , uma vez que esses podem ser obtidos de acordo com (2.10). Resta-nos, portanto, considerar a obtenção de limites superiores.

Em geral, dois tipos de limites superiores são freqüentemente considerados:



- Limites 'a priori', que são estabelecidos sobre a função objetivo ótima do modelo original depois de o problema agregado ter sido construído, mas antes de ele ter sido resolvido, e
- Limites 'a posteriori', que são estabelecidos depois de o problema agregado ter sido resolvido.

Devido ao fato de os limites *a posteriori* utilizarem a informação adicional da solução do problema agregado, estes são mais precisos que os limites *a priori*, ficando a aplicação desses últimos restrita à fase de construção do problema agregado, como ferramenta para se obter o nível de *clustering* desejado no processo de agregação (vide [4] e suas referências).

Neste texto, serão considerados somente os limites *a posteriori*, isto é, suporemos que o problema agregado tenha sido resolvido e que conhecemos sua solução ótima. Para uma abordagem a respeito de limites *a priori*, vide [1].

O teorema a seguir dá a chave para a obtenção de limites superiores para o valor de z^* , a partir de dados do problema agregado (2.6).

Teorema 2.1 [1]- Seja $\tilde{\sigma} = \{\tilde{S}_k : k = 1, \dots, \tilde{K}\}$ qualquer partição de $\{1, \dots, n\}$, não necessariamente a mesma partição σ utilizada na construção de (2.6), e suponha que são conhecidos números positivos $\{d_1, \dots, d_n\}$ e números não-negativos $\{p_1, \dots, p_{\tilde{K}}\}$ tais que:

$$\sum_{j \in \tilde{S}_k} d_j x_j^* \leq p_k, \quad k = 1, \dots, \tilde{K}. \quad (2.11)$$

Então:

$$z^* \leq u^t b + \sum_{k=1}^{\tilde{K}} \max_{j \in \tilde{S}_k} \left[\frac{c_j - u^t A^j}{d_j} \right]^+ p_k \quad (2.12)$$

onde u é qualquer vetor não-negativo do $\mathcal{R}^m$, A^j é a j -ésima coluna da matriz de restrições do problema original (2.1), e $[\alpha]^+ = \max\{0, \alpha\}$.

Demonstração - Considere o conjunto

$$W = \left\{ x \in \mathcal{R}^n / \sum_{j \in \tilde{S}_k} d_j x_j \leq p_k, \quad k = 1, \dots, \tilde{K}, \quad x \geq 0 \right\},$$

onde os números $d_j, j = 1, \dots, n$, $p_k, k = 1, \dots, \tilde{K}$, e a partição $\tilde{\sigma}$ possuem as características definidas pela hipótese da proposição.

Note que $W \neq \emptyset$, pois, devido à expressão (2.11), $x^* \in W$.



Além disso, para todo vetor $u \geq 0$, $u \in \mathcal{R}^m$, e para todo x factível para (2.1), tem-se:

$$\begin{aligned}
 z^* = c^t x^* &\leq c^t x^* + u^t(b - Ax^*), \text{ devido a não-negatividade de } u \text{ e à factibilidade de } x^* \\
 &= u^t b + (c^t - u^t A)x^* \\
 &\leq u^t b + \max_{x \in W} \{(c^t - u^t A)x\}, \text{ devido a } x^* \in W \\
 &= u^t b + \max_{x \in W} \left\{ \sum_{j=1}^n (c_j - u^t A^j) x_j \right\}
 \end{aligned} \tag{2.13}$$

onde A^j é a j -ésima coluna da matriz de restrições original.

Considere o problema de maximização em (2.13), e seja $\tilde{x}$ a sua solução ótima. Devido a não-negatividade do vetor $x \in W$, se o coeficiente $(c_j - u^t A^j)$ da variável x_j em (2.13) for negativo, então podemos dizer que $\tilde{x}_j = 0$. Dessa forma, o problema de otimização em (2.13) pode ser simplificado para

$$\max_{x \in W} \left\{ \sum_{j=1}^n [c_j - u^t A^j]^+ x_j \right\}, \text{ onde } [\alpha]^+ = \max\{0, \alpha\}. \tag{2.14}$$

Daí, levando-se em conta a partição $\tilde{\sigma}$, então, (2.13) pode ser escrito como:

$$z^* \leq u^t b + \sum_{k=1}^{\tilde{K}} \sum_{j \in \tilde{S}_k} [c_j - u^t A^j]^+ \tilde{x}_j \tag{2.15}$$

Considere, agora, os números $\{d_1, \dots, d_n\}$ com as características definidas pela hipótese da proposição. Como $d_j > 0$, $j = 1, \dots, n$, então esses números podem ser introduzidos na expressão (2.15), sem influência no seu resultado, da seguinte forma:

$$\begin{aligned}
 z^* &\leq u^t b + \sum_{k=1}^{\tilde{K}} \sum_{j \in \tilde{S}_k} \left[\frac{c_j - u^t A^j}{d_j} \right]^+ d_j \tilde{x}_j \\
 &\leq u^t b + \sum_{k=1}^{\tilde{K}} \max_{j \in \tilde{S}_k} \left[\frac{c_j - u^t A^j}{d_j} \right]^+ \sum_{j \in \tilde{S}_k} d_j \tilde{x}_j
 \end{aligned} \tag{2.16}$$

donde, levando-se em conta que $\tilde{x} \in W$, ou, equivalentemente, $\sum_{j \in \tilde{S}_k} d_j \tilde{x}_j \leq p_k$, $k = 1, \dots, \tilde{K}$, obtém-se:

$$z^* \leq u^t b + \sum_{k=1}^{\tilde{K}} \max_{j \in \tilde{S}_k} \left[\frac{c_j - u^t A^j}{d_j} \right]^+ p_k,$$

o que completa a prova da proposição. •

A suposição (2.11) requer algumas restrições em x^* sem resolver (2.1). Tais restrições podem aparecer explicitamente no problema (2.1) ou podem ser obtidas pela manipulação de suas restrições.

Um ponto chave para o cálculo do limite proposto em (2.12) é a escolha do vetor u .

Zipkin[1] tomou $u = U^*$, ou seja, tomou a solução ótima de (2.7), obtendo, assim:

$$z^* \leq U^{*t}b + \sum_{k=1}^{\tilde{K}} \max_{j \in \tilde{S}_k} \left[\frac{c_j - U^{*t}A^j}{d_j} \right]^+ p_k \quad (2.17)$$

Observe que, se U^* é factível para (2.2), então $c_j - U^{*t}A^j \leq 0$, $\forall j = 1, \dots, n$, já que U^* deve, necessariamente, verificar as restrições do problema dual (2.2). Daí, o problema de maximização em (2.14) tem função objetivo ótima nula, e, portanto, de (2.10) e (2.17), tem-se

$$U^{*t}b = \bar{c}^t X^* = Z^* \leq z^* \leq U^{*t}b,$$

donde segue que $z^* = Z^*$, o que significa que a solução desagregada obtida da solução ótima de (2.6) é também ótima para (2.1).

Mendelsohn (vide [4, 5] e suas referências) sugere a procura por soluções duais na direção do vetor U^* , ou seja, soluções duais do tipo $u = \theta U^*$, onde $\theta \in \mathcal{R}$ é o tamanho do passo na direção U^* , e deve ser escolhido de modo a minimizar o limite resultante.

Dessa abordagem, resulta:

$$z^* \leq \min_{\theta \geq 0, \theta \in \mathcal{R}} \left\{ \theta U^{*t}b + \sum_{k=1}^{\tilde{K}} \max_{j \in \tilde{S}_k} \left[\frac{c_j - \theta U^{*t}A^j}{d_j} \right]^+ p_k \right\} \quad (2.18)$$

A expressão sob minimização em (2.18) é convexa e linear por partes em θ , de modo que somente os pontos de não diferenciabilidade dessa expressão devem ser avaliados [4, 5], ou seja,

$$\frac{c_j - \theta U^{*t}A^j}{d_j} = 0, \quad j = 1, \dots, n, \quad \text{e} \quad (2.19)$$

$$\frac{c_j - \theta U^{*t}A^j}{d_j} = \frac{c_{j'} - \theta U^{*t}A^{j'}}{d_{j'}}, \quad j, j' \in \tilde{S}_k, k = 1, \dots, \tilde{K}. \quad (2.20)$$

Outra alternativa à avaliação das expressões (2.19) e (2.20) é a utilização de métodos de minimização para funções convexas, como por exemplo a *Busca de Fibonacci*[18]. Com isso, evita-se a avaliação dessas expressões, que pode ser trabalhosa, principalmente quando existem muitas variáveis em um mesmo grupo.



Taylor e Shetty (vide [4, 5] e suas referências) propuseram outra direção de busca por soluções duais não negativas, de modo a melhorar o limite (2.12). A partir do vetor U^* , escolheram o vetor b como direção de busca, e tomaram soluções duais da forma $u = U^* + \Phi b$, onde $\Phi \in \mathcal{R}$ é o tamanho do passo, e, como no caso anterior, deve ser escolhido de modo a minimizar o limite resultante.

Dessa forma, tem-se:

$$z^* \leq U^{*t}b + \min_{\substack{\Phi \in \mathcal{R} \\ U^* + \Phi b \geq 0}} \left\{ \Phi b^t b + \sum_{k=1}^{\tilde{K}} \max_{j \in \tilde{S}_k} \left[\frac{c_j - (U^* + \Phi b)^t A^j}{d_j} \right]^+ p_k \right\} \quad (2.21)$$

Devido à convexidade e linearidade por partes (em Φ) da expressão (2.21), a determinação da solução ótima do problema de minimização em (2.21) pode ser tratada similarmente ao problema (2.18).

Outras abordagens, tais como a utilização do vetor resultante da combinação linear convexa de duas ou mais soluções duais do problema (2.7) e a aplicação de subgradientes para minimizar (2.12), foram propostas em [5], inclusive com indicações numéricas sobre a qualidade dos limites resultantes.

Nossa abordagem, neste trabalho, consiste em investigar expressões alternativas para o limite (2.12) através da extensão da suposição (2.11) para outros tipos de restrições sobre x^* . Esse é o assunto do próximo capítulo.

Capítulo 3

Limites *a Posteriori* e Localizações Lineares

Um dos passos mais críticos no cálculo dos limites discutidos no capítulo anterior é a suposição da disponibilidade de restrições da forma:

$$\sum_{j \in \tilde{S}_k} d_j x_j^* \leq p_k, \quad k = 1, \dots, \tilde{K}, \quad (3.1)$$

para alguma partição $\tilde{\sigma} = \{\tilde{S}_k, k = 1, \dots, \tilde{K}\}$ do conjunto $\{1, \dots, n\}$ de índices das variáveis.

Restrições desse tipo definem um subconjunto do $\mathcal{R}^n$ que é genericamente chamado de *localização de x^** .

Mais precisamente, se é conhecido um conjunto $W \subseteq \mathcal{R}^n$, com a propriedade de que $x^* \in W$, então W é referido como *localização de x^** , ou, simplesmente, *localização*.

Neste capítulo, estenderemos o cálculo de limites superiores para localizações genéricas, baseados na teoria desenvolvida em [2] para agregação de variáveis em Programação Não-Linear.

A seção a seguir utiliza a teoria da dualidade em Programação Não-Linear [18, 17] na obtenção de limites superiores a partir de localizações quaisquer, enquanto a seção 3.2 faz um refinamento do limite obtido, através de alterações na solução dual utilizada no cálculo, e, por fim, a seção 3.3 considera a aplicação de localizações lineares particulares na construção de expressões analíticas para o limite superior.



3.1 Localizações Genéricas

Nesta seção, investigamos o cálculo de limites para o erro agregacional, em função de qualquer conjunto localização da solução ótima do problema inicial.

O teorema a seguir generaliza a abordagem de Zipkin, e fornece uma ferramenta para a computação de limites superiores a partir de localizações genéricas.

Teorema 3.1 [2] - *Seja $W \subseteq \mathcal{R}^n$, conjunto localização de x^* , tal que W seja fechado, convexo e limitado.*

Então para qualquer vetor $u \geq 0$, $u \in \mathcal{R}^m$, tem-se:

$$z^* \leq UB(u, W) \quad (3.2)$$

onde

$$UB(u, W) = \sum_{i=1}^m u_i b_i + \max_{x \in W} \left\{ \sum_{j=1}^n x_j \left(c_j - \sum_{i=1}^m u_i a_{ij} \right) \right\}. \quad (3.3)$$

Além disso,

$$z^* = \min_{u \geq 0} UB(u, W). \quad (3.4)$$

Demonstração - Pela definição do conjunto W , podemos acrescentar a condição $x \in W$ às restrições de (2.1), sem alterar a solução ótima do problema.

A função Lagrangeana [17] associada ao problema (2.1) é definida como

$$L(x, u) = c^t x - u^t (Ax - b), \text{ para algum } u \in \mathcal{R}^m, \quad u \geq 0.$$

Desde que o conjunto W é compacto, o teorema do ponto de sela [7, 17, 18] se verifica, de modo que:

$$z^* = \max_{x \in W} \min_{u \geq 0} L(x, u) \quad (3.5)$$

$$= \min_{u \geq 0} \max_{x \in W} L(x, u), \text{ devido à compacticidade de } W \quad (3.6)$$

$$\leq \max_{x \in W} L(x, u) \quad (3.7)$$

$$= \max_{x \in W} \{ (c^t - u^t A)x + u^t b \} \quad (3.8)$$

$$= \sum_{i=1}^m u_i b_i + \max_{x \in W} \left\{ \sum_{j=1}^n x_j \left(c_j - \sum_{i=1}^m u_i a_{ij} \right) \right\} = UB(u, W) \quad (3.9)$$

Além disso,

$$\begin{aligned} z^* &= \min_{u \geq 0} \max_{x \in W} L(x, u) \\ &= \min_{u \geq 0} UB(u, W), \end{aligned} \quad (3.10)$$

o que completa a prova do teorema. •

Corolário 3.1 [2] - Suponha que algumas restrições do problema original foram incluídas na definição de W , de modo que:

$$W = \left\{ x \in \mathcal{R}^n / \sum_{j=1}^n a_{ij}x_j - b_i \leq 0, i \in I \subseteq \{1, \dots, m\} = M, x \in \tilde{W} \right\} \quad (3.11)$$

onde $\tilde{W}$ é outra localização de x^* .

Então, para qualquer $u \geq 0$, temos:

$$\begin{aligned} UB(u, W) &= \max_{x \in W} \left\{ c^t x - \sum_{i \in M} u_i \left(\sum_{j=1}^n a_{ij}x_j - b_i \right) \right\} \\ &\geq \max_{x \in W} \left\{ c^t x - \sum_{i \in M \setminus I} u_i \left(\sum_{j=1}^n a_{ij}x_j - b_i \right) \right\} \geq z^*. \end{aligned} \quad (3.12)$$

Demonstração - Observe que, se $M = \{1, \dots, m\}$ e $I \subseteq M$, então:

$$\begin{aligned} UB(u, W) &= \max_{x \in W} \left\{ c^t x - \sum_{i \in M} u_i \left(\sum_{j=1}^n a_{ij}x_j - b_i \right) \right\} \\ &= \max_{x \in W} \left\{ c^t x - \sum_{i \in M \setminus I} u_i \left(\sum_{j=1}^n a_{ij}x_j - b_i \right) - \sum_{i \in I} u_i \left(\sum_{j=1}^n a_{ij}x_j - b_i \right) \right\} \\ &\geq \max_{x \in W} \left\{ c^t x - \sum_{i \in M \setminus I} u_i \left(\sum_{j=1}^n a_{ij}x_j - b_i \right) \right\}, \end{aligned}$$

pois $\sum_{i \in I} u_i \left(\sum_{j=1}^n a_{ij}x_j - b_i \right) < 0$, $\forall x \in W$, e $\forall u \geq 0, u \in \mathcal{R}^m$, o que valida a primeira desigualdade de (3.12).

Considere, agora, a Lagrangeana definida como:

$$L(x, u) = c^t x - \sum_{i \in M \setminus I} u_i \left(\sum_{j=1}^n a_{ij}x_j - b_i \right), \quad (3.13)$$

onde incluímos somente aquelas restrições que não são usadas na definição de W .



Daí, similarmente à demonstração do teorema (3.1), podemos escrever:

$$\begin{aligned}
 z^* &= \max_{x \in W} \min_{u \geq 0} L(x, u) \\
 &= \min_{u \geq 0} \max_{x \in W} L(x, u), \text{ devido à compacticidade de } W \\
 &\leq \max_{x \in W} L(x, u) \\
 &= \max_{x \in W} \left\{ c^t x - \sum_{i \in M \setminus I} u_i \left(\sum_{j=1}^n a_{ij} x_j - b_i \right) \right\},
 \end{aligned}$$

o que verifica a segunda desigualdade de (3.12), completando, assim, a demonstração do corolário. •

O corolário acima mostra que se a localização W é definida por alguma restrição original, então não precisamos incluir esta restrição na função objetivo do problema de maximização em (3.3) para computar o limite superior.

Daqui para frente, a menos que mencionado o contrário, suporemos que a condição $x \geq 0$ sempre faça parte do conjunto W .

Suponha, agora, que o limite $UB(u, W)$ tenha sido calculado para algum $u \geq 0$, digamos $u = U^*$. Então, é natural procurarmos por outro valor de u que forneça um melhor valor para $UB(u, W)$. Esse é o assunto da próxima seção.

3.2 Refinamento do Limite Superior

Consideramos, nesta seção, um processo de redução do valor do limite (3.3), através de alterações nas soluções duais utilizadas no seu cálculo.

O processo de refinamento de $UB(u, W)$ baseia-se na obtenção de u que resolva (aproximadamente) o problema $\min_{u \geq 0} UB(u, W)$.

Descrevemos, a seguir, um processo iterativo para tratar esse problema, proposto em [2] para o caso de agregação de variáveis em Programação Não-Linear, e aqui adaptado para tratar o caso linear.

Denote por $\Theta(u)$ a função objetivo do problema de maximização em (3.3), ou seja,

$$\Theta(u) = \max_{x \in W} \left\{ \sum_{j=1}^n x_j \left(c_j - \sum_{i=1}^m u_i a_{ij} \right) \right\}. \quad (3.14)$$

Suponha que (3.14) tenha uma solução única $\tilde{x}$ para $u = U^*$.

Então, pelo teorema do valor marginal (vide [2] e suas referências), $\Theta(u)$ é diferenciável em $u = U^*$, de modo que:

$$\nabla_{u_i} \Theta(U^*) = - \sum_{j=1}^n a_{ij} \tilde{x}_j. \quad (3.15)$$

onde ∇_{u_i} indica a i -ésima componente do gradiente de $\Theta(u)$.

Dessa forma,

$$\nabla_{u_i} UB(U^*, W) = b_i - \sum_{j=1}^n a_{ij} \tilde{x}_j, \quad i = 1, \dots, m. \quad (3.16)$$

Considere, agora, uma iteração do método do gradiente projetado [7, 17, 18] aplicado ao problema

$$\min_{u \geq 0} UB(u, W),$$

começando de $u = U^*$. Isto dá:

$$u(\rho) = \Pi_+(U^* - \rho \nabla_u UB(U^*, W)), \quad (3.17)$$

onde $\rho \geq 0$ é o tamanho do passo e $\Pi_+(\gamma) = \{[\gamma_i]^+, i = 1, \dots, m\}$, sendo, por definição, $[\gamma_i]^+ = \max\{0, \gamma_i\}$, o que garante que $u(\rho) \geq 0$.

O tamanho do passo $\tilde{\rho}$, associado com o método do gradiente projetado, pode ser definido por uma pesquisa unidimensional:

$$\tilde{\rho} = \arg \min \{UB(u(\rho), W) / \rho \geq 0\}. \quad (3.18)$$

Note que a função de ρ $UB(u(\rho), W)$ é convexa e linear por partes $\forall \rho$ tal que $u(\rho) \geq 0$, e, portanto, técnicas padrões, como a *Busca de Fibonacci* [18, 17], podem ser aplicadas.

A escolha $\tilde{u} = u(\tilde{\rho})$ garante um decréscimo estrito do limite superior, isto é, ou $UB(\tilde{u}, W) < UB(\bar{u}, W)$ ou $UB(\bar{u}, W) = \min_{u \geq 0} UB(u, W)$.

Essa abordagem proposta para decrescer o limite superior pode ser realizada iterativamente. Abaixo, nós especificamos os passos básicos do algoritmo.

Passo 0 (Inicialização) - Faça $k:=0$, e determine um valor inicial para u , por exemplo, $u^0 = U^*$.

Passo 1 - Para $u = u^k$, resolva (3.14) e encontre sua solução ótima $\tilde{x}^k$ (é suposto que $\tilde{x}^k$ é único). Compute o limite superior $UB(u^k, W)$ por (3.3):

$$UB(u^k, W) = \sum_{i=1}^m u_i^k b_i + \sum_{j=1}^n \tilde{x}_j^k \left(c_j - \sum_{i=1}^m u_i^k a_{ij} \right).$$



Passo 2 - Para $\tilde{x} = \tilde{x}^k$, compute os componentes do gradiente em (3.16):

$$\nabla_{u_i} UB(u^k, W) = b_i - \sum_{j=1}^n a_{ij} \tilde{x}_j^k, \quad i = 1, \dots, m.$$

Passo 3 - Para $u^k(\rho) = \Pi_+(u^k - \rho \nabla_u UB(u^k, W))$, compute

$$\rho_k = \arg \min \{UB(u^k(\rho), W) / \rho \geq 0\}.$$

Passo 4 - Caso o critério de parada tenha sido atingido, pare. Senão, faça $u^{k+1} := u^k(\rho_k)$ e $k := k + 1$ e volte ao Passo 1.

Para terminar o algoritmo, podemos fixar um número máximo de iterações ou utilizar outros critérios de parada, como, por exemplo, parar de iterar quando:

$$\frac{UB(u^k, W) - UB(u^{k+1}, W)}{UB(u^k, W)} \leq \gamma$$

onde $\gamma > 0$ é uma precisão prefixada.

Se a solução ótima $\tilde{x}$ de (3.14) é não única, então, pelo teorema do valor marginal (vide [2] e suas referências), $\Theta(u)$ é direcionalmente diferenciável, e técnicas de subgradiente podem ser aplicadas para decrescer o limite superior, similarmente ao método proposto em [5].

Como vimos na seção final do capítulo anterior, a idéia de melhorar limites superiores através de mudanças no vetor dual tem sido usada por muitos autores. Mendelsohn utilizou soluções duais da forma $u(\rho) = \rho U^*$, enquanto Shetty e Taylor tomaram $u(\rho) = \Pi_+(U^* + \rho b)$, onde b é o lado direito das restrições lineares originais (vide maiores detalhes em [4, 5] e suas referências).

Essas abordagens, entretanto, não garantem decréscimo estrito do limite superior, como acontece com a abordagem iterativa que acabamos de expor.

3.3 Localizações Lineares

A computação do limite superior para z^* proposto na seção anterior requer o conhecimento de um conjunto localização W .

De (3.3), segue que, quanto menor o conjunto W , melhor a estimativa para o erro cometido em virtude da agregação dos dados iniciais.



Sendo assim, definindo W pelas restrições originais, devemos utilizar tantas restrições quanto possível, de modo a obter uma melhor precisão para o limite resultante.

Por outro lado, a localização W deve ser 'simples' o suficiente para garantir que o esforço computacional associado com o cálculo do limite superior não seja caro demais.

Abaixo, consideramos algumas escolhas de W e as resultantes expressões analíticas dos limites superiores, ainda baseados no desenvolvimento feito em [2] para Programação Não-Linear.

Vimos que, para computar o limite superior UB , deduzido na seção 3.1, devemos maximizar uma função linear sob a condição $x \in W$.

Para simplificar os cálculos posteriores, denote por δ_j os coeficientes do problema de maximização em (3.3), ou seja, $\delta_j = c_j - \sum_{i=1}^m u_i a_{ij}$, $j = 1, \dots, n$. Os valores δ_j dependem, portanto, do vetor u .

Suponha que W é definido como:

$$W = \left\{ x \in \mathcal{R}^n / \sum_{j=1}^n d_j x_j \leq p, \quad p, d_j > 0, \quad j = 1, \dots, n, \quad x \geq 0 \right\} \quad (3.19)$$

Restrições lineares desse tipo podem ser tomadas diretamente das restrições do problema original e/ou por manipulação dessas restrições.

Note que W assim definido é convexo, fechado e limitado para quaisquer valores de d_j e p , e, portanto, o teorema (3.1) se aplica.

Para computar o limite superior (3.3) usando a localização (3.19), devemos resolver o problema:

$$\begin{aligned} \max \quad & \sum_{j=1}^n \delta_j x_j \\ \text{s.a.} \quad & \sum_{j=1}^n d_j x_j \leq p \\ & x_j \geq 0, j = 1, \dots, n, \end{aligned} \quad (3.20)$$

que se trata do como o *problema da mochila contínuo* [22].

Seja $\tilde{x}$ a solução ótima de (3.20).

Observe que, se $\delta_j \leq 0$ para algum $j \in \{1, \dots, n\}$, então, a correspondente componente de $\tilde{x}$ é tal que $\tilde{x}_j = 0$, devido a não negatividade de x .

Considere $\delta_j > 0$, para $j \in I \subseteq \{1, \dots, n\}$ e sejam $f(x) = \sum_{j=1}^n \delta_j x_j$ e $x_{n+1} \geq 0$, de modo que $\sum_{j=1}^n d_j x_j + x_{n+1} = p$. Como, por hipótese, $d_j > 0$, para $j = 1, \dots, n$, então, podemos escrever:

$$x_k = (p - \sum_{\substack{j=1 \\ j \neq k}}^n d_j x_j) / d_k - \frac{x_{n+1}}{d_k}, \text{ para algum } k \in I \subseteq \{1, \dots, n\}.$$

Com isso, $f(x)$ pode ser escrita como:

$$\begin{aligned} f(x) &= \sum_{\substack{j=1 \\ j \neq k}}^n \delta_j x_j + p \frac{\delta_k}{d_k} - \sum_{\substack{j=1 \\ j \neq k}}^n \left(\frac{\delta_k}{d_k} d_j \right) x_j - \left(\frac{\delta_k}{d_k} \right) x_{n+1} \\ &= p \left(\frac{\delta_k}{d_k} \right) + \sum_{\substack{j=1 \\ j \neq k}}^n \left\{ \frac{\delta_j d_k - \delta_k d_j}{d_k} \right\} x_j - \frac{\delta_k}{d_k} x_{n+1} \end{aligned} \quad (3.21)$$

Suponha que $\frac{\delta_k}{d_k} = \max_{j=1, \dots, n} \left\{ \frac{\delta_j}{d_j} \right\}$. Daí, $\frac{\delta_j d_k - \delta_k d_j}{d_k} \leq 0$, $j = 1, \dots, n$, e, devido a não negatividade de x e de (3.21), podemos escrever:

$$f(x) \leq p \frac{\delta_k}{d_k}, \forall x \in W, \quad (3.22)$$

donde se observa que o maior valor admissível para $f(x)$ é $p \frac{\delta_k}{d_k}$.

Por outro lado, se tomarmos x tal que $x_k = \frac{p}{d_k}$ e $x_j = 0$, $j = 1, \dots, k-1, k+1, \dots, n+1$, então, por (3.21), tem-se $f(x) = p \frac{\delta_k}{d_k}$, de modo que, por (3.22), segue-se que essa solução é ótima para (3.20).

Portanto, tem-se:

$$f^*(\tilde{x}) = \sum_{j=1}^n d_j \tilde{x}_j = p \frac{\delta_k}{d_k}, \text{ onde } \frac{\delta_k}{d_k} = \max_{j=1, \dots, n} \left\{ \frac{\delta_j}{d_j} \right\}, \quad (3.23)$$

de forma que (3.3) é facilmente computado utilizando-se esse resultado.

Suponha, agora, que são conhecidos números $\{r_1, \dots, r_n\}$, tal que $0 \leq x_j \leq r_j$, $j = 1, \dots, n$, e defina W como:

$$W = \left\{ x \in \mathcal{R}^n / \sum_{j=1}^n d_j x_j \leq p, p, d_j > 0, 0 \leq x_j \leq r_j, r_j > 0, j = 1, \dots, n \right\} \quad (3.24)$$

Os limites superiores r_j para as variáveis x_j em (3.24) podem ser tomados diretamente das restrições do problema original e/ou por manipulação dessas restrições.

Podemos fazer, por exemplo, $r_j = \min_{i=1, \dots, m} \{b_i / a_{ij}\}$, para $a_{ij}, b_i > 0$, $j = 1, \dots, n$.

O problema de maximização em (3.3) é agora escrito como:

$$\begin{aligned} \max \quad & \sum_{j=1}^n \delta_j x_j \\ \text{s.a.} \quad & \sum_{j=1}^n d_j x_j \leq p \\ & 0 \leq x_j \leq r_j, j = 1, \dots, n, \end{aligned} \quad (3.25)$$

que é o problema da mochila contínuo com limites para as variáveis.

Uma solução ótima $\tilde{x}$ de (3.25) é facilmente computada da solução do problema (3.20). A forma da solução reflete a bem-conhecida idéia da mochila: os objetos devem ser colocados na mochila começando do mais 'caro' até que o limite superior r_j seja atingido ou a mochila fique completamente preenchida.

Suponha que as variáveis estão ordenadas como:

$$\frac{\delta_1}{d_1} \geq \frac{\delta_2}{d_2} \geq \dots \geq \frac{\delta_n}{d_n} \quad (3.26)$$

Então, a solução ótima $\tilde{x}$ de (3.25) é tal que:

$$\delta_j \tilde{x}_j = [\delta_j]^+ \min \left\{ r_j, \frac{p - \sum_{l=1}^{j-1} d_l \tilde{x}_l}{d_j} \right\}, \quad j = 1, \dots, n, \quad (3.27)$$

onde, por definição, $[\gamma]^+ = \max \{0, \gamma\}$, e os termos $[\delta_j]^+$ em (3.27) garantem que $\tilde{x}_j = 0$ para $\delta_j < 0$.

Se $\delta_j = 0$, então $\tilde{x}_j$ é não único, e satisfaz:

$$0 \leq \tilde{x}_j \leq \min \left\{ r_j, \frac{p - \sum_{l=1}^{j-1} d_l \tilde{x}_l}{d_j} \right\}, \quad (3.28)$$

mas isso não influencia o valor ótimo da função objetivo de (3.25) ou de (3.20).

Considere, agora, que, em vez de uma restrição linear em (3.19), nós temos um número qualquer de restrições lineares independentes.

Seja $\tilde{\sigma} = \{\tilde{S}_k, k = 1, \dots, \tilde{K}\}$ qualquer partição de $\{1, \dots, n\}$, não necessariamente a mesma partição σ , utilizada para construir o problema agregado, e suponha que W é definido como:

$$W = \left\{ x \in \mathcal{R}^n / \sum_{j \in \tilde{S}_k} d_j x_j \leq p_k, \quad k = 1, \dots, \tilde{K}, x_j \geq 0, \forall j \in \tilde{S}_k \right\} \quad (3.29)$$

Localizações do tipo (3.29) foram utilizadas por Zipkin [1] nos seus trabalhos pioneiros sobre agregação.

Utilizando-se a localização (3.29), então o problema (3.20) se decompõe em $\tilde{K}$ subproblemas independentes do tipo:

$$\begin{aligned} \max \quad & \sum_{j \in \tilde{S}_k} \delta_j x_j \\ \text{s.a.} \quad & \sum_{j \in \tilde{S}_k} d_j x_j \leq p_k \\ & x_j \geq 0, j \in \tilde{S}_k, \end{aligned} \quad (3.30)$$

e uma solução ótima para cada subproblema é computada por (3.23), para $k = 1, \dots, \tilde{K}$. O objetivo ótimo global é a soma dos objetivos ótimos dos subproblemas.

O mesmo ocorre se W for definido como:

$$W = \left\{ x \in \mathcal{R}^n / \sum_{j \in \tilde{S}_k} d_j x_j \leq p_k, \quad 0 \leq x_j \leq r_j, \quad j \in \tilde{S}_k, \quad k = 1, \dots, \tilde{K}, \right\}, \quad (3.31)$$

sendo que, utilizando-se W dessa forma, o problema de maximização em (3.3) se decompõe em $\tilde{K}$ subproblemas com o formato:

$$\begin{aligned} \max \quad & \sum_{j \in \tilde{S}_k} \delta_j x_j \\ \text{s.a.} \quad & \sum_{j \in \tilde{S}_k} d_j x_j \leq p_k \\ & 0 \leq x_j \leq r_j, \quad j \in \tilde{S}_k, \end{aligned} \quad (3.32)$$

cujas soluções são computadas analogamente a (3.27).

Observe que, se a localização é decomponível, e pode ser escrita como

$$W = W_1 \times W_2 \times \dots \times W_L, \text{ onde } W_l \text{ é convexo e limitado em } \mathcal{R}^{n_l}, \text{ e } \sum_{l=1}^L n_l = n,$$

então, o problema de maximização em (3.3) também se decompõe em L subproblemas independentes, sendo que o l -ésimo subproblema tem n_l variáveis.

Mesmo se não for possível resolver os subproblemas analiticamente, podemos tratar este caso como um caso fácil, uma vez que a dimensão desses subproblemas pode ser bem menor que a dimensão do problema original.

Vários tipos de localizações decomponíveis podem ser obtidas de problemas com estrutura especial, e formam parte importante das técnicas de agregação-decomposição iterativa (para informações, vide [3]).

Limites superiores mais precisos podem ser obtidos quando utilizamos um número maior de restrições na definição da localização, pois isso pode reduzir o tamanho de W .

No entanto, já não é possível garantir, dessa forma, a unicidade da solução do problema de otimização em (3.3), e nem obter expressões analíticas para o limite resultante da utilização desse tipo de localização.

Faz-se necessário, portanto, a definição de localizações mais suaves, que garantam a unicidade da solução do problema de maximização em (3.3), além de fornecer expressões de fácil computação para o limite (3.3).

Com esse intuito, no capítulo a seguir, investigamos as localizações não-lineares tipo elipsoidais, e abordamos a sua aplicação na análise do erro devido à agregação.



Capítulo 4

Localizações Elipsoidais

Este capítulo introduz as localizações não-lineares tipo elipsoidais, que serão, posteriormente, aplicadas à análise do erro devido à agregação de variáveis em Programação Linear.

Considerando-se localizações elipsoidais em (3.3), então, a solução ótima do problema de maximização em (3.3) é única e pode ser computada analiticamente (vide capítulo 5), donde segue a grande vantagem da utilização desse tipo de localização na computação de limites superiores para a função objetivo do problema original.

Chama-se elipsóide ou conjunto elipsoidal ao conjunto:

$$E = \{x \in \mathcal{R}^n / g + h^t x + x^t H x \leq \text{constante}\}, \quad (4.1)$$

onde $g \in \mathcal{R}$, $h, x \in \mathcal{R}^n$, e $H \in \mathcal{R}^{n \times n}$, H simétrica e definida positiva [20].

Os eixos dos elipsóides assim definidos têm a direção dos autovetores de H e comprimentos iguais ao inverso dos autovalores correspondentes [20].

Neste trabalho, serão considerados conjuntos elipsoidais do tipo:

$$E = \{x \in \mathcal{R}^n / (x - \tilde{x})^t H (x - \tilde{x}) \leq R^2\}, \quad (4.2)$$

onde H é matriz diagonal e definida positiva, $\tilde{x} \in \mathcal{R}^n$ é o centro e $R > 0$ é o raio do elipsóide.

Observe que os eixos do elipsóide (4.2) são paralelos aos eixos coordenados do sistema cartesiano, e têm comprimento igual ao inverso dos elementos diagonais de H .

Considere, agora, o problema (2.1), e suponha que a ele foram acrescentadas variáveis de



folga $r \in \mathcal{R}^m$, $r \geq 0$, de modo que o problema fica com o seguinte formato:

$$\begin{aligned} \max \quad & c^t x \\ \text{s.a.} \quad & Ax + r = b \\ & x, r \geq 0, \end{aligned} \quad (4.3)$$

que pode ser escrito, alternativamente, como:

$$\begin{aligned} \max \quad & \hat{c}^t \hat{x} \\ \text{s.a.} \quad & \hat{A} \hat{x} = b \\ & \hat{x} \geq 0, \end{aligned} \quad (4.4)$$

onde $\hat{A} \in \mathcal{R}^{m \times (m+n)}$, $\hat{A} = [A \ I]$, I é a matriz identidade de ordem m , $\hat{c} \in \mathcal{R}^{m+n}$, $\hat{c} = [c \ 0]$, 0 é o vetor nulo do $\mathcal{R}^m$, e $\hat{x} \in \mathcal{R}^{m+n}$, $\hat{x} = (x, r)$.

O dual de (4.4) é escrito como:

$$\begin{aligned} \min \quad & b^t \hat{u} \\ \text{s.a.} \quad & \hat{A}^t \hat{u} - \hat{s} = \hat{c} \\ & \hat{s} \geq 0, \hat{u} \text{ irrestrito,} \end{aligned} \quad (4.5)$$

onde $\hat{u} \in \mathcal{R}^m$, $\hat{s} \in \mathcal{R}^{m+n}$, e $\hat{A}$ e $\hat{c}$ definidos como em (4.4).

As soluções ótimas de (4.4) e (4.5), serão denotadas $\hat{x}^*$ e $(\hat{u}^*, \hat{s}^*)$, respectivamente. No que segue, será suposto que as regiões factíveis de (4.4) e (4.5) são limitadas e que seu interior é não vazio. Será suposto, ainda, que posto($\hat{A}$) = m .

Além disso, daqui por diante, $\hat{X}$ e $\hat{S}$ indicarão as matrizes diagonais cujos elementos diagonais são os elementos dos vetores $\hat{x}$ e $\hat{s}$, respectivamente, e $\|\cdot\|$ indicará a norma euclidiana.

A seção a seguir utiliza essa notação na construção de localizações elipsoidais de $\hat{x}^*$. Esses elipsóides, juntamente com elipsóides que contêm todas as soluções de folga duais ótimas de (4.5), têm sido aplicados na dedução de critérios para identificar variáveis primais básicas e não-básicas ótimas, o que é muito útil nas modernas técnicas de pontos interiores aplicadas à resolução de problemas de programação linear (vide abordagens de Todd[14] e Ye[15], além de Choi e Goldfarb[10]).

Neste texto, propomos a utilização desses elipsóides na análise do erro devido à agregação das variáveis do problema (2.1) (vide capítulo 5).



4.1 Elipsóides Primais

Apresentamos, nesta seção, os *elipsóides primais*, assim chamados os elipsóides que contêm todas as soluções ótimas $\hat{x}^*$ de (4.4).

O teorema a seguir caracteriza os elipsóides desenvolvidos em [10], aqui chamados *elipsóides tipo Choi-Goldfarb*.

Teorema 4.1 [10] - Sejam $\hat{x}$ e $(\hat{u}, \hat{s})$ soluções factíveis para os problemas (4.4) e (4.5), respectivamente, com $\hat{s} > 0$. Então, o elipsóide:

$$E = \{x \in \mathcal{R}^{m+n} / \|\hat{S}(x - \hat{x})\| \leq \varepsilon\}, \quad (4.6)$$

onde

$$\varepsilon^2 = (\hat{x}^t \hat{s})^2 + \|\hat{S}\hat{x}\|^2, \quad (4.7)$$

contém todas as soluções ótimas $\hat{x}^*$ do problema (4.4).

Demonstração - Observe que, para todo $\hat{x}^*$ ótimo de (4.4), tem-se:

$$\begin{aligned} \|\hat{S}(\hat{x}^* - \hat{x})\|^2 &= \|\hat{S}\hat{x}^*\|^2 + \|\hat{S}\hat{x}\|^2 - 2\hat{x}^t \hat{S}^2 \hat{x}^* \\ &\leq \|\hat{S}\hat{x}^*\|^2 + \|\hat{S}\hat{x}\|^2, \text{ devido à não-negatividade de } \hat{x}^*, \hat{x} \text{ e } \hat{s} \\ &\leq (\hat{s}^t \hat{x}^*)^2 + \|\hat{S}\hat{x}\|^2, \text{ também devido às restrições de não-negatividade} \\ &= [(\hat{A}^t \hat{u} - \hat{c})^t \hat{x}^*]^2 + \|\hat{S}\hat{x}\|^2, \text{ devido à factibilidade de } (\hat{u}, \hat{s}) \\ &= (\hat{u}^t \hat{A}\hat{x}^* - \hat{c}^t \hat{x}^*)^2 + \|\hat{S}\hat{x}\|^2 \\ &= (\hat{u}^t \hat{b} - \hat{c}^t \hat{x}^*)^2 + \|\hat{S}\hat{x}\|^2, \text{ devido à factibilidade de } \hat{x}^* \\ &\leq (\hat{u}^t \hat{b} - \hat{c}^t \hat{x})^2 + \|\hat{S}\hat{x}\|^2, \text{ pois } \hat{u}^t \hat{b} \geq \hat{c}^t \hat{x}^* \geq \hat{c}^t \hat{x}, \forall \hat{x} \text{ factível para (4.4)} \\ &= (\hat{u}^t \hat{A}\hat{x} - \hat{c}^t \hat{x})^2 + \|\hat{S}\hat{x}\|^2, \text{ pois } \hat{x} \text{ é factível para (4.4)} \\ &= [(\hat{A}^t \hat{u} - \hat{c})^t \hat{x}]^2 + \|\hat{S}\hat{x}\|^2 \\ &= (\hat{s}^t \hat{x})^2 + \|\hat{S}\hat{x}\|^2, \text{ pois } (\hat{u}, \hat{s}) \text{ é factível para (4.5)} \\ &= \varepsilon^2 \end{aligned}$$

o que completa a prova do teorema. •

Observação 4.1 -

1. Para a obtenção dos elipsóides descritos no teorema 4.1, são necessárias soluções $\hat{x}$ e $(\hat{u}, \hat{s})$ factíveis para os problemas (4.4) e (4.5), respectivamente.

A solução $\hat{x}$ pode ser tomada como $\hat{x} = (\bar{x}, \bar{r})$, onde $\bar{x}$ é a solução desagregada obtida da desagregação da solução ótima de (2.6), e $\bar{r} = b - A\bar{x}$.

Para obter a solução $(\hat{u}, \hat{s})$, basta aplicar a agregação de variáveis ao problema:

$$\begin{aligned} \min \quad & b^t \hat{u} \\ \text{s.a.} \quad & \hat{A}^t \hat{u} \geq \hat{c} \\ & \hat{u} \text{ irrestrito,} \end{aligned} \quad (4.8)$$

e tomar $(\hat{u}, \hat{s}) = (\bar{u}, \bar{s})$, onde $\bar{u}$ é a solução desagregada obtida da desagregação da solução ótima de (4.8), e $\bar{s} = \hat{A}^t \bar{u} - \hat{c}$.

Observe que, para que o elipsóide do teorema (4.1) fique bem definido, é necessário que $\hat{S}$ seja definida positiva, o que requer que $\hat{s}$ seja positivo.

Portanto, se acaso alguma componente de $\hat{s}$ se anular ao calcularmos $\bar{s} = \hat{A}^t \bar{u} - \hat{c}$, então uma estratégia para obter $\hat{s} > 0$ é alterar um pouco alguma componente de $\bar{u}$ e recalcular $\bar{s}$.

2. Note que, $\hat{s}^t \hat{x} \rightarrow 0$, quando $\hat{x} \rightarrow \hat{x}^*$ e $(\hat{u}, \hat{s}) \rightarrow (\hat{u}^*, \hat{s}^*)$. Dessa forma, os elipsóides do teorema (4.1) têm raio pequeno, se $\hat{x}$ e $(\hat{u}, \hat{s})$ estiverem próximos da solução ótima de (4.4) e (4.5), respectivamente.
3. Observe, ainda, que se é conhecido um limite inferior para o valor ótimo da função objetivo de (4.4), ou seja, se se conhece LB tal que $\hat{c}^t \hat{x}^* \geq LB$, então

$$0 \leq \hat{s}^t \hat{x}^* = (\hat{A}^t \hat{u} - \hat{c})^t \hat{x}^* = \hat{u}^t \hat{A} \hat{x}^* - \hat{c}^t \hat{x}^* \leq \hat{u}^t b - LB,$$

de modo que o raio dos elipsóides descritos no teorema (4.1) também podem ser calculados como

$$\epsilon^2 = (\hat{u}^t b - LB)^2 + \|\hat{S} \hat{x}\|^2 \quad (4.9)$$

É claro que, se o raio definido em (4.9) for menor que o raio proposto no teorema (4.1), então é preferível utilizar a forma (4.9), pois desejamos obter localizações tão pequenas e simples quanto possível para aplicação na análise do erro devido à agregação. Note, entretanto, que essa situação só ocorrerá quando LB for maior que Z^* .



Com o objetivo de obter localizações elipsoidais reduzidas, os teoremas a seguir propõem algumas técnicas para diminuir o raio dos elipsóides tipo Choi-Goldfarb, propostos no teorema (4.1).

Teorema 4.2 [12] - Seja $W \subseteq \mathcal{R}^{m+n}$ um conjunto localização de $\hat{x}^*$ e sejam $\hat{x}$ e $(\hat{u}, \hat{s}), \hat{s} > 0$, soluções factíveis para os problemas (4.4) e (4.5), respectivamente. Então, o elipsóide:

$$E = \{x \in \mathcal{R}^{m+n} / \|\hat{S}(x - \hat{x})\| \leq \varepsilon(W)\}, \text{ onde} \quad (4.10)$$

$$\varepsilon^2(W) = \min \{\tilde{\varepsilon}^2(W), \hat{\varepsilon}^2(W)\}, \quad (4.11)$$

$$\tilde{\varepsilon}^2(W) = \|\hat{S}\hat{x}\|^2 + \max_{x \in W} \{\|\hat{S}x\|^2 - 2(\hat{S}^2\hat{x})^t x\}, \text{ e} \quad (4.12)$$

$$\hat{\varepsilon}^2(W) = \|\hat{S}\hat{x}\|^2 + (\hat{s}^t\hat{x})^2 - 2 \min_{x \in W} \{(\hat{S}^2\hat{x})^t x\}, \quad (4.13)$$

contém todas as soluções ótimas $\hat{x}^*$ de (4.4).

Demonstração - Observe que, para todo $\hat{x}^*$ ótimo de (4.4), tem-se:

$$\|\hat{S}(\hat{x}^* - \hat{x})\|^2 = \|\hat{S}\hat{x}\|^2 + \|\hat{S}\hat{x}^*\|^2 - 2(\hat{S}^2\hat{x})^t \hat{x}^* \quad (4.14)$$

$$\leq \|\hat{S}\hat{x}\|^2 + \max_{x \in W} \{\|\hat{S}x\|^2 - 2(\hat{S}^2\hat{x})^t x\} = \tilde{\varepsilon}^2(W) \quad (4.15)$$

onde a desigualdade segue de $\hat{x}^* \in W$.

Além disso, como $\hat{s}^t\hat{x}^* \leq \hat{s}^t\hat{x}$, $\forall \hat{x}$ e $(\hat{u}, \hat{s})$ factíveis para (4.4) e (4.5), respectivamente (vide demonstração do teorema (4.1)), então, de (4.14), tem-se:

$$\begin{aligned} \|\hat{S}(\hat{x}^* - \hat{x})\|^2 &\leq \|\hat{S}^t\hat{x}\|^2 + (\hat{s}^t\hat{x}^*)^2 - 2(\hat{S}^2\hat{x})^t \hat{x}^* \\ &\leq \|\hat{S}^t\hat{x}\|^2 + (\hat{s}^t\hat{x})^2 - 2(\hat{S}^2\hat{x})^t \hat{x}^* \\ &\leq \|\hat{S}\hat{x}\|^2 + (\hat{s}^t\hat{x})^2 - 2 \min_{x \in W} \{(\hat{S}^2\hat{x})^t x\} = \hat{\varepsilon}^2(W), \end{aligned}$$

onde a última desigualdade segue de $\hat{x}^* \in W$.

Tomando-se $\varepsilon^2(W) = \min \{\tilde{\varepsilon}^2(W), \hat{\varepsilon}^2(W)\}$, obtemos:

$$\|\hat{S}(\hat{x}^* - \hat{x})\|^2 \leq \varepsilon^2(W), \quad (4.16)$$

o que completa a prova do teorema. •

As proposições a seguir obtêm expressões analíticas para os raios $\tilde{\varepsilon}(W)$ e $\hat{\varepsilon}(W)$ do teorema (4.2), com a utilização de algumas localizações específicas.

Proposição 4.1 [12] - Considere o elipsóide definido no teorema (4.2) e suponha que o conjunto W é definido como $W = \{x \in \mathcal{R}^{m+n} / \hat{s}^t x \leq \hat{s}^t \hat{x}, x \geq 0\}$, onde $\hat{x}$ e $(\hat{u}, \hat{s})$, $\hat{s} > 0$ são soluções factíveis para os problemas (4.4) e (4.5), respectivamente. Então,

$$\varepsilon^2(W) = \|\hat{S}\hat{x}\|^2 + (\hat{s}^t \hat{x})^2 - 2(\hat{s}^t \hat{x}) \min_{i=1, \dots, m+n} \hat{s}_i \hat{x}_i.$$

Demonstração - Note que, da demonstração do teorema (4.1), tem-se $\hat{s}^t \hat{x}^* \leq \hat{s}^t \hat{x}$, $\forall \hat{x}$ e $(\hat{u}, \hat{s})$ factíveis para os problemas (4.4) e (4.5), respectivamente. Conseqüentemente, podemos dizer que $\hat{x}^* \in W$, de modo que W define uma localização de $\hat{x}^*$.

Além disso, considerando-se que $\hat{s} > 0$, então, $\forall x = (x_j) \in W$, podemos escrever $0 \leq x_j \leq \frac{\hat{s}^t \hat{x}}{\hat{s}_j}$, $j = 1, \dots, m+n$, donde segue que W é limitado.

Para calcular $\tilde{\varepsilon}(W)$ em (4.12), devemos resolver o problema:

$$\max_{x \in W} \{ \|\hat{S}x\|^2 - 2(\hat{S}^2 \hat{x})^t x \} \quad (4.17)$$

A função objetivo do problema (4.17) é estritamente convexa, e, portanto, a solução ótima é alcançada em um vértice de W .

Cada vértice i de W têm todas as componentes nulas, exceto a i -ésima componente, que vale $\hat{s}^t \hat{x} / \hat{s}_i$. Substituindo-se x em (4.17) pelo i -ésimo vértice de W , então o problema (4.17) toma a forma:

$$\max_{x \in W} \{ (\hat{s}^t \hat{x})^2 - 2(\hat{s}_i \hat{x}_i)(\hat{s}^t \hat{x}) \} \quad (4.18)$$

Observe que o vértice i de W que maximiza (4.18) é aquele que minimiza o produto $\hat{s}_i \hat{x}_i$. Portanto, (4.18) é equivalente a:

$$(\hat{s}^t \hat{x})^2 - 2(\hat{s}^t \hat{x})^2 \min_{i=1, \dots, m+n} (\hat{s}_i \hat{x}_i) \quad (4.19)$$

Daí, por (4.12), tem-se:

$$\tilde{\varepsilon}^2(W) = \|\hat{S}\hat{x}\|^2 + (\hat{s}^t \hat{x})^2 - 2(\hat{s}^t \hat{x}) \min_{i=1, \dots, m+n} \hat{s}_i \hat{x}_i,$$

validando, assim, a primeira parte da tese da proposição.

Considere, agora, a obtenção do segundo raio. Para calcular $\hat{\varepsilon}(W)$, devemos resolver o problema:

$$\min_{x \in W} (\hat{S}^2 \hat{x})^t x \quad (4.20)$$

A solução do problema (4.20) é, obviamente, o ponto $x = 0$, devido a não-negatividade dos vetores $\hat{s}$ e $\hat{x}$. Dessa forma, por (4.13), tem-se:

$$\hat{\varepsilon}^2(W) = \|\hat{S}\hat{x}\|^2 + (\hat{s}^t \hat{x})^2, \quad (4.21)$$

logicamente maior que o anterior, completando, assim, a demonstração da proposição. •

Note que, se $\hat{x} > 0$ na proposição (4.1), então $\tilde{\varepsilon}(W)$ é menor do que o raio do elipsóide proposto no teorema (4.1), o que implica que obtivemos uma localização elipsoidal menor que a equivalente localização tipo Choi-Goldfarb.

Proposição 4.2 [12] - Seja $\sum_{j=1}^{m+n} d_j x_j = p$, $d_j, p > 0$ uma das restrições de (4.4), ou uma combinação linear delas, e considere o elipsóide definido no teorema (4.2). Defina a localização W como $W = \{x \in \mathcal{R}^{m+n} / \sum_{j=1}^{m+n} d_j x_j = p, x \geq 0\}$. Então:

$$\tilde{\varepsilon}^2(W) = \|\hat{S}\hat{x}\|^2 + p^2 \max_{i=1, \dots, m+n} \left\{ \left(\frac{\hat{s}_i}{d_i} \right)^2 \left[1 - 2 \frac{d_i \hat{x}_i}{p} \right] \right\},$$

$$\hat{\varepsilon}^2(W) = \|\hat{S}\hat{x}\|^2 + (\hat{s}^t \hat{x})^2 - 2p \min_{i=1, \dots, m+n} \frac{\hat{s}_j^2 \hat{x}_j}{d_j},$$

e

$$\varepsilon(W) = \min \{ \tilde{\varepsilon}(W), \hat{\varepsilon}(W) \}$$

Demonstração - Note que, pela maneira como W foi definido, tem-se $\hat{x}^* \in W$, de modo que W define uma localização de $\hat{x}^*$. Além disso, $\forall x = (x_j) \in W, 0 \leq x_j \leq p/d_j, j = 1, \dots, m+n$, donde segue que W é limitado.

Analogamente às considerações da proposição anterior, o valor ótimo da função objetivo de (4.17) é atingido em um vértice de W .

O i -ésimo vértice de W possui todas as componentes nulas, exceto a i -ésima componente, que vale p/d_i . Substituindo x em (4.17) pelo i -ésimo vértice de W , o problema (4.17) toma a forma:

$$\max_{x \in W} \left\{ \left(\frac{\hat{s}_i p}{d_i} \right)^2 - 2(\hat{s}_i^2 \hat{x}_i) \frac{p}{d_i} \right\} \quad (4.22)$$

que é equivalente a

$$p^2 \max_{i=1, \dots, m+n} \left\{ \left(\frac{\hat{s}_i}{d_i} \right)^2 \left[1 - 2 \frac{d_i \hat{x}_i}{p} \right] \right\}.$$

Daí, por (4.12), tem-se:

$$\tilde{\varepsilon}^2(W) = \|\hat{S}\hat{x}\|^2 + p^2 \max_{i=1, \dots, m+n} \left\{ \left(\frac{\hat{s}_i}{d_i} \right)^2 \left[1 - 2 \frac{d_i \hat{x}_i}{p} \right] \right\},$$

como afirma a primeira parte da tese da proposição.

Agora, para obter $\hat{\varepsilon}(W)$, devemos resolver o problema:

$$\min_{x \in W} (\hat{S}^2 \hat{x})^t x \quad (4.23)$$

Esse problema pode ser tratado como um problema de mochila contínuo, cuja resolução foi discutida na seção 3.3.

A solução ótima de (4.23) é um ponto que possui todas as componentes nulas, exceto a k -ésima componente, que vale p/d_k , onde k é tal que:

$$\frac{\hat{s}_k^2 \hat{x}_k}{d_k} = \min_{j=1, \dots, m+n} \left\{ \frac{\hat{s}_j^2 \hat{x}_j}{d_j} \right\} \quad (4.24)$$

Substituindo esse ponto na função objetivo de (4.23), obtemos:

$$\min (\hat{S}^2 \hat{x})^t x = \frac{\hat{s}_k^2 \hat{x}_k}{d_k}$$

o que, por (4.24), dá:

$$\min (\hat{S}^2 \hat{x})^t x = p \min_{j=1, \dots, m+n} \frac{\hat{s}_j^2 \hat{x}_j}{d_j}$$

donde segue, por (4.13), que:

$$\hat{\varepsilon}^2(W) = \|\hat{S}\hat{x}\|^2 + (\hat{s}\hat{x})^2 - 2p \min_{j=1, \dots, m+n} \frac{\hat{s}_j^2 \hat{x}_j}{d_j}$$

Tomando-se o mínimo entre os dois raios calculados, obtemos o resultado requerido. •

Observe que, se o conjunto localização W do teorema (4.2) for definido como

$$W = \left\{ x \in \mathcal{R}^{m+n} / \sum_{j=1}^{m+n} d_j x_j = p, \quad r_j \leq x_j \leq R_j, \quad r_j \geq 0, \quad R_j > 0, \quad j = 1, \dots, m+n \right\},$$

Então, para obter $\hat{\varepsilon}(W)$, devemos resolver o problema:

$$\max_{x \in W} \left\{ \|\hat{S}x\|^2 - 2(\hat{S}^2 \hat{x})^t x \right\} \quad (4.25)$$

Observe que podemos acrescentar a constante $\|\hat{S}\hat{x}\|^2$ à função objetivo de (4.25) sem alterar a sua solução ótima, ficando, assim, com o problema:

$$\max_{x \in W} \left\{ \|\hat{S}\hat{x}\|^2 + \|\hat{S}x\|^2 - 2(\hat{S}^2 \hat{x})^t x \right\} \quad (4.26)$$

que é equivalente a

$$\max_{x \in W} \|\hat{S}(x - \hat{x})\|^2 \quad (4.27)$$

Observe, ainda, que a solução de (4.27) é atingida em um vértice de W , devido à convexidade de W e devido ao fato de que a função objetivo é um elipsóide, cuja matriz hessiana é definida positiva.

Se $\gamma_j = x_j - r_j$, tem-se $x_j = \gamma_j + r_j, j = 1, \dots, m+n$, de modo que $\sum_{j=1}^{m+n} d_j x_j = p$ pode ser escrito como $\sum_{j=1}^{m+n} d_j (\gamma_j + r_j) = p$, ou, equivalentemente, $\sum_{j=1}^{m+n} d_j \gamma_j = \hat{p}$, onde $\hat{p} = p - \sum_{j=1}^{m+n} d_j r_j$. Além disso, para cada $j = 1, \dots, m+n$, tem-se $r_j \leq x_j \leq R_j$, o que implica que $0 \leq \gamma_j \leq \hat{R}_j$, onde $\hat{R}_j = R_j - r_j$.

Portanto, o problema (4.27) é equivalente a

$$\max_{x \in \hat{W}} \|\hat{S}(\gamma + r - \hat{x})\|^2 \quad (4.28)$$

onde $\gamma = (\gamma_j)$, $r = (r_j)$, e $\hat{W} = \{\gamma \in \mathcal{R}^{m+n} / \sum_{j=1}^{m+n} d_j \gamma_j = \hat{p}, 0 \leq \gamma_j \leq \hat{R}_j\}$.

Observe que o problema (4.28) pode ser escrito, alternativamente, como:

$$\max_{x \in \hat{W}} \sum_{j=1}^{m+n} \hat{s}_j^2 (\gamma_j + r_j - \hat{x}_j)^2 \quad (4.29)$$

Uma solução de (4.28) pode ser computada com o seguinte algoritmo:

Passo 0 - Seja $J = \{1, \dots, m+n\}$ e $\gamma^* = (0, \dots, 0)$.

Passo 1 - Faça $l_j = \min\{\hat{R}_j, \frac{\hat{p}}{d_j}\}, \forall j \in J$

Passo 2 - Obtenha k de modo que $s_k^2(l_k + r_k - \hat{x}_k)^2 = \max_{j \in J} \{s_j^2(l_j + r_j - \hat{x}_j)^2\}$

Faça $\gamma_k = l_k$ e redefina o vetor $\hat{p}$ e o conjunto J como $\hat{p} = \hat{p} - d_k l_k, J = J - \{k\}$

Passo 3 - Se $\hat{p} = 0$, pare, pois γ^* é ótimo. Senão, retorne ao Passo 1.

Considere, agora, o cálculo de $\hat{\varepsilon}(W)$.

Para obter $\hat{\varepsilon}(W)$, devemos resolver o problema:

$$\min_{x \in W} (\hat{S}^2 \hat{x})^t x \quad (4.30)$$

Utilizando a mesma mudança de variáveis $\gamma_j = x_j - r_j, j = 1, \dots, m+n$, o problema (4.30) é transformado em:

$$\min_{\gamma \in \hat{W}_2} \{(\hat{S}^2 \hat{x})^t \gamma + (\hat{S}^2 \hat{x})^t r\}, \quad (4.31)$$

onde $\gamma = (\gamma_j)$, $r = (r_j)$, e $\hat{W}_2 = \{\gamma \in \mathcal{R}^{m+n} / \sum_{j=1}^{m+n} d_j \gamma_j = \hat{p}, 0 \leq \gamma_j \leq \hat{R}_j\}$.

Observe que o problema (4.31) é o problema da mochila contínuo, com limites para as variáveis, cuja solução foi discutida na seção 3 do capítulo anterior.

Suponha que as variáveis de (4.31) estão ordenadas de acordo com:

$$\frac{\hat{s}_1^2 \hat{x}_1}{d_1} \leq \frac{\hat{s}_2^2 \hat{x}_2}{d_2} \leq \dots \leq \frac{\hat{s}_{m+n}^2 \hat{x}_{m+n}}{d_{m+n}}$$

Então:

$$\gamma_j^* = \min \left\{ \hat{R}_j, \left[\frac{\hat{p} - \sum_{k=1}^{j-1} d_k \gamma_k}{d_j} \right]^+ \right\}, \quad j = 1, \dots, m+n, \quad (4.32)$$

de modo que

$$\min_{\gamma \in \tilde{W}_2} \left\{ (\hat{S}^2 \hat{x})^t \gamma + (\hat{S}^2 \hat{x})^t r \right\} = \sum_{j=1}^{m+n} \hat{s}_j^2 \hat{x}_j (\gamma_j^* + r_j),$$

onde γ_j^* é calculado por (4.32).

Combinando esse resultado com (4.13), obtém-se:

$$\hat{\varepsilon}(W) = \|\hat{S} \hat{x}\|^2 + (\hat{s}^t \hat{x})^2 - 2 \sum_{j=1}^{m+n} \hat{s}_j \hat{x}_j (r_j + \gamma_j^*)$$

onde γ_j^* é dado por (4.32).

Com isso, podemos computar facilmente o raio do elipsóide proposto no teorema (4.2).

Observação 4.2

1. Note que a expressão para o raio $\varepsilon(W)$ em (4.12) é válida para quaisquer vetores $\hat{x}$, $\hat{s}$ do $\mathcal{R}^{m+n}$, com $\hat{s} > 0$, de modo que $\hat{x}$ e $(\hat{u}, \hat{s})$ não precisam, necessariamente, ser factíveis para os problemas (4.4) e (4.5), respectivamente.

Com base nessas considerações, podemos obter uma localização elipsoidal de x^* , ótima de (2.1) da seguinte forma:

Sejam $\tilde{x}$ e $\tilde{s}$ vetores do $\mathcal{R}^n$ com $\tilde{s} > 0$, e considere $W \subseteq \mathcal{R}^n$ tal que $x^* \in W$, para todo x^* ótimo de (2.1).

Então, podemos escrever:

$$\begin{aligned} \|\tilde{S}(x^* - \tilde{x})\|^2 &= \|\tilde{S}\tilde{x}\|^2 + \|\tilde{S}x^*\|^2 - 2\tilde{x}^t \tilde{S}^2 x^* \\ &\leq \|\tilde{S}\tilde{x}\|^2 + \max_{x \in W} \left\{ \|\tilde{S}x\|^2 - 2\tilde{x}^t \tilde{S}^2 x \right\} = \varepsilon^2(W), \end{aligned} \quad (4.33)$$

onde a desigualdade segue de $x^* \in W$.

Portanto, podemos dizer que o elipsóide:

$$E = \{x \in \mathcal{R}^n / \|\tilde{S}(x - \tilde{x})\|^2 \leq \varepsilon^2(W)\}, \quad (4.34)$$

onde $\varepsilon(W)$ é definido por (4.33), contém todas as soluções ótimas x^* de (2.1).

Observe que o elipsóide (4.34) tem dimensão menor que os elipsóides definidos nos teoremas (4.1) e (4.2), e pode ser obtido com menor esforço computacional, já que $\tilde{x}$ e $(\tilde{u}, \tilde{s})$ não precisam, necessariamente, ser factíveis para os problemas (2.1) e (2.2), respectivamente.

Note, entretanto, que $W \subseteq E$, pois $\|\tilde{S}(x - \tilde{x})\|^2 \leq \|\tilde{S}\tilde{x}\|^2 + \max_{x \in W} \{\|\tilde{S}x\| - 2\tilde{x}^t \tilde{S}^2 x\}$, $\forall x \in W$.

Por isso, a utilização do elipsóide assim definido na análise do erro agregacional pode incorrer em um limite mais impreciso quando comparado com o limite obtido com a utilização do conjunto W .

2. Uma expressão mais simples para o raio do elipsóide (4.34) pode ser obtida se considerarmos $\tilde{x} \geq 0$, de modo que:

$$\begin{aligned} \|\tilde{S}(x^* - \tilde{x})\|^2 &= \|\tilde{S}\tilde{x}\|^2 + \|\tilde{S}x^*\|^2 - 2\tilde{x}^t \tilde{S}^2 x^* \\ &\leq \|\tilde{S}\tilde{x}\|^2 + \|\tilde{S}x^*\|^2, \text{ devido a não-negatividade de } \tilde{x}, \tilde{s} \text{ e } x^* \\ &\leq \|\tilde{S}\tilde{x}\|^2 + (\tilde{s}^t x^*)^2, \text{ também devido às restrições de não-negatividade} \\ &\leq \|\tilde{S}\tilde{x}\|^2 + \left(\max_{x \in W} \tilde{s}^t x\right)^2, \text{ devido a } x^* \in W \end{aligned} \quad (4.35)$$

Podemos afirmar, portanto, que o elipsóide:

$$E = \{x \in \mathcal{R}^n / \|\tilde{S}(x - \tilde{x})\|^2 \leq \varepsilon^2(W)\}, \quad (4.36)$$

onde

$$\varepsilon^2(W) = \|\tilde{S}\tilde{x}\|^2 + \left(\max_{x \in W} \tilde{s}^t x\right)^2 \quad (4.37)$$

contém todas as soluções ótimas x^* de (2.1).

Observe que se W for definido por alguma restrição de (2.1) ou uma combinação dessas restrições, então, o problema de maximização em (4.37) se confunde com o problema da mochila contínuo, de fácil resolução analítica, conforme discutido na seção 3.3.

3. Podemos simplificar ainda mais o raio (4.37), se considerarmos $\tilde{x} = 0$ em (4.36), de modo que:

$$\varepsilon(W) = \max_{x \in W} \tilde{s}^t x.$$

Daí, podemos dizer que o elipsóide

$$E = \left\{ \|\tilde{S}x\| \leq \max_{x \in W} \tilde{s}^t x \right\} \quad (4.38)$$

define uma localização da solução ótima x^* de (2.1).

O teorema a seguir apresenta outra técnica para reduzir o raio dos elipsóides propostos no teorema (4.1).

Teorema 4.3 [13] - Sejam $\hat{x}$ e $(\hat{u}, \hat{s}), \hat{s} > 0$, soluções factíveis para os problemas (4.4) e (4.5), respectivamente. Seja $W \subseteq \mathcal{R}^{m+n}$ conjunto convexo e fechado, tal que $\hat{x}^* \in W, \forall \hat{x}^*$ ótimo de (4.4). Então o elipsóide:

$$E = \left\{ x \in \mathcal{R}^{m+n} / \|\hat{S}(x - x_o)\| \leq \varepsilon(W) \right\}, \quad (4.39)$$

onde x_o é a solução ótima única do problema:

$$\min_{x \in W} \|\hat{S}x\|^2 \quad (4.40)$$

e

$$\varepsilon^2(W) = (\hat{s}^t \hat{x})^2 - \|\hat{S}x_o\|^2, \quad (4.41)$$

contém todas as soluções ótimas $\hat{x}^*$ de (4.4).

Demonstração - Observe que, para qualquer $\gamma \in \mathcal{R}^{m+n}$, temos:

$$\begin{aligned} \|\hat{S}(\hat{x}^* - \gamma)\|^2 &= \|\hat{S}\hat{x}^*\|^2 + \|\hat{S}\gamma\|^2 - 2\gamma^t \hat{S}^2 \hat{x}^* \\ &\leq (\hat{s}^t \hat{x}^*)^2 + \|\hat{S}\gamma\|^2 - 2\gamma^t \hat{S}^2 \hat{x}^* \\ &\leq (\hat{s}^t \hat{x})^2 + \|\hat{S}\gamma\|^2 - 2\gamma^t \hat{S}^2 \hat{x}^*, \end{aligned} \quad (4.42)$$

onde a primeira desigualdade segue da não-negatividade de $\hat{s}$ e $\hat{x}^*$, e a última segue de $0 \leq \hat{s}^t \hat{x}^* \leq \hat{s}^t \hat{x}, \forall \hat{x}$ factível para (4.4), conforme demonstração do teorema (4.1).

Considere W com as características definidas pela hipótese do teorema. De $\hat{x}^* \in W$ e de (4.42), tem-se:

$$\|\hat{S}(\hat{x}^* - \gamma)\|^2 \leq (\hat{s}^t \hat{x})^2 + \max_{x \in W} \left\{ \|\hat{S}\gamma\|^2 - 2\gamma^t \hat{S}^2 x \right\} \quad (4.43)$$

Note que (4.43) se verifica $\forall \gamma \in \mathcal{R}^{m+n}$.

Suponha que γ é escolhido dentre as soluções do problema:

$$\min_{\gamma \in \mathcal{R}^{m+n}} \max_{x \in W} \{ \|\hat{S}\gamma\|^2 - 2\gamma^t \hat{S}^2 x \}, \quad (4.44)$$

de modo que (4.43) pode ser escrito como

$$\|\hat{S}(\hat{x}^* - \gamma)\|^2 \leq (\hat{s}^t \hat{x})^2 + \min_{\gamma \in \mathcal{R}^{m+n}} \max_{x \in W} \{ \|\hat{S}\gamma\|^2 - 2\gamma^t \hat{S}^2 x \} \quad (4.45)$$

A função $f(\gamma, x) = \|\hat{S}\gamma\|^2 - 2\gamma^t \hat{S}^2 x$ é convexa em γ e linear em x , e o conjunto W é fechado e convexo, por hipótese. Supondo-se que o conjunto das soluções ótimas do problema

$$\max_{x \in W} \varphi(x) \quad (4.46)$$

é limitado, onde

$$\varphi(x) = \min_{\gamma \in \mathcal{R}^n} f(\gamma, x), \quad (4.47)$$

então, pelo *teorema minimax generalizado* (vide [13] e suas referências), temos

$$\min_{\gamma \in \mathcal{R}^n} \max_{x \in W} f(\gamma, x) = \max_{x \in W} \min_{\gamma \in \mathcal{R}^n} f(\gamma, x), \quad (4.48)$$

As condições de otimalidade (condições KKT) para o problema

$$\min_{\gamma \in \mathcal{R}^{m+n}} \{ \|\hat{S}\gamma\|^2 - 2\gamma^t \hat{S}^2 x \} \quad (4.49)$$

resultam em $2\hat{S}^2\gamma - 2\hat{S}^2x = 0$, ou, equivalentemente, $\hat{s}_j^2(\gamma_j - x_j) = 0$, $j = 1, \dots, m+n$. Desde que, por hipótese, $\hat{s} > 0$, então, $\gamma_j = x_j$, $j = 1, \dots, m+n$, donde segue que $\gamma = x$, na solução ótima de (4.49).

Conseqüentemente, por (4.48), temos:

$$\max_{x \in W} \varphi(x) = - \min_{x \in W} \|\hat{S}x\|^2, \quad (4.50)$$

de modo que (4.45) pode ser escrita como:

$$\|\hat{S}(\hat{x}^* - \gamma)\|^2 \leq (\hat{s}^t \hat{x})^2 - \min_{x \in W} \|\hat{S}x\|^2 \quad (4.51)$$

Observe que $\gamma = x$ é a solução do problema de minimização em (4.51). Dessa forma, fazendo $\gamma = x_o$ em (4.51), tem-se:

$$\|\hat{S}(\hat{x}^* - x_o)\|^2 \leq (\hat{s}^t \hat{x})^2 - \|\hat{S}x_o\|^2,$$

como afirma a hipótese do teorema. •

Observe que os parâmetros raio e centro da localização elipsoidal proposta no teorema (4.3), diferentemente dos elipsóides propostos por Choi-Goldfarb, não são fixos, e dependem de algum conjunto localização da solução ótima $\hat{x}^*$.

Além disso, nota-se da expressão (4.41), que o raio do elipsóide (4.39) é com certeza menor que o raio do elipsóide proposto por Choi e Goldfarb (vide teorema (4.1)).

As proposições a seguir obtêm expressões analíticas para o centro e o raio dos elipsóides propostos no teorema (4.3), a partir de algumas localizações de $\hat{x}^*$.

Proposição 4.3 [13] - Considere o elipsóide proposto no teorema (4.3) e seja W definido como $W = \{x \in \mathcal{R}^{m+n} / \hat{A}x = b\}$. Então, o centro do respectivo elipsóide é:

$$x_{01} = \hat{S}^{-2} \hat{A}^t (\hat{A} \hat{S}^{-2} \hat{A}^t)^{-1} b \quad (4.52)$$

e

$$\varepsilon^2(W) = (\hat{s}^t \hat{x})^2 - b^t (\hat{A} \hat{S}^{-2} \hat{A}^t)^{-1} b. \quad (4.53)$$

Demonstração - Observe que, da maneira como W foi definido, tem-se $\hat{x}^* \in W$, de modo que W define uma localização de $\hat{x}^*$. Para obter x_{01} , devemos resolver o problema:

$$\min \{0.5 \|\hat{S}x\|^2, \text{ tal que } \hat{A}x = b\}, \quad (4.54)$$

onde o multiplicador 0.5 foi acrescentado para simplificar expressões posteriores.

As condições KKT aplicadas ao problema (4.54) dão:

$$\hat{S}^2 x_{01} - \hat{A}^t u = 0 \quad (4.55)$$

e

$$\hat{A}x_{01} = b, \quad (4.56)$$

para algum $u \geq 0$, $u \in \mathcal{R}^m$.

De (4.55), tem-se $x_{01} = \hat{S}^{-2} \hat{A}^t u$, o que, substituído em (4.56) dá $\hat{A} \hat{S}^{-2} \hat{A}^t u = b$, donde segue que $u = (\hat{A} \hat{S}^{-2} \hat{A}^t)^{-1} b$.

Dessa forma, tem-se $x_{01} = \hat{S}^{-2} \hat{A}^t (\hat{A} \hat{S}^{-2} \hat{A}^t)^{-1} b$, conforme requerido pela tese da proposição.

Além disso,

$$\begin{aligned} \|\hat{S}x_{01}\|^2 &= x_{01}^t \hat{S}^2 x_{01} \\ &= b^t (\hat{A} \hat{S}^{-2} \hat{A}^t)^{-1} \hat{A} \hat{S}^{-2} \hat{S}^2 \hat{S}^{-2} \hat{A}^t (\hat{A} \hat{S}^{-2} \hat{A}^t)^{-1} b \\ &= b^t (\hat{A} \hat{S}^{-2} \hat{A}^t)^{-1} \hat{A} \hat{S}^{-2} \hat{A}^t (\hat{A} \hat{S}^{-2} \hat{A}^t)^{-1} b \\ &= b^t (\hat{A} \hat{S}^{-2} \hat{A}^t)^{-1} b, \end{aligned}$$

o que finaliza a demonstração da proposição. •

Proposição 4.4 [13] - Considere o elipsóide proposto no teorema (4.3) e sejam os vetores $d_1 = (\hat{A}\hat{S}^{-2}\hat{A}^t)^{-1}b$ e $d_2 = -(\hat{A}\hat{S}^{-2}\hat{A}^t)^{-1}(\hat{A}\hat{S}^{-1})e$, para algum $(\hat{u}, \hat{s})$, $\hat{s} > 0$, factível para (4.5), e $e = (1, \dots, 1) \in \mathcal{R}^{m+n}$. Sejam, ainda, UB e LB tais que $LB \leq \hat{c}^t \hat{x}^* \leq UB$. Considere a localização W definida como

$$W = \{x \in \mathcal{R}^{m+n} / \hat{A}x = b, b^t \hat{u} - UB \leq \hat{s}^t x \leq b^t \hat{u} - LB\}$$

Considere, ainda, que v_p indique a projeção do vetor v sobre o núcleo da matriz $\hat{A}\hat{S}^{-1}$.

Então para o centro x_{02} do respectivo elipsóide, tem-se:

(a) Se $b^t \hat{u} - UB \leq -b^t d_2 \leq b^t \hat{u} - LB$, então $x_{02} = x_{01}$

(b) Se $-b^t d_2 < b^t \hat{u} - UB$, então:

$$x_{02} = \hat{S}^{-1} \left[\hat{S}^{-1} \hat{A}^t \left(d_1 + \frac{b^t \hat{u} - UB + b^t d_2}{\|e_p\|^2} d_2 \right) + \frac{b^t \hat{u} - UB + b^t d_2}{\|e_p\|^2} e \right] \quad (4.57)$$

e

$$\varepsilon^2(W) = (\hat{s}^t \hat{x})^2 - \|\hat{S}\hat{x}\|^2 + \|(\hat{S}\hat{x})_p\|^2 - \frac{(b^t \hat{u} - UB + b^t d_2)^2}{\|e_p\|^2} \quad (4.58)$$

(c) se $-b^t d_2 > b^t \hat{u} - LB$, então:

$$x_{02} = \hat{S}^{-1} \left[\hat{S}^{-1} \hat{A}^t \left(d_1 + \frac{b^t \hat{u} - LB + b^t d_2}{\|e_p\|^2} d_2 \right) + \frac{b^t \hat{u} - LB + b^t d_2}{\|e_p\|^2} e \right] \quad (4.59)$$

e

$$\varepsilon^2(W) = (\hat{s}^t \hat{x})^2 - \|\hat{S}\hat{x}\|^2 + \|(\hat{S}\hat{x})_p\|^2 - \frac{(b^t \hat{u} - LB + b^t d_2)^2}{\|e_p\|^2}. \quad (4.60)$$

Demonstração - Como $\hat{s}^t \hat{x}^* = (A^t \hat{u} - c)^t \hat{x}^* = b^t \hat{u} - \hat{c}^t \hat{x}^*$, então, pela definição de UB e LB , tem-se $b^t \hat{u} - UB \leq \hat{s}^t \hat{x}^* \leq b^t \hat{u} - LB$, o que implica que $\hat{x}^* \in W$. Portanto, W define uma localização de $\hat{x}^*$.

Para obter x_{02} de acordo com o teorema 4.3, devemos resolver o problema:

$$\min \{0.5 \|\hat{S}x\|^2, x \in W\}. \quad (4.61)$$

onde acrescentamos o multiplicador 0.5 para simplificar expressões posteriores.

Observe que, se x_{01} dado pela proposição (4.3) é factível para (4.61), ele também é ótimo de (4.61), uma vez que a localização utilizada na proposição (4.3) é uma relaxação da localização ora utilizada. Além disso, observando-se que $\hat{s}^t x_{01} = -b^t d_2$, então, se $b^t \hat{u} - UB \leq$

$-b^t d_2 \leq b^t \hat{u} - LB$, tem-se que x_{01} é factível para (4.61). Daí, podemos dizer que, se $b^t \hat{u} - UB \leq -b^t d_2 \leq b^t \hat{u} - LB$, então $x_{01} = x_{02}$, o que prova a parte (a) da proposição.

Considere, agora, as condições KKT para (4.61):

$$\hat{A}x = b \quad (4.62)$$

$$b^t \hat{u} - UB \leq \hat{s}^t x \leq b^t \hat{u} - LB \quad (4.63)$$

$$\hat{S}^2 x - \hat{A}^t \hat{u} - (\theta - \lambda) \hat{s} = 0 \quad (4.64)$$

$$\lambda(\hat{s}^t x + LB - b^t \hat{u}) = 0 \quad (4.65)$$

$$\theta(b^t \hat{u} - UB - \hat{s}^t x) = 0 \quad (4.66)$$

$$\lambda \geq 0, \theta \geq 0 \quad (4.67)$$

De (4.64), temos:

$$x = \hat{S}^{-2} [\hat{A}^t \hat{u} + (\theta - \lambda) \hat{s}]. \quad (4.68)$$

Substituindo (4.68) em (4.62), obtemos $\hat{u} = (\hat{A} \hat{S}^{-2} \hat{A}^t)^{-1} [b - (\theta - \lambda) \hat{A} \hat{S}^{-1} e]$, onde $e = (1, \dots, 1) \in \mathcal{R}^{m+n}$.

Utilizando essa expressão em (4.68), obtemos:

$$x = \hat{S}^{-2} \hat{A}^t (\hat{A} \hat{S}^{-2} \hat{A}^t)^{-1} b + (\theta - \lambda) \hat{S}^{-1} [I - \hat{S}^{-1} \hat{A}^t (\hat{A} \hat{S}^{-2} \hat{A}^t)^{-1} \hat{A} \hat{S}^{-1}] e. \quad (4.69)$$

Pela proposição 4.3, o primeiro termo aqui é x_{01} , enquanto o segundo termo é igual a $(\theta - \lambda) \hat{S}^{-1} e_p$, onde e_p indica a projeção do vetor e no espaço nulo da matriz $\hat{A} \hat{S}^{-1}$ (para maiores informações acerca da operação de projeção, vide [19]).

Portanto,

$$x = x_{01} + (\theta - \lambda) \hat{S}^{-1} e_p \quad (4.70)$$

e, daí,

$$\hat{s}^t x = \hat{s}^t x_{01} + (\theta - \lambda) \hat{s}^t \hat{S}^{-1} e_p = -b^t d_2 + (\theta - \lambda) e^t e_p = -b^t d_2 + (\theta - \lambda) \|e_p\|^2 \quad (4.71)$$

pois, pela definição de projeção, $(e - e_p)^t e_p = 0$, donde segue que $e^t e_p = \|e_p\|^2$.

Suponha $-b^t d_2 < b^t \hat{u} - UB$. Então, (4.71) pode ser escrita como $\hat{s}^t x < b^t \hat{u} - UB + (\theta - \lambda) \|e_p\|^2$. Mas, de $x \in W$, deve-se ter, também, $\hat{s}^t x \geq b^t \hat{u} - UB$, donde segue que $\theta > 0$.

Daí, $\hat{s}^t x = b^t \hat{u} - UB$ por (4.66) e $\lambda = 0$ por (4.65).

Combinando este resultado com (4.71), obtemos $\theta = (b^t \hat{u} - UB + b^t d_2) / \|e_p\|^2$, de modo que, por (4.70):

$$x_{02} = x_{01} + \frac{b^t \hat{u} - UB + b^t d_2}{\|e_p\|^2} \hat{S}^{-1} e_p \quad (4.72)$$

Como $e_p = [I - \hat{S}^{-1} \hat{A}^t (\hat{A} \hat{S}^{-2} \hat{A}^t)^{-1} \hat{A} \hat{S}^{-1}] e = e + \hat{S}^{-1} \hat{A}^t d_2$, tem-se, por (4.72):

$$\begin{aligned} x_{02} &= \hat{S}^{-2} \hat{A}^t (\hat{A} \hat{S}^{-2} \hat{A}^t)^{-1} b + \frac{b^t \hat{u} - UB + b^t d_2}{\|e_p\|^2} \hat{S}^{-1} e_p \\ &= \hat{S}^{-2} \hat{A}^t d_1 + \frac{b^t \hat{u} - UB + b^t d_2}{\|e_p\|^2} \hat{S}^{-1} (e + \hat{S}^{-1} \hat{A}^t d_2) \\ &= \hat{S}^{-1} \left[\hat{S}^{-1} \hat{A}^t \left(d_1 + \frac{b^t \hat{u} - UB + b^t d_2}{\|e_p\|^2} d_2 \right) + \frac{b^t \hat{u} - UB + b^t d_2}{\|e_p\|^2} e \right]. \end{aligned}$$

Observe que:

$$\begin{aligned} x_{01} &= \hat{S}^{-2} \hat{A}^t (\hat{A} \hat{S}^{-2} \hat{A}^t)^{-1} b \\ &= \hat{S}^{-1} \left[\hat{S}^{-1} \hat{A}^t (\hat{A} \hat{S}^{-2} \hat{A}^t)^{-1} \hat{A} \hat{S}^{-1} \right] \hat{S} \hat{x} \\ &= \hat{S}^{-1} \left[I - (I - \hat{S}^{-1} \hat{A}^t (\hat{A} \hat{S}^{-2} \hat{A}^t)^{-1} \hat{A} \hat{S}^{-1}) \right] \hat{S} \hat{x} \\ &= \hat{S}^{-1} \left[\hat{S} \hat{x} - (\hat{S} \hat{x})_p \right] \end{aligned}$$

Daí, por (4.72) e devido à ortogonalidade dos vetores e_p e $\hat{S} \hat{x} - (\hat{S} \hat{x})_p$, e utilizando, ainda, o teorema de Pitágoras, temos:

$$\begin{aligned} \|\hat{S} x_{02}\|^2 &= \|\hat{S} \hat{x} - (\hat{S} \hat{x})_p + \frac{b^t \hat{u} - UB + b^t d_2}{\|e_p\|^2} e_p\|^2 \\ &= (\hat{S} \hat{x} - (\hat{S} \hat{x})_p + \frac{b^t \hat{u} - UB + b^t d_2}{\|e_p\|^2} e_p)^t (\hat{S} \hat{x} - (\hat{S} \hat{x})_p + \frac{b^t \hat{u} - UB + b^t d_2}{\|e_p\|^2} e_p) \\ &= (\hat{S} \hat{x} - (\hat{S} \hat{x})_p)^t (\hat{S} \hat{x} - (\hat{S} \hat{x})_p) + \frac{(b^t \hat{u} - UB + b^t d_2)^2}{\|e_p\|^4} \|e_p\|^2 \\ &= \|\hat{S} \hat{x}\|^2 - \|(\hat{S} \hat{x})_p\|^2 + \frac{(b^t \hat{u} - UB + b^t d_2)^2}{\|e_p\|^2}, \end{aligned}$$

donde se obtém

$$\varepsilon^2(W) = (\hat{S} \hat{x})^2 - \|\hat{S} \hat{x}\|^2 + \|(\hat{S} \hat{x})_p\|^2 - \frac{(b^t \hat{u} - UB + b^t d_2)^2}{\|e_p\|^2},$$

o que prova a parte (b) da proposição.



Suponha, agora, $-b^t d_2 > b^t \hat{u} - LB$. Então, de $\hat{s}^t x \leq b^t \hat{u} - LB$ e (4.71), segue que $\lambda > 0$. Portanto, $\hat{s}^t x = b^t \hat{u} - LB$ por (4.65) e $\theta = 0$ por (4.66).

Daí, de (4.71), obtemos $\lambda = -(b^t \hat{u} - LB + b^t d_2) / \|e_p\|^2$, de modo que, por (4.70):

$$x_{02} = \hat{S}^{-1} \left[\hat{S}^{-1} A^t \left(d_1 + \frac{b^t \hat{u} - LB + b^t d_2}{\|e_p\|^2} d_2 \right) + \frac{b^t \hat{u} - LB + b^t d_2}{\|e_p\|^2} e \right]$$

e

$$\|\hat{S} x_{02}\|^2 = \|\hat{S} \hat{x}\|^2 - \|(\hat{S} \hat{x})_p\|^2 + \frac{(b^t \hat{u} - LB + b^t d_2)^2}{\|e_p\|^2},$$

donde se obtém

$$\varepsilon^2(W) = (\hat{s}^t \hat{x})^2 - \|\hat{S} \hat{x}\|^2 + \|(\hat{S} \hat{x})_p\|^2 - \frac{(b^t \hat{u} - LB + b^t d_2)^2}{\|e_p\|^2},$$

o que completa a demonstração da proposição. •

Observação 4.3

1. Seja $\hat{s} \in \mathcal{R}^{m+n}$, com $\hat{s} > 0$. Seja, ainda, W conjunto localização da solução ótima $\hat{x}^*$ de (4.4). Então, de

$$\hat{s}^t \hat{x}^* \leq \max_{x \in W} \hat{s}^t x$$

e da demonstração do teorema (4.3), podemos dizer que o elipsóide

$$E = \{x \in \mathcal{R}^{m+n} / \|\hat{S}(x - x_o)\|^2 \leq \varepsilon^2(W)\}, \quad (4.73)$$

onde x_o é a solução ótima única do problema $\min \{\|\hat{S}x\|^2, x \in W\}$, e

$$\varepsilon^2(W) = \left(\max_{x \in W} \hat{s}^t x \right)^2 - \|\hat{S} x_o\|^2, \quad (4.74)$$

contém todas as soluções ótimas $\hat{x}^*$ de (4.4).

2. A forma (4.74) para o raio do elipsóide é interessante porque não requer $(\hat{u}, \hat{s})$ factível para (4.5), de modo que o elipsóide (4.73) pode ser obtido com menor esforço computacional que o elipsóide proposto no teorema (4.3).
3. Note que se W em (4.73) for definido por uma restrição do problema (4.4) ou por uma combinação dessas restrições, então, o problema de maximização em (4.74) se confunde com o problema da mochila contínuo, de fácil resolução analítica, conforme discutido na seção 3.3.

A seção a seguir investiga as localizações das variáveis de folga $\hat{s}^*$ ótimas de (4.5).



4.2 Elipsóides Duais

Analogamente ao desenvolvimento feito na seção anterior, podemos definir os *elipsóides duais*, aqui chamados os elipsóides que contêm todas as variáveis de folga $\hat{s}^*$ ótimas de (4.5).

O teorema a seguir caracteriza os elipsóides duais tipo Choi-Goldfarb.

Teorema 4.4 [10] - Sejam $\hat{x} > 0$ e $(\hat{u}, \hat{s})$ soluções factíveis para os problemas (4.4) e (4.5), respectivamente. Então, o elipsóide :

$$E = \left\{ s \in \mathcal{R}^{m+n} / \|\hat{X}(s - \hat{s})\| \leq \varepsilon \right\}, \quad (4.75)$$

onde

$$\varepsilon^2 = (\hat{x}^t \hat{s})^2 + \|\hat{X} \hat{s}\|^2, \quad (4.76)$$

contém todas as soluções duais de folga $\hat{s}^*$ ótimas de (4.5).

Demonstração - Observe que, para todo $\hat{s}^*$ ótimo de (4.5), e $\forall \hat{x}$ factível para (4.4) e $(\hat{u}, \hat{s})$ factível para (4.5), tem-se:

$$\begin{aligned} \|\hat{X}(\hat{s}^* - \hat{s})\|^2 &= \|\hat{X} \hat{s}^*\|^2 + \|\hat{X} \hat{s}\|^2 - 2\hat{s}^t \hat{X}^2 \hat{s}^* \\ &\leq \|\hat{X} \hat{s}^*\|^2 + \|\hat{X} \hat{s}\|^2, \text{ devido a não-negatividade de } \hat{s}, \hat{x} \text{ e } \hat{s}^* \\ &\leq (\hat{x}^t \hat{s}^*)^2 + \|\hat{X} \hat{s}\|^2, \text{ também devido às restrições de não-negatividade} \\ &= [\hat{x}^t (\hat{A}^t \hat{u}^* - \hat{c})]^2 + \|\hat{X} \hat{s}\|^2, \text{ devido à factibilidade de } (\hat{u}^*, \hat{s}^*) \text{ para (4.5)} \\ &= [(\hat{A} \hat{x})^t \hat{u}^* - \hat{c}^t \hat{x}]^2 + \|\hat{X} \hat{s}\|^2 \\ &= (\hat{u}^{*t} \hat{b} - \hat{c}^t \hat{x})^2 + \|\hat{X} \hat{s}\|^2, \text{ pois } \hat{x}^* \text{ é factível para (4.4)} \\ &\leq (\hat{u}^t \hat{b} - \hat{c}^t \hat{x})^2 + \|\hat{X} \hat{s}\|^2, \text{ pois } 0 \leq \hat{u}^{*t} \hat{b} - \hat{c}^t \hat{x} \leq \hat{u}^t \hat{b} - \hat{c}^t \hat{x} \\ &= (\hat{u}^t \hat{A} \hat{x} - \hat{c}^t \hat{x})^2 + \|\hat{X} \hat{s}\|^2, \text{ devido à factibilidade de } \hat{x} \text{ para (4.4)} \\ &= [(\hat{A}^t \hat{u} - \hat{c})^t \hat{x}]^2 + \|\hat{X} \hat{s}\|^2 \\ &= (\hat{s}^t \hat{x})^2 + \|\hat{X} \hat{s}\|^2, \text{ devido à factibilidade de } (\hat{u}, \hat{s}) \text{ para (4.5)} \\ &= \varepsilon^2, \end{aligned}$$

o que finaliza a demonstração do teorema. •

Observação 4.4

1. As soluções $\hat{x}$ e $(\hat{u}, \hat{s})$ factíveis para os problemas (4.4) e (4.5), respectivamente, requeridas para a obtenção do elipsóide (4.75), podem ser obtidas similarmente às considerações do item 1 da observação (4.1).
2. Se é conhecido UB tal que $\hat{c}^t \hat{x}^* \leq UB$, então, podemos escrever $0 \leq \hat{x}^t \hat{s}^* \leq \hat{x}^t (\hat{A}^t \hat{u}^* - \hat{c}) = \hat{u}^{*t} b - \hat{c}^t \hat{x} \leq UB - \hat{c}^t \hat{x}$, e, portanto, o raio do elipsóide (4.75) pode ser calculado, alternativamente, como:

$$\varepsilon^2 = (UB - \hat{c}^t \hat{x})^2 + \|\hat{X} \hat{s}\|^2 \quad (4.77)$$

3. A mesma técnica de redução do raio dos elipsóides primais proposta no teorema (4.2) pode ser aplicada também aos elipsóides duais.

Observe, por exemplo, que $0 \leq \hat{x}^t \hat{s}^* \leq \hat{x}^t \hat{s}$ (vide demonstração do teorema (4.4)).

Com isso, podemos dizer que o conjunto $\{s \in \mathcal{R}^{m+n} / \hat{x}^t s \leq \hat{x}^t \hat{s}, s \geq 0\}$ define uma localização de $\hat{s}^*$. Daí, utilizando os mesmos argumentos do teorema (4.2) e da proposição (4.1), obtemos a localização elipsoidal de $\hat{s}^*$:

$$E = \left\{ \|\hat{X}(s - \hat{s})\|^2 \leq \|\hat{S} \hat{x}\|^2 + (\hat{s}^t \hat{x})^2 - 2(\hat{s}^t \hat{x}) \min_{i=1, \dots, m+n} \hat{s}_i \hat{x}_i \right\} \quad (4.78)$$

Observe que, se $\hat{s} > 0$ em (4.78), então o elipsóide (4.78) é menor que o elipsóide proposto por Choi e Goldfarb (compare com o teorema (4.4)).

4. Suponha que é conhecido $V \subseteq \mathcal{R}^n$ tal que $s^* \in V$, onde s^* é a variável de folga ótima de (2.2). Sejam, ainda, $\bar{x}$ e $\bar{s}$ vetores do $\mathcal{R}^n$, com $\bar{x} > 0$. Então, podemos dizer que $0 \leq \bar{x}^t s^* \leq \max \{\bar{x}^t s, s \in V\}$.

Com isso, podemos afirmar que o elipsóide:

$$E = \left\{ s \in \mathcal{R}^n / \|\bar{X}(s - \bar{s})\|^2 \leq \left(\max_{s \in V} \bar{x}^t s \right)^2 + \|\bar{X} \bar{s}\|^2 \right\} \quad (4.79)$$

contém todas as variáveis de folga s^* ótimas de (2.2).

Observe que o elipsóide (4.79) tem dimensão menor que o elipsóide definido pelo teorema (4.4), e pode ser obtido com menos esforço computacional, uma vez que não requer a factibilidade de $\bar{x}$ e $(\bar{u}, \bar{s})$ para os problemas (2.1) e (2.2), respectivamente.

O teorema a seguir, similarmente ao teorema (4.3), propõe uma técnica para reduzir o raio dos elipsóides propostos no teorema (4.4).

Teorema 4.5 [12] - Sejam $\hat{x} > 0$ e $(\hat{u}, \hat{s})$ soluções factíveis para os problemas (4.4) e (4.5), respectivamente. Considere, ainda, o conjunto $V \subseteq \mathcal{R}^{m+n}$, de modo que V é convexo, fechado e limitado, e $\hat{s}^* \in V$, $\forall \hat{s}^*$ ótimo de (4.5). Então, o elipsóide:

$$E = \{s \in \mathcal{R}^{m+n} / \|\hat{X}(s - s_o)\| \leq \varepsilon(V)\}, \quad (4.80)$$

onde s_o é a solução ótima única do problema:

$$\min_{s \in V} \|\hat{X}s\|^2 \quad (4.81)$$

e

$$\varepsilon^2(V) = (\hat{s}^t \hat{x})^2 - \|\hat{X}s_o\|^2, \quad (4.82)$$

contém todas as soluções de folga $\hat{s}^*$ ótimas de (4.5).

Demonstração - Observe que, para qualquer $\gamma \in \mathcal{R}^{m+n}$, temos:

$$\begin{aligned} \|\hat{X}(\hat{s}^* - \gamma)\|^2 &= \|\hat{X}\hat{s}^*\|^2 + \|\hat{X}\gamma\|^2 - 2\gamma^t \hat{X}^2 \hat{s}^* \\ &\leq (\hat{x}^t \hat{s}^*)^2 + \|\hat{X}\gamma\|^2 - 2\gamma^t \hat{X}^2 \hat{s}^*, \text{ devido a não-negatividade de } \hat{x}, \hat{s}^* \text{ e } \gamma \\ &\leq (\hat{x}^t \hat{s})^2 + \|\hat{X}\gamma\|^2 - 2\gamma^t \hat{X}^2 \hat{s}^*, \text{ devido a } 0 \leq \hat{x}^t \hat{s}^* \leq \hat{x}^t \hat{s}, \text{ vide teorema (4.4)} \\ &\leq (\hat{x}^t \hat{s})^2 + \max_{s \in V} \{\|\hat{X}\gamma\|^2 - 2\gamma^t \hat{X}^2 s\}, \text{ devido a } \hat{s}^* \in V \end{aligned} \quad (4.83)$$

O problema de maximização em (4.83) pode ser tratado similarmente ao problema de maximização em (4.43), donde resulta a expressão (4.82) para o raio do elipsóide (4.80) (vide demonstração do teorema (4.3)). •

Similarmente ao proposto para os elipsóides primais, as proposições a seguir utilizam algumas localizações de $\hat{s}^*$ na obtenção de expressões analíticas para o centro e o raio do elipsóide (4.80).

Proposição 4.5 [13] - Considere o elipsóide definido no teorema (4.5), e seja V definido como $V = \{s \in \mathcal{R}^{m+n} / \hat{A}^t \hat{u} - s = \hat{c}\}$, onde a matriz $\hat{A}$ e o vetor $\hat{c}$ são definidos como em (4.4). Então, o centro s_{01} do respectivo elipsóide é dado por:

$$s_{01} = -\hat{X}^{-1} \bar{c}_p,$$

onde $\bar{c}_p$ indica a projeção do vetor $\bar{c} = \hat{X}\hat{c}$ no espaço nulo da matriz $\hat{A}\hat{X}$, e

$$\varepsilon^2(V) = (\hat{x}^t \hat{s})^2 - \|\bar{c}_p\|^2.$$

Demonstração - Observe que, do modo como V foi definido, tem-se $\hat{s}^* \in V$, de modo que V define uma localização de $\hat{s}^*$.

Para obter s_{01} , devemos resolver o problema

$$\min \{0.5 \|\hat{X}s\|^2, \text{ tal que } \hat{A}^t u - s = \hat{c}\}, \quad (4.84)$$

onde o multiplicador 0.5 foi acrescentado para simplificar expressões posteriores.

A aplicação das condições KKT ao problema (4.84) resulta em:

$$\hat{X}^2 s - \beta = 0 \quad (4.85)$$

e

$$\hat{A}\beta = 0 \quad (4.86)$$

onde β é o multiplicador de Lagrange ótimo para (4.84).

De (4.85), segue que $s = \hat{X}^{-2}\beta$. Daí, de $\hat{A}^t u - \hat{s} = \hat{c}$, segue que:

$$\hat{A}^t u - \hat{X}^{-2}\beta = \hat{c} \quad (4.87)$$

Multiplicando (4.87) por $\hat{A}\hat{X}^2$, vem:

$$(\hat{A}\hat{X}^2\hat{A}^t)u - \hat{A}\beta = \hat{A}\hat{X}^2\hat{c} \quad (4.88)$$

o que, por (4.86), dá:

$$u = (\hat{A}\hat{X}^2\hat{A}^t)^{-1}\hat{A}\hat{X}^2\hat{c} \quad (4.89)$$

Substituindo (4.89) nas restrições de (4.84), obtemos:

$$\begin{aligned} s_{01} &= \hat{A}^t(\hat{A}\hat{X}^2\hat{A}^t)^{-1}\hat{A}\hat{X}^2\hat{c} - \hat{c} \\ &= -(I - \hat{A}^t(\hat{A}\hat{X}^2\hat{A}^t)^{-1}\hat{A}\hat{X}^2)\hat{c} \\ &= -\hat{X}^{-1}(I - \hat{X}\hat{A}^t(\hat{A}\hat{X}^2\hat{A}^t)^{-1}\hat{A}\hat{X})\hat{X}\hat{c} \\ &= -\hat{X}^{-1}\bar{c}_p. \end{aligned}$$

Além disso, $\|\hat{X}s_{01}\|^2 = \|\bar{c}_p\|^2$, de modo que, por (4.82), $\varepsilon^2(V) = (\hat{s}^t\hat{x})^2 - \|\bar{c}_p\|^2$, como afirma a tese da proposição. •

Proposição 4.6 [13] - Considere o elipsóide definido no teorema (4.5), e seja V definido como $V = \{s \in \mathcal{R}^{m+n} / \hat{A}^t\hat{u} - s = \hat{c}, LB - \hat{c}^t\hat{x} \leq \hat{x}^t s \leq UB - \hat{c}^t\hat{x}\}$. Considere, ainda, que v_p indique a projeção do vetor v no núcleo da matriz $\hat{A}\hat{X}$, $\bar{c} = \hat{X}\hat{c}$ e $e = (1, \dots, 1) \in \mathcal{R}^{m+n}$.

Então, para o centro s_{02} do respectivo elipsóide, temos:

(a) se $LB - \hat{c}^t \hat{x} \leq -e^t \bar{c}_p \leq UB - \hat{c}^t \hat{x}$, então $s_{02} = s_{01}$;

(b) se $-e^t \bar{c}_p > UB - \hat{c}^t \hat{x}$, então:

$$s_{02} = \hat{X}^{-1} \left[\left(\frac{\hat{c}^t \hat{x} - UB - e^t \bar{c}_p}{\|e - e_p\|^2} \right) (e - e_p) - \bar{c}_p \right].$$

e

$$\varepsilon^2(V) = (\hat{x}^t \hat{s})^2 - \left(\frac{\hat{c}^t \hat{x} - UB - e^t \bar{c}_p}{\|e - e_p\|} \right)^2 + \|\bar{c}_p\|^2,$$

(c) se $-e^t \bar{c}_p < LB - \hat{c}^t \hat{x}$, então:

$$s_{02} = \hat{X}^{-1} \left[\frac{LB - \hat{c}^t \hat{x} + e^t \bar{c}_p}{\|e - e_p\|^2} (e - e_p) - \bar{c}_p \right].$$

e

$$\varepsilon^2(V) = (\hat{x}^t \hat{s})^2 - \left(\frac{LB - \hat{c}^t \hat{x} + e^t \bar{c}_p}{\|e - e_p\|} \right)^2 + \|\bar{c}_p\|^2,$$

Demonstração - Observe que $\hat{x}^t \hat{s}^* = \hat{x}^t (\hat{A}^t \hat{u}^* - \hat{c}) = \hat{u}^{*t} b - \hat{c}^t \hat{x}$. Daí, considerando LB e UB com as características definidas pela hipótese da proposição, tem-se $LB - \hat{c}^t \hat{x} \leq \hat{x}^t \hat{s}^* \leq UB - \hat{c}^t \hat{x}$, o que significa que $\hat{s}^* \in V$. Portanto, V define uma localização de $\hat{s}^*$.

Para encontrar s_{02} pelo teorema (4.5), deve-se resolver o problema:

$$\min_{s \in V} \{0.5 \|\hat{X}s\|^2\}, \quad (4.90)$$

onde usamos o multiplicador 0.5 para simplificar expressões subseqüentes.

Desde que a localização ora utilizada é uma relaxação da localização utilizada pela proposição (4.5), se s_{01} é factível para o problema (4.90), então s_{01} é também a solução ótima de (4.90). Observando, ainda, que $\hat{x}^t s_{01} = -\hat{x}^t \hat{X}^{-1} \bar{c}_p = -e^t \bar{c}_p$, podemos dizer que, se $LB - \hat{c}^t \hat{x} \leq -e^t \bar{c}_p \leq UB - \hat{c}^t \hat{x}$, então s_{01} é factível para o problema (4.90), de modo que $s_{02} = s_{01}$, o que prova a parte (a) da proposição.

As condições KKT para (4.90) são:

$$\hat{A}^t \hat{u} - s = \hat{c} \quad (4.91)$$

$$LB - \hat{c}^t \hat{x} \leq \hat{x}^t s \leq UB - \hat{c}^t \hat{x} \quad (4.92)$$

$$\hat{X}^2 s - \beta - (\theta - \lambda) \hat{x} = 0 \quad (4.93)$$

$$\hat{A} \beta = 0 \quad (4.94)$$

$$\lambda(\hat{x}^t s - UB + \hat{c}^t \hat{x}) = 0 \quad (4.95)$$

$$\theta(LB - \hat{c}^t \hat{x} - \hat{x}^t s) = 0 \quad (4.96)$$

$$\lambda \geq 0, \theta \geq 0 \quad (4.97)$$

De (4.93), temos:

$$s = \hat{X}^{-2}[\beta + (\theta - \lambda)\hat{x}]. \quad (4.98)$$

Substituindo (4.98) em (4.91), obtemos:

$$\hat{u} = (\hat{A}\hat{X}^2\hat{A}^t)^{-1}(\hat{A}\hat{X}^2\hat{c} + (\theta - \lambda)b).$$

Usando esta expressão em (4.91), obtemos:

$$\begin{aligned} s &= \hat{A}^t(\hat{A}\hat{X}^2\hat{A}^t)^{-1}\hat{A}\hat{X}^2\hat{c} + (\theta - \lambda)\hat{A}^t(\hat{A}\hat{X}^2\hat{A}^t)^{-1}b - \hat{c} \\ &= (\theta - \lambda)\hat{A}^t(\hat{A}\hat{X}^2\hat{A}^t)^{-1}b - \hat{X}^{-1}[I - \hat{X}\hat{A}^t(\hat{A}\hat{X}^2\hat{A}^t)^{-1}\hat{A}\hat{X}]\hat{X}\hat{c} \\ &= (\theta - \lambda)\hat{A}^t(\hat{A}\hat{X}^2\hat{A}^t)^{-1}b - \hat{X}^{-1}\bar{c}_p, \end{aligned} \quad (4.99)$$

$$(4.100)$$

onde $\bar{c}_p$ representa a projeção do vetor $\bar{c} = \hat{X}\hat{c}$ no núcleo da matriz $\hat{A}\hat{X}$.

Pela definição de projeção sobre o espaço nulo, tem-se:

$$\begin{aligned} \hat{A}^t(\hat{A}\hat{X}^2\hat{A}^t)^{-1}b &= \hat{X}^{-1}\hat{X}\hat{A}^t(\hat{A}\hat{X}^2\hat{A}^t)^{-1}\hat{A}\hat{X}\hat{X}^{-1}\hat{x} \\ &= \hat{X}^{-1}[\hat{X}\hat{A}^t(\hat{A}\hat{X}^2\hat{A}^t)^{-1}\hat{A}\hat{X}]\hat{X}^{-1}\hat{x} \\ &= \hat{X}^{-1}[\hat{X}\hat{A}^t(\hat{A}\hat{X}^2\hat{A}^t)^{-1}\hat{A}\hat{X}]e \\ &= \hat{X}^{-1}[I - (I - \hat{X}\hat{A}^t(\hat{A}\hat{X}^2\hat{A}^t)^{-1}\hat{A}\hat{X})]e \\ &= \hat{X}^{-1}(e - e_p), \end{aligned}$$

de modo que, por (4.99), tem-se:

$$s = \hat{X}^{-1}[(\theta - \lambda)(e - e_p) - \bar{c}_p]. \quad (4.101)$$

Daí,

$$0 \leq \hat{x}^t s = (\theta - \lambda)\|e - e_p\|^2 - e^t \bar{c}_p, \quad (4.102)$$

pois $e^t(e - e_p) = \|e - e_p\|^2$, devido às propriedades de projeção.

Suponha que $-e^t \bar{c}_p > UB - \hat{c}^t \hat{x}$. Então, combinando a desigualdade $\hat{x}^t s \leq UB - \hat{c}^t \hat{x}$, da definição de V , e a expressão (4.102), obtemos

$$(\theta - \lambda)\|e - e_p\|^2 - e^t \bar{c}_p \leq UB - \hat{c}^t \hat{x} \quad (4.103)$$

Mas, por hipótese, $-e^t \bar{c}_p > UB - \hat{c}^t \hat{x}$, donde segue que $(\theta - \lambda)\|e - e_p\|^2$ deve ser negativo em (4.103), o que exige que $\lambda > 0$.

Com isso, $\hat{x}^t s = UB - \hat{c}^t \hat{x}$ por (4.95) e, supondo-se $UB \neq LB$, tem-se $\theta = 0$ por (4.96).

Substituindo este resultado em (4.102), obtemos $\lambda = (\hat{c}^t \hat{x} - UB - e^t \bar{c}_p) / \|e - e_p\|^2$, e, por (4.101), tem-se:

$$s_{02} = \hat{X}^{-1} \left[\left(\frac{UB - \hat{c}^t \hat{x} + e^t \bar{c}_p}{\|e - e_p\|^2} \right) (e - e_p) - \bar{c}_p \right].$$

Daí:

$$\begin{aligned} \|\hat{X} s_{02}\|^2 &= \left\| \frac{UB - \hat{c}^t \hat{x} + e^t \bar{c}_p}{\|e - e_p\|^2} (e - e_p) - \bar{c}_p \right\|^2 \\ &= \left(\frac{UB - \hat{c}^t \hat{x} + e^t \bar{c}_p}{\|e - e_p\|^2} (e - e_p) - \bar{c}_p \right)^t \left(\frac{UB - \hat{c}^t \hat{x} + e^t \bar{c}_p}{\|e - e_p\|^2} (e - e_p) - \bar{c}_p \right) \\ &= \frac{(UB - \hat{c}^t \hat{x} + e^t \bar{c}_p)^2}{\|e - e_p\|^4} \|e - e_p\|^2 - \|\bar{c}_p\|^2, \text{ devido à ortogonalidade de } e - e_p \text{ e } \bar{c}_p \\ &= \frac{(UB - \hat{c}^t \hat{x} + e^t \bar{c}_p)^2}{\|e - e_p\|^2} - \|\bar{c}_p\|^2 \end{aligned}$$

Portanto, por (4.82), tem-se:

$$\varepsilon^2(V) = (\hat{x}^t \hat{s})^2 - \left(\frac{\hat{c}^t \hat{x} - UB - e^t \bar{c}_p}{\|e - e_p\|} \right)^2 + \|\bar{c}_p\|^2,$$

o que prova a parte (b) da proposição.

Agora, seja $-e^t \bar{c}_p < LB - \hat{c}^t \hat{x}$. Então, devido à factibilidade de s , tem-se $\hat{x}^t s \geq LB - \hat{c}^t \hat{x}$.

Daí, por (4.102), segue que

$$(\theta - \lambda) \|e - e_p\|^2 - e^t \bar{c}_p \geq LB - \hat{c}^t \hat{x} \quad (4.104)$$

Mas, por hipótese, $-e^t \bar{c}_p < LB - \hat{c}^t \hat{x}$, o que implica que $(\theta - \lambda) \|e - e_p\|^2$ deve ser positivo em (4.104), o que exige que $\theta > 0$.

Daí, $\hat{x}^t s = LB - \hat{c}^t \hat{x}$ por (4.96), e, supondo-se $UB \neq LB$, tem-se $\lambda = 0$ por (4.95). De (4.102), temos $\theta = (LB - \hat{c}^t \hat{x} + e^t \bar{c}_p) / \|e - e_p\|^2$, o que, substituído em (4.101), dá:

$$s_{02} = \hat{X}^{-1} \left[\frac{LB - \hat{c}^t \hat{x} + e^t \bar{c}_p}{\|e - e_p\|^2} (e - e_p) - \bar{c}_p \right].$$

Daí, da mesma forma que no caso anterior, por (4.82), tem-se:

$$\varepsilon^2(V) = (\hat{x}^t \hat{s})^2 - \left(\frac{LB - \hat{c}^t \hat{x} + e^t \bar{c}_p}{\|e - e_p\|} \right)^2 + \|\bar{c}_p\|^2,$$

o que completa a demonstração da proposição. •

O capítulo a seguir aborda a aplicação das localizações elipsoidais que acabamos de analisar na metodologia de cálculo do erro agregacional.



Capítulo 5

Limites *a Posteriori* e Localizações Elipsoidais

Aplicamos, neste capítulo, as localizações elipsoidais abordadas no capítulo 4, à análise do erro devido à agregação de variáveis em Programação Linear.

Uma das formas de se utilizar os elipsóides na análise de erro é tomar W em (3.3) como um conjunto elipsoidal, com o intuito de se obter uma expressão analítica para o erro resultante dessa abordagem, o que será feito na seção 5.1.

Pode-se, ainda, utilizar os elipsóides na construção de localizações tipo caixa para aplicação no cálculo de limites superiores para z^* . Esse é o assunto da seção 5.2.

Por fim, a seção 5.3 apresenta um estudo numérico para avaliar a qualidade dos limites obtidos com a aplicação das localizações elipsoidais na análise do erro agregacional.

5.1 Análise de Erro sob Localização Elipsoidal

Considere a localização W definida como $W = \{x / \|\tilde{S}(x - \tilde{x})\|^2 \leq \varepsilon^2\}$, obtida por alguma das técnicas expostas no capítulo anterior.

O cálculo do limite *a posteriori* dado por (3.3) requer a resolução do seguinte problema:

$$\begin{aligned} \max \quad & \sum_{j=1}^n \delta_j x_j \\ \text{s.a.} \quad & \|\tilde{S}(x - \tilde{x})\| \leq \varepsilon \end{aligned} \tag{5.1}$$

onde $\delta_j = c_j - \bar{u}^t A^j$, para algum $\bar{u} \geq 0$, $\bar{u} \in \mathcal{R}^m$ e A^j indica a j -ésima coluna da matriz de restrições do problema (2.1).



A resolução do problema (5.1) é facilmente obtida com a utilização das condições KKT, que dão:

$$\delta - 2\lambda \tilde{S}^2(x - \tilde{x}) = 0 \quad (5.2)$$

e

$$\lambda[(x - \tilde{x})^t \tilde{S}^2(x - \tilde{x}) - \varepsilon^2] = 0 \quad (5.3)$$

Multiplicando (5.3) por $(x - \tilde{x})$, tem-se:

$$\delta^t(x - \tilde{x}) - 2\lambda(x - \tilde{x})^t \tilde{S}^2(x - \tilde{x}) = 0$$

o que, por (5.3), dá $\delta^t(x - \tilde{x}) - 2\lambda\varepsilon^2 = 0$, donde segue que:

$$\lambda = \frac{\delta^t(x - \tilde{x})}{2\varepsilon^2} \quad (5.4)$$

É evidente que $\lambda \neq 0$, pois a restrição elipsoidal deve estar ativa, de modo que, de (5.2), tem-se:

$$x - \tilde{x} = \frac{1}{2\lambda} \tilde{S}^{-2} \delta \quad (5.5)$$

o que, por (5.4) dá:

$$x - \tilde{x} = \frac{\varepsilon^2}{\delta^t(x - \tilde{x})} \tilde{S}^{-2} \delta \quad (5.6)$$

Multiplicando (5.6) por δ^t , tem-se:

$$[\delta^t(x - \tilde{x})]^2 = \varepsilon^2 \delta^t \tilde{S}^{-2} \delta$$

e, daí,

$$\delta^t(x - \tilde{x}) = \varepsilon(\delta^t \tilde{S}^{-2} \delta)^{1/2}$$

o que implica que o valor ótimo da função objetivo do problema (5.1) é:

$$\delta^t x = \delta^t \tilde{x} + \varepsilon \left(\sum_{j=1}^n \left(\frac{\delta_j}{\tilde{s}_j} \right)^2 \right)^{1/2}$$

Dessa forma, tomando-se a localização $W = \{x / \|\tilde{S}(x - \tilde{x})\|^2 \leq \varepsilon^2\}$, o limite $UB(u, W)$, dado por (3.3), é calculado como:

$$z^* \leq \bar{u}^t b + \sum_{j=1}^n \delta_j \tilde{x}_j + \varepsilon \left[\sum_{j=1}^n (\delta_j / \tilde{s}_j)^2 \right]^{\frac{1}{2}}, \quad (5.7)$$

onde $\delta_j = c_j - \bar{u}^t A^j$, para algum $\bar{u} \geq 0$, $\bar{u} \in \mathcal{R}^m$.

Por outro lado, se $\delta_j = 0$, $\forall j = 1, \dots, n$, então (5.1) não tem solução ótima única, mas, nesse caso, a sua função objetivo ótima é nula.

Dessa forma, o limite (3.3) se torna $z^* \leq \bar{u}^t b$.

Daí, se tomarmos $\bar{u} = U^*$, então por (2.10) podemos escrever $U^{*t} b = Z^* \leq z^* \leq U^{*t} b$, donde segue que $z^* = U^{*t} b$, e a solução desagregada obtida da solução ótima de (2.6) é também ótima para (2.1).

Portanto, podemos concluir que, utilizando-se localizações elipsoidais, o problema de maximização em (3.3) tem solução ótima única, exceto para o caso trivial onde $\delta_j = 0$, $\forall j = 1, \dots, n$, caso em que a solução do problema agregado, quando desagregada, recupera a solução do problema original.

Isto significa que $\Theta(u)$ em (3.14) é diferenciável em $u = U^*$, e nós podemos decrescer o limite (3.3) por alterações no vetor dual (vide seção 3.2, onde foi proposto um processo iterativo que trata dessa abordagem).

Considere, agora, que o centro $\tilde{x}$ do elipsóide seja a solução factível obtida da desagregação da solução ótima do problema agregado, e que $\bar{u} = U^*$. Considere, ainda, que g indique a matriz peso, com estrutura bloco-diagonal, formada pelos elementos dos vetores peso utilizados na obtenção do problema agregado.

Então:

$$\begin{aligned} \delta^t \tilde{x} &= (c - A^t U^*)^t \tilde{x} \\ &= c^t \tilde{x} - U^{*t} A \tilde{x} \\ &= c^t g X^* - U^{*t} A g X^*, \text{ pela definição de desagregação} \\ &= \bar{c}^t X^* - U^{*t} \bar{A} X^*, \text{ devido à construção do problema agregado} \\ &= (\bar{c} - \bar{A}^t U^*)^t X^* \\ &= 0, \text{ pelo teorema das folgas complementares} \end{aligned}$$

Dessa forma, a expressão (5.7) se torna, simplesmente:

$$z^* \leq \bar{u}^t b + \varepsilon \left[\sum_{j=1}^n (\delta_j / \tilde{s}_j)^2 \right]^{\frac{1}{2}}, \quad (5.8)$$

onde $\delta_j = c_j - U^{*t} A^j$ e A^j indica a j -ésima coluna da matriz de restrições original.



Observação 5.1

1. Observe que o cálculo do limite (3.3) é definido no espaço $\mathcal{R}^n$, e alguns dos elipsóides descritos no capítulo 4 são do espaço $\mathcal{R}^{m+n}$.

Entretanto, a utilização desses elipsóides não resulta em qualquer dificuldade na resolução do problema de maximização em (3.3), que pode ser escrito como:

$$\begin{aligned} \max \quad & \sum_{j=1}^n \delta_j \hat{x}_j \\ \text{s.a.} \quad & \hat{x} \in W \end{aligned} \tag{5.9}$$

onde W é um elipsóide do $\mathcal{R}^{m+n}$.

Se $\tilde{x} \in \mathcal{R}^{m+n}$ é a solução ótima de (5.9), então, podemos dizer que $\tilde{x}_j = 0$, $j = n + 1, \dots, m + n$; as outras componentes de $\tilde{x}$ podem ser calculadas similarmente ao problema (5.1).

2. O vetor $\tilde{s}$ desempenha papel significativo na computação do limite (5.7). Esse vetor influi no cálculo do raio do elipsóide, de modo que, quanto menor o valor das componentes de $\tilde{s}$, menor o raio do elipsóide resultante (vide capítulo 4). Por outro lado, nota-se, da expressão (5.7), que valores pequenos para as componentes de $\tilde{s}$ tendem a aumentar o limite para z^* .

Existe, portanto, um valor ótimo para o limite (5.7) em função do vetor $\tilde{s}$ utilizado no seu cálculo.

Se o vetor $\tilde{s}$ não precisa ser factível para o problema (4.5), como acontece com alguns dos elipsóides abordados no capítulo 4, então é possível definir suas componentes $\tilde{s}_j$ em função de δ_j . Podemos, fazer, por exemplo:

$$\tilde{s}_j = \begin{cases} \min_{j=1, \dots, n} \{\delta_j\}, & \text{se } \delta_j < 0 \\ \delta_j, & \text{se } \delta_j \geq 0 \end{cases} \tag{5.10}$$

Essa abordagem não garante, entretanto, que o limite (5.7) seja mínimo. Outras escolhas de $\tilde{s}$ são possíveis. De fato, pode-se utilizar uma abordagem *heurística* para minimizar o limite (5.7) em função do vetor $\tilde{s}$. Não desenvolveremos essa abordagem aqui. No entanto, maiores detalhes sobre *algoritmos heurísticos* podem ser encontrados em [22].



5.2 Localizações tipo caixa

Aplicamos, nesta seção, as localizações elipsoidais na construção de localizações de x^* , ótimo de (2.1), do tipo:

$$W = \{x \in \mathcal{R}^n / l_j \leq x_j \leq L_j, \quad j = 1, \dots, n\} \quad (5.11)$$

e localizações de s^* , ótimo de (2.2), do tipo:

$$V = \{s \in \mathcal{R}^n / r_j \leq s_j \leq R_j, \quad j = 1, \dots, n\} \quad (5.12)$$

A localização (5.11) será aqui chamada de *localização primal tipo caixa*, enquanto (5.12) será referida como *localização dual tipo caixa*.

5.2.1 Localização primal tipo caixa

Observe que localizações do tipo (5.11) podem ser obtidas facilmente das restrições de (2.1) se considerarmos $l_j = 0$ e $L_j = \min_{i=1, \dots, n} \left\{ \frac{b_i}{a_{ij}}, \quad b_i, a_{ij} > 0 \right\}, \quad j = 1, \dots, n$.

Entretanto, esses limites para as variáveis podem se tornar mais precisos quando considerarmos as localizações elipsoidais e um corte para o valor da função objetivo. As proposições a seguir discutem essa abordagem.

Proposição 5.1 [10, 12] - Considere o conjunto elipsoidal $E = \{x \in \mathcal{R}^{m+n} / \|\hat{S}(x - \hat{x})\|^2 \leq \varepsilon^2\}$, obtido por alguma das técnicas da seção 4.1, com centro $\hat{x}$ tal que $\hat{A}\hat{x} = b$. Considere, ainda, para cada $i = 1, \dots, m+n$, os problemas:

$$\begin{aligned} \min \quad & e_i^t x \\ \text{s.a.} \quad & \|\hat{S}(x - \hat{x})\|^2 \leq \varepsilon^2 \\ & \hat{A}x = b \\ & \hat{c}^t x \geq \hat{c}^t \hat{x}, \end{aligned} \quad (5.13)$$

onde $e_i \in \mathcal{R}^{m+n}$ é o i -ésimo vetor unitário.

Então, a solução ótima x_i^- do i -ésimo problema (5.13) é como segue:

$$x_i^- = \hat{x}_i - \varepsilon \sqrt{e_i^t P_{\hat{A}\hat{S}-1} e_i - \frac{(e_i^t P_{\hat{A}\hat{S}-1} \hat{c})^2}{\hat{c}^t P_{\hat{A}\hat{S}-1} \hat{c}}}, \quad \text{se } \hat{c}^t P_{\hat{A}\hat{S}-1} e_i > 0, \quad (5.14)$$

e

$$x_i^- = \hat{x}_i - \varepsilon \sqrt{e_i^t P_{\hat{A}\hat{S}-1} e_i}, \quad \text{se } \hat{c}^t P_{\hat{A}\hat{S}-1} e_i \leq 0, \quad (5.15)$$

onde $P_{\hat{A}\hat{S}-1} = S^{-2}[I - \hat{A}^t(\hat{A}\hat{S}^{-2}\hat{A}^t)^{-1}\hat{A}\hat{S}^{-2}]$.

Demonstração - A Lagrangeana associada com (5.13) é:

$$L(x, \lambda, \hat{u}, \theta) = e_i^t x + \lambda \left[(x - \hat{x})^t \hat{S}^2 (x - \hat{x}) - \varepsilon^2 \right] + \hat{u}^t (\hat{A}x - b) + \theta (\hat{c}^t \hat{x} - \hat{c}^t x)$$

As condições de otimalidade de primeira ordem (condições KKT) para (5.13) dão:

$$e_i + 2\lambda \hat{S}^2 (x - \hat{x}) - \theta \hat{c} + \hat{A}^t \hat{u} = 0 \quad (5.16)$$

$$\lambda [(x - \hat{x})^t \hat{S}^2 (x - \hat{x}) - \varepsilon^2] = 0, \quad (5.17)$$

$$\theta (\hat{c}^t x - \hat{c}^t \hat{x}) = 0, \quad (5.18)$$

$$(5.19)$$

Multiplicando (5.16) por $x - \hat{x}$ e observando que $\hat{A}(x - \hat{x}) = 0$, obtemos:

$$\lambda = \frac{\hat{x}_i - x_i}{2\varepsilon^2} \quad (5.20)$$

Multiplicando (5.16) por $\hat{A}\hat{S}^{-2}$, tem-se:

$$\hat{u} = -(\hat{A}\hat{S}^{-2}\hat{A}^t)^{-1}\hat{A}\hat{S}^{-2}(e_i - \theta\hat{c})$$

De (5.16), temos:

$$\begin{aligned} 2\lambda(\hat{x} - x) &= \hat{S}^{-2}(\hat{A}^t\hat{u} + e_i - \theta\hat{c}) \\ &= P_{\hat{A}\hat{S}^{-1}}(e_i - \theta\hat{c}), \end{aligned} \quad (5.21)$$

onde $P_{\hat{A}\hat{S}^{-1}} = S^{-2}[I - \hat{A}^t(\hat{A}\hat{S}^{-2}\hat{A}^t)^{-1}\hat{A}\hat{S}^{-2}]$.

Multiplicando (5.21) por $\hat{c}^t$, obtemos:

$$0 \geq 2\lambda\hat{c}^t(\hat{x} - x) = \hat{c}^t P_{\hat{A}\hat{S}^{-1}} e_i - \theta \hat{c}^t P_{\hat{A}\hat{S}^{-1}} \hat{c}. \quad (5.22)$$

Se $\hat{c}^t P_{\hat{A}\hat{S}^{-1}} e_i > 0$, então, segue de (5.22) que $\theta > 0$, e, portanto, pelas condições de complementaridade, tem-se $\hat{c}^t(\hat{x} - x) = 0$. Então, de (5.22), temos $\theta = \frac{\hat{c}^t P_{\hat{A}\hat{S}^{-1}} e_i}{\hat{c}^t P_{\hat{A}\hat{S}^{-1}} \hat{c}}$.

Substituindo as expressões para λ e θ em (5.21), vem:

$$2 \left(\frac{\hat{x}_i - x_i}{2\varepsilon^2} \right) (\hat{x} - x) = P_{\hat{A}\hat{S}^{-1}} \left(e_i - \frac{\hat{c}^t P_{\hat{A}\hat{S}^{-1}} e_i}{\hat{c}^t P_{\hat{A}\hat{S}^{-1}} \hat{c}} \hat{c} \right). \quad (5.23)$$

Multiplicando (5.23) por e_i^t , tem-se:

$$\left(\frac{\hat{x}_i - x_i}{\varepsilon^2} \right) (\hat{x}_i - x_i) = e_i^t P_{\hat{A}\hat{S}^{-1}} e_i - \frac{(e_i^t P_{\hat{A}\hat{S}^{-1}} \hat{c})^2}{\hat{c}^t P_{\hat{A}\hat{S}^{-1}} \hat{c}},$$

pois $P_{\hat{A}\hat{S}-1}$ é simétrica e, portanto, $e_i^t P_{\hat{A}\hat{S}-1} \hat{c} = \hat{c}^t P_{\hat{A}\hat{S}-1} e_i$.

Daí, $(\hat{x}_i - x_i)^2 = \varepsilon^2 \left[e_i^t P_{\hat{A}\hat{S}-1} e_i - \frac{(e_i^t P_{\hat{A}\hat{S}-1} \hat{c})^2}{\hat{c}^t P_{\hat{A}\hat{S}-1} \hat{c}} \right]$, donde segue que:

$$x_i^- = \hat{x}_i - \varepsilon \sqrt{e_i^t P_{\hat{A}\hat{S}-1} e_i - \frac{(e_i^t P_{\hat{A}\hat{S}-1} \hat{c})^2}{\hat{c}^t P_{\hat{A}\hat{S}-1} \hat{c}}}.$$

Suponha, agora, que $\hat{c}^t P_{\hat{A}\hat{S}-1} e_i \leq 0$, e que relaxamos a restrição $\hat{c}^t x \geq \hat{c}^t \hat{x}$ em (5.13), ou, equivalentemente, ponha $\theta = 0$ na Lagrangeana.

Então, similar a (5.21), obtemos:

$$2\lambda(\hat{x} - x) = P_{\hat{A}\hat{S}-1} e_i \quad (5.24)$$

Observe que, $\lambda > 0$, pois a restrição elipsoidal deve estar ativa.

Multiplicando (5.24) por $\hat{c}^t$, da hipótese de que $\hat{c}^t P_{\hat{A}\hat{S}-1} e_i \leq 0$, segue que $\hat{c}^t x \leq \hat{c}^t \hat{x}$, o que implica na factibilidade da solução do problema relaxado para o problema (5.13).

Portanto, desde que a solução do problema relaxado é factível para (5.13), então podemos afirmar que ela é também a solução ótima de (5.13).

Dessa forma, multiplicando (5.24) por e_i^t e substituindo λ , obtemos:

$$2 \left(\frac{\hat{x}_i - x_i}{2\varepsilon^2} \right) (\hat{x}_i - x_i) = e_i^t P_{\hat{A}\hat{S}-1} e_i,$$

donde segue que $x_i^- = \hat{x}_i - \varepsilon \sqrt{e_i^t P_{\hat{A}\hat{S}-1} e_i}$, como afirma a tese da proposição.

Observe que, para que a solução do problema (5.13) esteja bem definida, é necessário que os termos da raiz quadrada de (5.14) e (5.15) sejam não-negativos.

Para mostrar que isso sempre acontece, provemos, antes, que a matriz $P_{\hat{A}\hat{S}-1}$ é definida positiva, ou seja, $x^t P_{\hat{A}\hat{S}-1} x > 0$, $\forall x \in \mathcal{R}^{m+n}$, $x \neq 0$.

Note que, $\forall x \in \mathcal{R}^{m+n}$, temos:

$$x^t P_{\hat{A}\hat{S}-1} x = x^t \hat{S}^{-2} x - x^t \hat{S}^{-2} \hat{A}^t (\hat{A} \hat{S}^{-2} \hat{A})^{-1} \hat{A} \hat{S}^{-2} x \quad (5.25)$$

Seja $y = \hat{S}^{-1} x$. Então, de (5.25):

$$\begin{aligned} x^t P_{\hat{A}\hat{S}-1} x &= \|y\|^2 - y^t [\hat{S}^{-1} \hat{A}^t (\hat{A} \hat{S}^{-1} \hat{A}^t)^{-1} \hat{A} \hat{S}^{-1}] y \\ &= y^t [I - \hat{S}^{-1} \hat{A}^t (\hat{A} \hat{S}^{-2} \hat{A}^t)^{-1} \hat{A} \hat{S}^{-1}] y \end{aligned} \quad (5.26)$$

Seja $K = \hat{A} \hat{S}^{-1}$. Daí, de (5.26), tem-se:

$$x^t P_{\hat{A}\hat{S}-1} x = y^t [I - K^t (K K^t)^{-1} K] y$$



Observe que, como $I - K^t(KK^t)^{-1}K$ indica a matriz de projeção sobre o núcleo de K , então $y^t[I - K^t(KK^t)^{-1}K]y > 0$, $\forall y \in \mathcal{R}^{m+n}, y \neq 0$ (vide propriedades das matrizes de projeção em [19]). Portanto, $x^t P_{\hat{A}\hat{S}^{-1}} x > 0$, $\forall x \in \mathcal{R}^{m+n}, x \neq 0$, de modo que $P_{\hat{A}\hat{S}^{-1}}$ é definida positiva e admite uma decomposição de Cholesky da forma $P_{\hat{A}\hat{S}^{-1}} = L^t L$, onde L é triangular superior (vide [19]).

Daí, tomando-se $v = Le_i$ e $u = L\hat{c}$, tem-se:

$$\begin{aligned} e_i^t P_{\hat{A}\hat{S}^{-1}} e_i - \frac{(\hat{c}^t P_{\hat{A}\hat{S}^{-1}} e_i)^2}{\hat{c}^t P_{\hat{A}\hat{S}^{-1}} \hat{c}} &= e_i^t L^t L e_i - \frac{(e_i^t L^t L \hat{c})^2}{\hat{c}^t L^t L \hat{c}} \\ &= \|v\|^2 - \frac{(v^t u)^2}{\|u\|^2} \\ &= \frac{\|u\|^2 \|v\|^2 - (v^t u)^2}{\|u\|^2} \geq 0, \end{aligned}$$

devido à desigualdade de Cauchy-Schwarz [19], o que mostra a positividade do termo sob a raiz quadrada em (5.14).

Da mesma forma, em (5.15), temos $e_i^t P_{\hat{A}\hat{S}^{-1}} e_i = e_i^t L^t L e_i = \|v\|^2 \geq 0$.

Observe o que acontece, agora, quando $\hat{c}^t P_{\hat{A}\hat{S}^{-1}} e_i < 0$ e $\hat{c}^t P_{\hat{A}\hat{S}^{-1}} \hat{c} = 0$. Nessas condições, devido às propriedades das matrizes de projeção, tem-se que $y = \hat{S}^{-1} \hat{c}$ é ortogonal ao núcleo da matriz $K = \hat{A}\hat{S}^{-1}$, o que implica que $\exists w$ tal que $y = Kw$, ou seja, $\hat{S}^{-1} \hat{c} = \hat{S}^{-1} \hat{A}w$, ou, equivalentemente, $\hat{c} = \hat{A}w$, para algum $w \in \mathcal{R}^{m+n}$. Daí,

$$\hat{c}^t x = w^t \hat{A}x = w^t b \quad (5.27)$$

e

$$\hat{c}^t \hat{x} = w^t \hat{A}\hat{x} = w^t b \quad (5.28)$$

de modo que podemos desconsiderar a restrição $\hat{c}^t x \geq \hat{c}^t \hat{x}$ em (5.13), o que significa tomar $\theta = 0$ na Lagrangeana.

Daí, multiplicando (5.24) por $\hat{c}^t$ e utilizando (5.27) e (5.28), tem-se $\hat{c}^t P_{\hat{A}\hat{S}^{-1}} e_i = 0$, o que contraria nossa hipótese de que $\hat{c}^t P_{\hat{A}\hat{S}^{-1}} e_i < 0$, de modo que as situações $\hat{c}^t P_{\hat{A}\hat{S}^{-1}} e_i < 0$ e $\hat{c}^t P_{\hat{A}\hat{S}^{-1}} \hat{c} = 0$ não podem ocorrer simultaneamente.

Portanto, com base nessas considerações, podemos dizer que as expressões (5.14) e (5.15) estão bem definidas, finalizando, assim a demonstração da proposição. •

Considere, agora, a obtenção de limites superiores para $\hat{x}^*$.

Proposição 5.2 [10, 12] - Considere a localização elipsoidal $E = \{x \in \mathcal{R}^{m+n} / \|\hat{S}(x - \hat{x})\|^2 \leq \varepsilon^2\}$ obtida por alguma das técnicas propostas na seção 4.1, com centro $\hat{x}$ tal que $\hat{A}\hat{x} = b$. Então,

para cada $i = 1, \dots, m + n$, a solução ótima x_i^+ do problema:

$$\begin{aligned} \max \quad & e_i^t x \\ \text{s.a.} \quad & \|\hat{S}(x - \hat{x})\|^2 \leq \varepsilon^2 \\ & \hat{A}x = b \\ & \hat{c}^t x \geq \hat{c}^t \hat{x} \end{aligned} \quad (5.29)$$

é como segue:

$$x_i^+ = \hat{x}_i + \varepsilon \sqrt{e_i^t P_{\hat{A}\hat{S}^{-1}} e_i - \frac{(\hat{c}^t P_{\hat{A}\hat{S}^{-1}} e_i)^2}{\hat{c}^t P_{\hat{A}\hat{S}^{-1}} \hat{c}}}, \text{ se } \hat{c}^t P_{\hat{A}\hat{S}^{-1}} e_i < 0, \quad (5.30)$$

e

$$x_i^+ = \hat{x}_i + \varepsilon \sqrt{e_i^t P_{\hat{A}\hat{S}^{-1}} e_i}, \text{ se } \hat{c}^t P_{\hat{A}\hat{S}^{-1}} e_i \geq 0, \quad (5.31)$$

onde $P_{\hat{A}\hat{S}^{-1}}$ é definida como na proposição (5.1).

Demonstração - A Lagrangeana associada com (5.29) é:

$$L(x, \hat{u}, \lambda, \theta) = e_i^t x - \lambda \left[(x - \hat{x})^t \hat{S}^2 (x - \hat{x}) - \varepsilon^2 \right] + \hat{u}^t (\hat{A}x - b) - \theta (\hat{c}^t \hat{x} - \hat{c}^t x)$$

Da mesma forma que na proposição (5.1), obtemos:

$$\lambda = (x_i - \hat{x}_i) / 2\varepsilon^2, \quad (5.32)$$

$$\hat{u} = -(\hat{A}\hat{S}^{-2}\hat{A}^t)^{-1} \hat{A}\hat{S}^{-2}(\theta\hat{c} + e_i), \text{ e} \quad (5.33)$$

$$2\lambda(x - \hat{x}) = P_{\hat{A}\hat{S}^{-1}}(e_i + \theta\hat{c}), \quad (5.34)$$

onde $P_{\hat{A}\hat{S}^{-1}} = S^{-2}[I - \hat{A}^t(\hat{A}\hat{S}^{-2}\hat{A}^t)^{-1}\hat{A}\hat{S}^{-2}]$.

Multiplicando (5.34) por $\hat{c}^t$, obtemos:

$$0 \leq 2\lambda(x - \hat{x}) = \hat{c}^t P_{\hat{A}\hat{S}^{-1}} e_i + \theta \hat{c}^t P_{\hat{A}\hat{S}^{-1}} \hat{c} \quad (5.35)$$

Se $\hat{c}^t P_{\hat{A}\hat{S}^{-1}} e_i < 0$, então deve-se ter $\theta > 0$. Daí, pelas condições de complementaridade, tem-se $\hat{c}^t x = \hat{c}^t \hat{x}$, donde segue, por (5.35) que $\theta = -\frac{\hat{c}^t P_{\hat{A}\hat{S}^{-1}} e_i}{\hat{c}^t P_{\hat{A}\hat{S}^{-1}} \hat{c}}$.

Substituindo as expressões de λ e θ em (5.34), e multiplicando o resultado por e_i^t , tem-se, analogamente à proposição (5.1),

$$(x_i - \hat{x}_i)^2 = \varepsilon^2 \left(e_i^t P_{\hat{A}\hat{S}^{-1}} e_i - \frac{(\hat{c}^t P_{\hat{A}\hat{S}^{-1}} e_i)^2}{\hat{c}^t P_{\hat{A}\hat{S}^{-1}} \hat{c}} \right).$$

Daí, $x_i^+ = \hat{x}_i + \varepsilon \sqrt{e_i^t P_{\hat{A}\hat{S}^{-1}} e_i - \frac{(\hat{c}^t P_{\hat{A}\hat{S}^{-1}} e_i)^2}{\hat{c}^t P_{\hat{A}\hat{S}^{-1}} \hat{c}}}$, o que demonstra a primeira parte da proposição.

Considere, agora, que $\hat{c}^t P_{\hat{A}\hat{S}-1} e_i \geq 0$ e que relaxamos a restrição $\hat{c}^t x \geq \hat{c}^t \hat{x}$ em (5.29), ou, equivalentemente, ponha $\theta = 0$ na Lagrangeana.

Então, similar a (5.34), obtemos:

$$2\lambda(x - \hat{x}) = P_{\hat{A}\hat{S}-1}(e_i) \quad (5.36)$$

Observe que, se $\lambda = 0$, então, por (5.20), tem-se $x_i^+ = \hat{x}_i$.

Considere $\lambda > 0$ em (5.36). Então, da hipótese de que $\hat{c}^t P_{\hat{A}\hat{S}-1} e_i \geq 0$, segue que $\hat{c}^t x \geq \hat{c}^t \hat{x}$, donde se observa que a solução do problema relaxado é factível para (5.29), e, portanto, também é ótima para (5.29).

Dessa forma, multiplicando (5.36) por e_i^t , e substituindo a expressão de λ , obtemos $x_i^+ = \hat{x}_i + \varepsilon \sqrt{e_i^t P_{\hat{A}\hat{S}-1} e_i}$, conforme requerido. Note que essas soluções estão bem definidas por que a matriz $P_{\hat{A}\hat{S}-1}$ é definida positiva, completando, dessa forma, a demonstração da proposição. •

Com os resultados fornecidos pela proposições (5.1) e (5.2), o conjunto (5.11) é facilmente construído. Basta fazer $l_j = [x_j^-]^+ = \max\{0, x_j^-\}$ e $L_j = x_j^+$, $j = 1, \dots, m+n$.

A localização (5.11) pode ser utilizada diretamente no cálculo do limite (3.3) ou pode ser utilizada em conjunto com outra localização.

Podemos, por exemplo, acrescentar à localização (5.11) a i -ésima restrição do problema original, ficando assim com a localização:

$$W = \left\{ x \in \mathcal{R}^n / \sum_{j=1}^n a_{ij} x_j \leq b_i, \quad l_j \leq x_j \leq L_j, \quad j = 1, \dots, n, \quad \text{e algum } i \in \{1, \dots, m\} \right\} \quad (5.37)$$

Podemos, em seguida, utilizar a localização (5.37) no cálculo da expressão (3.3), resultando, dessa abordagem, o seguinte problema:

$$\begin{aligned} \max \quad & \sum_{j=1}^n \delta_j x_j \\ \text{s.a.} \quad & \sum_{j=1}^n a_{ij} x_j \leq b_i, \quad \text{para algum } i \in \{1, \dots, m\} \\ & l_j \leq x_j \leq L_j, \quad j = 1, \dots, n \end{aligned} \quad (5.38)$$

que é facilmente transformado no problema da mochila contínuo com limites superiores para as variáveis, se tomarmos, em (5.38), $y_j = x_j - l_j$, ficando, assim, com o problema:

$$\begin{aligned} \max \quad & \sum_{j=1}^n \delta_j y_j + \sum_{j=1}^n \delta_j l_j \\ \text{s.a.} \quad & \sum_{j=1}^n a_{ij} y_j \leq b_i - \sum_{j=1}^n a_{ij} l_j, \quad \text{para algum } i \in \{1, \dots, m\} \\ & 0 \leq y_j \leq L_j - l_j, \quad j = 1, \dots, n \end{aligned} \quad (5.39)$$



cuja resolução foi discutida na seção 3.3.

Outra alternativa seria utilizar a localização (5.11) juntamente com uma restrição do tipo $c^t x \leq UB$, onde UB é tal que $c^t x^* \leq UB$, obtendo, assim, a localização:

$$W = \left\{ x \in \mathcal{R}^n / \sum_{j=1}^n c_j x_j \leq UB, \ l_j \leq x_j \leq L_j, \ x_j \geq 0, \ j = 1, \dots, n \right\} \quad (5.40)$$

A utilização da localização (5.40) em (3.3) dá origem ao seguinte problema:

$$\begin{aligned} \max \quad & \sum_{j=1}^n \delta_j x_j \\ \text{s.a.} \quad & \sum_{j=1}^n c_j x_j \leq UB \\ & l_j \leq x_j \leq L_j, \ j = 1, \dots, n \end{aligned} \quad (5.41)$$

Considere $y_j = x_j - l_j$ em (5.41). Com isso, o problema (5.41) pode ser reescrito como:

$$\begin{aligned} \max \quad & \sum_{j=1}^n \delta_j y_j + \sum_{j=1}^n \delta_j l_j \\ \text{s.a.} \quad & \sum_{j=1}^n c_j y_j \leq UB - \sum_{j=1}^n c_j l_j \\ & 0 \leq y_j \leq L_j - l_j, \ j = 1, \dots, n, \end{aligned} \quad (5.42)$$

cuja solução é facilmente computada de acordo com as considerações da seção 3.3.

Observe que o cálculo de limites superiores para z^* utilizando essa última abordagem pode ser feito iterativamente. De fato, se se conhece um valor inicial para UB , então, pode-se utilizar esse valor em (5.41) e computar um novo valor para UB por (3.3). Em seguida, podemos utilizar esse novo valor de UB em (5.41) e recalculamos o limite por (3.3). Esse processo continua até que o valor de UB permaneça inalterado.

Entretanto, não podemos garantir a convergência de UB para z^* . Note, porém, que esse processo pode resultar em um melhoramento significativo para o valor de UB (vide resultados dos testes computacionais na seção 5.4).

5.2.2 Localização dual tipo caixa

As proposições a seguir consideram a obtenção de localizações duais tipo caixa a partir dos elipsóides duais.

Proposição 5.3 [10] - Considere o conjunto elipsoidal $E = \{s \in \mathcal{R}^{m+n} / \|\hat{X}(s - \hat{s})\| \leq \varepsilon\}$, obtido por alguma das técnicas descritas na seção 4.2. Considere, ainda, para cada $i = 1, \dots, m + n$, os problemas:



$$\begin{aligned}
& \min e_i^t s \\
& \text{s.a. } \|\hat{X}(s - \hat{s})\|^2 \leq \varepsilon^2 \\
& \quad s = \hat{A}^t u - \hat{c} \\
& \quad b^t u \leq b^t \hat{u}
\end{aligned} \tag{5.43}$$

Então, a solução ótima s_i^- do i -ésimo problema (5.43) é como segue:

$$s_i^- = \hat{s}_i - \varepsilon \sqrt{e_i^t \hat{A}^t (\hat{A} \hat{X}^2 \hat{A}^t)^{-1} \left(\hat{A} e_i - \frac{b^t (\hat{A} \hat{X}^2 \hat{A}^t)^{-1} \hat{A} e_i}{b^t (\hat{A} \hat{X}^2 \hat{A}^t)^{-1} b} \right)}, \text{ se } b^t (\hat{A} \hat{X}^2 \hat{A}^t)^{-1} \hat{A} e_i < 0,$$

e

$$s_i^- = \hat{s}_i - \varepsilon \sqrt{e_i^t \hat{A}^t (\hat{A} \hat{X}^2 \hat{A}^t)^{-1} \hat{A} e_i}, \text{ se } b^t (\hat{A} \hat{X}^2 \hat{A}^t)^{-1} \hat{A} e_i \geq 0.$$

Demonstração - A Lagrangeana associada com o i -ésimo problema acima é:

$$L(u, s, \lambda, \pi, \sigma) = e_i^t s + \lambda [(s - \hat{s})^t \hat{X}^2 (s - \hat{s}) - \varepsilon^2] + \pi (s - \hat{A}^t u + \hat{c}) + \sigma (b^t u - b^t \hat{u}).$$

As condições KKT para (5.43) são:

$$e_i + \pi + 2\lambda [\hat{X}^2 (s - \hat{s})] = 0 \tag{5.44}$$

$$-\hat{A}^t \pi + \sigma b = 0 \tag{5.45}$$

$$\sigma b^t (u - \hat{u}) = 0 \tag{5.46}$$

$$\lambda [(s - \hat{s})^t \hat{X}^2 (s - \hat{s}) - \varepsilon^2] = 0 \tag{5.47}$$

$$\lambda \geq 0, \sigma \geq 0 \tag{5.48}$$

De (5.45) e de (5.46), tem-se:

$$\pi^t (s - \hat{s}) = \pi^t \hat{A}^t (u - \hat{u}) = \sigma b^t (u - \hat{u}) = 0 \tag{5.49}$$

Multiplicando (5.44) por $(s - \hat{s})^t$, temos:

$$s_i - \hat{s}_i + \pi^t (s - \hat{s}) + 2\lambda (s - \hat{s})^t \hat{X}^2 (s - \hat{s}) = 0.$$

Portanto, por (5.49) e (5.47), obtemos $s_i - \hat{s}_i + 2\lambda \varepsilon^2 = 0$, donde segue que:

$$\lambda = \frac{\hat{s}_i - s_i}{2\varepsilon^2}. \tag{5.50}$$

Agora, multiplicando (5.44) por $\hat{A}$, tem-se:

$$\hat{A} e_i + \sigma b + 2\lambda \hat{A} \hat{X}^2 \hat{A}^t (u - \hat{u}) = 0$$

Daí,

$$u - \hat{u} = -\frac{1}{2\lambda}(\hat{A}\hat{X}^2\hat{A}^t)^{-1}(\hat{A}e_i + \sigma b). \quad (5.51)$$

Multiplicando (5.51) por b^t , vem:

$$0 \geq b^t(u - \hat{u}) = -\frac{1}{2\lambda}b^t(\hat{A}\hat{X}^2\hat{A}^t)^{-1}(\hat{A}e_i + \sigma b) \quad (5.52)$$

Portanto, se $b^t(\hat{A}\hat{X}^2\hat{A}^t)^{-1}\hat{A}e_i < 0$ em (5.52), deve-se ter $\sigma > 0$.

Então, por (5.46), tem-se $b^t(u - \hat{u}) = 0$, de modo que, por (5.52),

$$\sigma = -\frac{b^t(\hat{A}\hat{X}^2\hat{A}^t)^{-1}\hat{A}e_i}{b^t(\hat{A}\hat{X}^2\hat{A}^t)^{-1}b} \quad (5.53)$$

Daí, de (5.51), temos:

$$u - \hat{u} = -\frac{\varepsilon^2}{\hat{s}_i - s_i}(\hat{A}\hat{X}^2\hat{A}^t)^{-1}\left(\hat{A}e_i - \frac{b^t(\hat{A}\hat{X}^2\hat{A}^t)^{-1}\hat{A}e_i}{b^t(\hat{A}\hat{X}^2\hat{A}^t)^{-1}b}b\right) \quad (5.54)$$

Multiplicando (5.54) por $e_i^t\hat{A}^t$, tem-se:

$$e_i^t\hat{A}^t(u - \hat{u}) = -\frac{\varepsilon^2}{\hat{s}_i - s_i}e_i^t\hat{A}^t(\hat{A}\hat{X}^2\hat{A}^t)^{-1}\left(\hat{A}e_i - \frac{b^t(\hat{A}\hat{X}^2\hat{A}^t)^{-1}\hat{A}e_i}{b^t(\hat{A}\hat{X}^2\hat{A}^t)^{-1}b}b\right),$$

que é equivalente a

$$e_i^t(\hat{s} - s) = \frac{\varepsilon^2}{\hat{s}_i - s_i}e_i^t\hat{A}^t(\hat{A}\hat{X}^2\hat{A}^t)^{-1}\left(\hat{A}e_i - \frac{b^t(\hat{A}\hat{X}^2\hat{A}^t)^{-1}\hat{A}e_i}{b^t(\hat{A}\hat{X}^2\hat{A}^t)^{-1}b}b\right),$$

ou seja,

$$(\hat{s}_i - s_i)^2 = \varepsilon^2 \left[e_i^t\hat{A}^t(\hat{A}\hat{X}^2\hat{A}^t)^{-1}\left(\hat{A}e_i - \frac{b^t(\hat{A}\hat{X}^2\hat{A}^t)^{-1}\hat{A}e_i}{b^t(\hat{A}\hat{X}^2\hat{A}^t)^{-1}b}b\right) \right],$$

donde segue que:

$$s_i^- = \hat{s}_i - \varepsilon \sqrt{e_i^t\hat{A}^t(\hat{A}\hat{X}^2\hat{A}^t)^{-1}\left(\hat{A}e_i - \frac{b^t(\hat{A}\hat{X}^2\hat{A}^t)^{-1}\hat{A}e_i}{b^t(\hat{A}\hat{X}^2\hat{A}^t)^{-1}b}b\right)}$$

o que demonstra a primeira parte da proposição.

Considere, agora, que $b^t(\hat{A}\hat{X}^2\hat{A}^t)^{-1}\hat{A}e_i \geq 0$, e suponha que relaxamos a restrição $b^tu \leq b^t\hat{u}$ em (5.43), ou, equivalentemente, suponha $\sigma = 0$ na Lagrangeana.

Então, de (5.51), segue que:

$$u - \hat{u} = -\frac{\varepsilon^2}{\hat{s}_i - s_i}(\hat{A}\hat{X}^2\hat{A}^t)^{-1}\hat{A}e_i \quad (5.55)$$

daí, como $\hat{s}_i - s_i > 0$ em (5.55), e da hipótese de que $b^t(\hat{A}\hat{X}^2\hat{A}^t)^{-1}\hat{A}e_i \geq 0$, multiplicando (5.55) por b^t , segue que $b^tu \leq b^t\hat{u}$.

Logo, como a solução s_i^- do problema relaxado é factível para o problema (5.43), então, ela é também ótima para (5.43).

Daí, multiplicando (5.55) por $e_i^t\hat{A}^t$, obtém-se:

$$e_i^t\hat{A}^t(\hat{u} - u) = \frac{\varepsilon^2}{\hat{s}_i - s_i} e_i^t\hat{A}^t(\hat{A}\hat{X}^2\hat{A}^t)^{-1}\hat{A}e_i$$

que é equivalente a

$$e_i^t(\hat{s} - s) = \frac{\varepsilon^2}{\hat{s}_i - s_i} e_i^t\hat{A}^t(\hat{A}\hat{X}^2\hat{A}^t)^{-1}\hat{A}e_i,$$

ou seja,

$$(\hat{s}_i - s_i)^2 = \varepsilon^2 e_i^t\hat{A}^t(\hat{A}\hat{X}^2\hat{A}^t)^{-1}\hat{A}e_i,$$

donde segue que:

$$s_i^- = \hat{s}_i - \varepsilon \sqrt{e_i^t\hat{A}^t(\hat{A}\hat{X}^2\hat{A}^t)^{-1}\hat{A}e_i},$$

completando, assim, a demonstração da proposição. •

A proposição a seguir considera a obtenção de limites superiores para $\hat{s}^*$.

Proposição 5.4 - Considere o conjunto elipsoidal $E = \{s \in \mathcal{R}^{m+n} / \|\hat{X}(s - \hat{s})\| \leq \varepsilon\}$, obtido por alguma das técnicas descritas na seção 4.2. Considere, ainda, para cada $i = 1, \dots, m+n$, os problemas:

$$\begin{aligned} \max \quad & e_i^t s \\ \text{s.a.} \quad & \|\hat{X}(s - \hat{s})\|^2 \leq \varepsilon^2 \\ & s = \hat{A}^t u - \hat{c} \\ & b^t u \leq b^t \hat{u} \end{aligned} \tag{5.56}$$

Então, a solução ótima s_i^+ do i -ésimo problema (5.56) é como segue:

$$s_i^+ = \hat{s}_i + \varepsilon \sqrt{e_i^t\hat{A}^t(\hat{A}\hat{X}^2\hat{A}^t)^{-1}\hat{A}e_i - \frac{(e_i^t\hat{A}^t(\hat{A}\hat{X}^2\hat{A}^t)^{-1}\hat{A}e_i)^2}{b^t(\hat{A}\hat{X}^2\hat{A}^t)^{-1}\hat{A}e_i}}, \text{ se } b^t(\hat{A}\hat{X}^2\hat{A}^t)^{-1}\hat{A}e_i > 0,$$

e

$$s_i^+ = \hat{s}_i + \varepsilon \sqrt{e_i^t\hat{A}^t(\hat{A}\hat{X}^2\hat{A}^t)^{-1}\hat{A}e_i}, \text{ se } b^t(\hat{A}\hat{X}^2\hat{A}^t)^{-1}\hat{A}e_i \leq 0.$$

Demonstração - A Lagrangeana associada com o i -ésimo problema (5.4) é:

$$L(u, s, \lambda, \pi, \sigma) = e_i^t s - \lambda \left[(s - \hat{s})^t \hat{X}^2 (s - \hat{s}) - \varepsilon^2 \right] + \pi (s - \hat{A}^t u + \hat{c}) - \sigma (b^t u - b^t \hat{u})$$

Da mesma forma que na proposição (5.43), obtemos:

$$\lambda = \frac{s_i - \hat{s}_i}{2\varepsilon^2} \quad (5.57)$$

e

$$u - \hat{u} = \frac{1}{2\lambda}(\hat{A}\hat{X}^2\hat{A}^t)^{-1}(\hat{A}e_i - \sigma b) \quad (5.58)$$

Multiplicando (5.58) por b^t , vem:

$$0 \geq b^t(u - \hat{u}) = \frac{1}{2\lambda}b^t(\hat{A}\hat{X}^2\hat{A}^t)^{-1}(\hat{A}e_i - \sigma b) \quad (5.59)$$

Portanto, se $b^t(\hat{A}\hat{X}^2\hat{A}^t)^{-1}\hat{A}e_i > 0$ em (5.59), segue de $\sigma > 0$. Daí, das condições de complementaridade, tem-se $b^tu = b^t\hat{u}$, donde segue, de (5.59), que $\sigma = \frac{b^t(\hat{A}\hat{X}^2\hat{A}^t)^{-1}\hat{A}e_i}{b^t(\hat{A}\hat{X}^2\hat{A}^t)^{-1}b}$.

Daí, utilizando as expressões obtidas para λ e σ em (5.58), tem-se:

$$u - \hat{u} = \frac{\varepsilon^2}{s_i - \hat{s}_i}(\hat{A}\hat{X}^2\hat{A}^t)^{-1} \left(\hat{A}e_i - \frac{b^t(\hat{A}\hat{X}^2\hat{A}^t)^{-1}\hat{A}e_i}{b^t(\hat{A}\hat{X}^2\hat{A}^t)^{-1}b}b \right) \quad (5.60)$$

Multiplicando (5.60) por $e_i^t\hat{A}^t$, vem:

$$(s_i - \hat{s}_i)^2 = \varepsilon^2 e_i^t \hat{A}^t (\hat{A}\hat{X}^2\hat{A}^t)^{-1} \left(\hat{A}e_i - \frac{b^t(\hat{A}\hat{X}^2\hat{A}^t)^{-1}\hat{A}e_i}{b^t(\hat{A}\hat{X}^2\hat{A}^t)^{-1}b}b \right),$$

donde segue que

$$s_i^+ = \hat{s}_i + \varepsilon \sqrt{e_i^t \hat{A}^t (\hat{A}\hat{X}^2\hat{A}^t)^{-1} \hat{A}e_i - \frac{(e_i^t \hat{A}^t (\hat{A}\hat{X}^2\hat{A}^t)^{-1}b)^2}{b^t(\hat{A}\hat{X}^2\hat{A}^t)^{-1}b}},$$

completando, assim, a demonstração da primeira parte da proposição.

Considere, agora, $b^t(\hat{A}\hat{X}^2\hat{A}^t)^{-1}\hat{A}e_i \leq 0$, e suponha que relaxamos a restrição $b^tu \leq b^t\hat{u}$ em (5.4), ou, equivalentemente, suponha $\sigma = 0$ na Lagrangeana.

Então, de (5.58), segue que:

$$u - \hat{u} = \frac{1}{2\lambda}(\hat{A}\hat{X}^2\hat{A}^t)^{-1}\hat{A}e_i \quad (5.61)$$

Como, por hipótese, $b^t(\hat{A}\hat{X}^2\hat{A}^t)^{-1}\hat{A}e_i \leq 0$, então, multiplicando (5.61) por b^t , tem-se $b^tu \leq b^t\hat{u}$, donde segue que a solução ótima do problema relaxado é factível para (5.56), e portanto, é também ótima para (5.56).

Daí, multiplicando (5.61) por $e_i^t\hat{A}^t$, vem:

$$s_i - \hat{s}_i = \frac{\varepsilon^2}{s_i - \hat{s}_i} e_i^t \hat{A}^t (\hat{A}\hat{X}^2\hat{A}^t)^{-1} \hat{A}e_i,$$

donde segue que:

$$s_i = \hat{s}_i + \varepsilon \sqrt{e_i^t \hat{A}^t (\hat{A} \hat{X}^2 \hat{A}^t)^{-1} \hat{A} e_i},$$

o que finaliza a demonstração da proposição. •

A localização (5.12) pode ser obtida facilmente do resultado das proposições (5.3) e (5.4). Basta fazer $r_j = [s_j^-]^+ = \max \{0, s_j^-\}$ e $R_j = s_j^+$, $j = 1, \dots, n$.

Dispondo da localização (5.12), se tomarmos os vetores s^- e s^+ do $\mathcal{R}^n$, dados por $s^- = ([s_j^-]^+)$ e $s^+ = (s_j^+)$, então, com base nas restrições do problema (2.2), temos $A^t u^* \leq c + s^*$. Como $s^- \leq s^* \leq s^+$, então, podemos escrever:

$$c + s^- \leq A^t u^* \leq c + s^+, \quad \forall u^* \text{ ótimo de (2.2)} \quad (5.62)$$

Observe que a expressão (5.62) define uma localização de u^* e pode ser utilizada para orientar a busca por soluções duais a serem utilizadas no cálculo do limite (3.3).

Podemos, por exemplo, considerar, similarmente à abordagem proposta na seção 3.2, o vetor u que resolva (aproximadamente) o problema:

$$\min \{UB(u, W)/c + s^- \leq A^t u \leq c + s^+\} \quad (5.63)$$

Para simplificar o problema (5.63), podemos considerar somente uma restrição do sistema (5.62) na definição da localização de u^* . Com isso, o problema (5.63) pode ser escrito como:

$$\min \left\{ UB(u, W)/c_j + s_j^- \leq \sum_{i=1}^m a_{ij} u_i \leq c_j + s_j^+ \right\}, \text{ para algum } j \in \{1, \dots, n\}. \quad (5.64)$$

e pode ser tratado pelo processo iterativo sugerido na seção 3.2

5.3 Exemplo

Passamos, agora, a exemplificar os conceitos desenvolvidos anteriormente, considerando o problema abaixo, extraído de Zipkin [1]:

$$\begin{aligned} \max \quad & 2.5x_1 + 3x_2 + 4x_3 + 5x_4 \\ \text{s.a.} \quad & 4x_1 + 5x_2 + 7x_3 + 10x_4 \leq 54 \\ & x_1 + 2x_2 + x_3 + 2x_4 \leq 10, \\ & x_1, x_2, x_3, x_4 \geq 0 \end{aligned} \quad (5.65)$$

que tem solução ótima $x^* = (\frac{16}{3}, 0, \frac{14}{3}, 0)$, o que dá $z^* = 32$.

Tomando a partição $\sigma = \{\{1, 2\}, \{3, 4\}\}$ e os vetores peso $g^1 = g^2 = (\frac{1}{2}, \frac{1}{2})$, então, o problema agregado é dado por:

$$\begin{aligned} \max \quad & 2.75X_1 + 4.5X_2 \\ \text{s.a.} \quad & 4.5X_1 + 8.5X_2 \leq 54 \\ & 1.5X_1 + 1.5X_2 \leq 10, \\ & X_1 \geq 0, X_2 \geq 0 \end{aligned} \quad (5.66)$$

A solução ótima de (5.66) é $X^* = (\frac{2}{3}, 6)$, e, conseqüentemente, $Z^* = 28\frac{5}{6}$. O dual do problema (5.66) tem solução $U^* = \frac{1}{48}(21, 25)$.

Uma solução desagregada $\bar{x}$, obtida da desagregação de peso fixo de X^* , é dada por $\bar{x}_j = g_j^k X_k^*, j \in S_k, k = 1, 2$, ou seja,

$$\begin{aligned} \bar{x}_1 &= g_1^1 X_1^* = \frac{1}{2} \times \frac{2}{3} = \frac{1}{3} \\ \bar{x}_2 &= g_2^1 X_1^* = \frac{1}{2} \times \frac{2}{3} = \frac{1}{3} \\ \bar{x}_3 &= g_1^2 X_2^* = \frac{1}{2} \times 6 = 3 \\ \bar{x}_4 &= g_2^2 X_2^* = \frac{1}{2} \times 6 = 3 \end{aligned}$$

Observe que $\bar{x} = (\frac{1}{3}, \frac{1}{3}, 3, 3)$ é factível para (5.65) e $c^t \bar{x} = 28\frac{5}{6} = Z^*$, conforme afirmam a proposição (2.1) e o corolário (2.1).

Obtivemos, desse modo, um limite inferior para z^* . Consideremos, agora, a obtenção de limites superiores.

Nos cálculos a seguir, consideraremos o vetor δ , dado por $\delta = c - A^t U^* = (\frac{11}{48}, -\frac{11}{48}, \frac{20}{48}, -\frac{20}{48})$.

Considerando-se a partição $\tilde{\sigma} = \{1, 2, 3, 4\}$, e os números p_k , e $d_j, j = 1, 2, 3, 4$ tomados da primeira restrição do problema (5.65), então, o limite (2.17), proposto por Zipkin [1], é dado por:

$$\begin{aligned} z^* &\leq U^{*t} b + \sum_{k=1}^{\bar{K}} \max_{j \in \tilde{S}_K} \left[\frac{c_j - U^{*t} A^j}{d_j} \right]^+ p_k \\ &= 28\frac{5}{6} + \max_{j=1, \dots, 4} \left\{ \frac{11}{192}, 0, \frac{20}{336}, 0 \right\} \times 54 \\ &= 28\frac{5}{6} + \frac{20}{336} \times 54 \\ &= 32,0476 \end{aligned}$$

Suponha, agora, que tomamos os números p_k , e $d_j, j = 1, \dots, 4$, da segunda restrição do problema original. Daí, por (2.17), temos:

$$z^* \leq U^{*t} b + \sum_{k=1}^{\bar{K}} \max_{j \in \tilde{S}_K} \left[\frac{c_j - U^{*t} A^j}{d_j} \right]^+ p_k$$



$$\begin{aligned}
&= 28\frac{5}{6} + \max_{j=1,\dots,4} \left\{ \frac{11}{48}, 0, \frac{20}{48}, 0 \right\} \times 10 \\
&= 28\frac{5}{6} + \frac{20}{48} \times 10 \\
&= 33
\end{aligned}$$

Dessa forma, a utilização da primeira restrição de (5.65) resulta em mais precisão para o limite proposto por Zipkin. Para computar os limites (2.18) e (2.21), propostos por Mendelsohn[4, 5] e Taylor/Shetty[5], respectivamente, podemos aplicar a Busca de Fibonacci aos problemas de minimização dessas expressões. Não descreveremos aqui esse processo de busca. Deixamos, entretanto, como referência, os textos de Bazaraa[17] e Luenberger[18], onde se pode encontrar uma descrição detalhada desse algoritmo.

Tomando-se os números p_k e $d_j, j = 1, \dots, 4$, da primeira restrição do problema original, então, aplicando-se a Busca de Fibonacci aos problemas de minimização em (2.18) e (2.21), obtemos $z^* \leq 32,0004$ com $\theta^* = 0,96$, para o limite de Mendelsohn, e $z^* \leq 32$, com $\Phi^* = -0,0021$, para o limite proposto por Taylor/Shetty.

Da mesma forma, se os números p_k e $d_j, j = 1, \dots, 4$, forem tomados da segunda restrição de (5.65), então, como resultado da Busca de Fibonacci, obtemos $z^* \leq 32,1860$, com $\theta^* = 1,1162$, para o limite de Mendelsohn, e $z^* \leq 32,07221$, com $\Phi^* = 0,001074$, para o limite de Taylor/Shetty.

Considere, agora, a aplicação do algoritmo proposto na seção 3.2.

Tomando-se $u^o = U^*$ e o conjunto W definido pela primeira restrição de (5.65), então, o problema (3.14) se torna:

$$\begin{aligned}
&\max \quad \frac{11}{48}x_1 - \frac{11}{48}x_2 + \frac{20}{48}x_3 - \frac{20}{48}x_4 \\
&\text{s.a.} \quad 4x_1 + 5x_2 + 7x_3 + 10x_4 \leq 54 \\
&\quad \quad x_i \geq 0, \quad i = 1, \dots, 4,
\end{aligned}$$

cujas solução é o ponto $\tilde{x}^o = (0, 0, \frac{54}{7}, 0)$. Dessa forma,

$$\begin{aligned}
UB(u^o, W) &= \sum_{i=1}^2 u_i^o b_i + \sum_{j=1}^4 \tilde{x}_j (c_j - \sum_{i=1}^2 u_i^o a_{ij}) \\
&= 28\frac{5}{6} + \frac{54}{7} \times \frac{20}{48} \\
&\simeq 32,0476
\end{aligned}$$



Por (3.16), temos:

$$\nabla_{u_i} UB(u^o, W) = b_i - \sum_{j=1}^4 a_{ij} \tilde{x}_j^o, \quad i = 1, 2,$$

e, portanto,

$$\nabla_u UB(u^o, W) = \begin{pmatrix} 54 - 7 \times \frac{54}{7} \\ 10 - \frac{54}{7} \end{pmatrix} = \begin{pmatrix} 0 \\ \frac{16}{7} \end{pmatrix}$$

Daí, se fizermos $u(\rho) = u^o - \rho \nabla_u UB(u^o, W)$, de modo que $u(\rho) \geq 0$, então, aplicando a Busca de Fibonacci ao problema $\min_{\rho \geq 0, u(\rho) \geq 0} \{UB(u(\rho), W)\}$, obtemos $\rho^* = 0.009113$ e $z^* \leq UB(u(\rho^*), W) = 32$. Observe que é impossível melhorar esse limite com mais iterações do algoritmo, pois o limite obtido é igual ao valor ótimo da função objetivo do problema original. Entretanto, essa situação nem sempre acontece na prática. Nesses casos, podemos tomar $u^1 = u(\rho^*)$ e reiniciar o algoritmo até que o valor de $UB(u(\rho^*), W)$ permaneça inalterado.

Considerando, agora, o conjunto W definido pela segunda restrição do problema (5.65), o problema (3.14) se torna:

$$\begin{aligned} \max \quad & \frac{11}{48}x_1 - \frac{11}{48}x_2 + \frac{20}{48}x_3 - \frac{20}{48}x_4 \\ \text{s.a.} \quad & x_1 + 2x_2 + x_3 + 2x_4 \leq 10 \\ & x_i \geq 0, \quad i = 1, \dots, 4, \end{aligned}$$

cujas solução é o ponto $\tilde{x}^o = (0, 0, 10, 0)$. Com isso,

$$\begin{aligned} UB(u^o, W) &= \sum_{i=1}^2 u_i^o b_i + \sum_{j=1}^4 \tilde{x}_j^o (c_j - \sum_{i=1}^2 u_i^o a_{ij}) \\ &= 28 \frac{5}{6} + 10 \times \frac{20}{48} \\ &= 33 \end{aligned}$$

Por (3.16), temos:

$$\nabla_{u_i} UB(u^o, W) = b_i - \sum_{j=1}^4 a_{ij} \tilde{x}_j^o, \quad i = 1, 2,$$

e, portanto,

$$\nabla_u UB(u^o, W) = \begin{pmatrix} 54 - 70 \\ 10 - 10 \end{pmatrix} = \begin{pmatrix} -16 \\ 0 \end{pmatrix}$$

Daí, fazendo $u(\rho) = u^o - \rho \nabla_u UB(u^o, W)$, de modo que $u(\rho) \geq 0$, então, aplicando a Busca de Fibonacci ao problema $\min_{\rho \geq 0, u(\rho) \geq 0} \{UB(u(\rho), W)\}$, obtemos $\rho^* = 0.00372$, o que dá $z^* \leq UB(u(\rho^*), W) \simeq 32,04762$. Podemos reiniciar o algoritmo com $u^1 = u(\rho^*)$ até que o valor de $UB(u(\rho^*), W)$ permaneça inalterado, ou por um número fixo de iterações.

Considere, agora, o problema original escrito no formato de igualdade:

$$\begin{aligned} \max \quad & 2.5x_1 + 3x_2 + 4x_3 + 5x_4 \\ \text{s.a.} \quad & 4x_1 + 5x_2 + 7x_3 + 10x_4 + x_5 = 54 \\ & x_1 + 2x_2 + x_3 + 2x_4 + x_6 = 10 \\ & x_i \geq 0, \quad i = 1, \dots, 6 \end{aligned} \quad (5.67)$$

O dual de (5.67) é escrito como:

$$\begin{aligned} \min \quad & 54\hat{u}_1 + 10\hat{u}_2 \\ \text{s.a.} \quad & 4\hat{u}_1 + \hat{u}_2 \geq 2.5 \\ & 5\hat{u}_1 + 2\hat{u}_2 \geq 3 \\ & 7\hat{u}_1 + \hat{u}_2 \geq 4 \\ & 10\hat{u}_1 + 2\hat{u}_2 \geq 2 \\ & \hat{u}_1 \geq 0, \hat{u}_2 \geq 0 \end{aligned} \quad (5.68)$$

Agregando-se as variáveis de (5.68) em um único *cluster*, conforme sugerido na observação (4.1), e tomando-se pesos iguais a $\frac{1}{2}$ para cada variável, obtemos o seguinte problema agregado:

$$\begin{aligned} \min \quad & 32\hat{U} \\ \text{s.a.} \quad & \frac{5}{2}\hat{U} \geq 2.5 \\ & \frac{7}{2}\hat{U} \geq 3 \\ & 4\hat{U} \geq 4 \\ & 6\hat{U} \geq 5 \\ & \hat{U} \geq 0 \end{aligned} \quad (5.69)$$

cuja solução é, obviamente, $\hat{U}^* = 1$.

Desagregando a solução ótima de (5.69) segundo a proposição (2.1), obtemos $\hat{u} = (\frac{1}{2}, \frac{1}{2})$, que é factível para (5.68).

Computando o vetor $\hat{s}$ por $\hat{s} = \hat{A}^t \hat{u} - \hat{c}$, obtemos $\hat{s} = (0, \frac{1}{2}, 0, 1, \frac{1}{2}, \frac{1}{2})^t$. Como o vetor $\hat{s}$ possui algumas componentes nulas, vamos alterar a primeira componente de $\hat{u}$ para $\frac{3}{5}$, ou seja, vamos tomar $\hat{u} = (\frac{3}{5}, \frac{1}{2})$. Daí, $\hat{s} = \hat{A}^t \hat{u} - \hat{c} = (\frac{2}{5}, 1, \frac{7}{10}, 2, \frac{3}{5}, \frac{1}{2})^t$. Observe que, agora,



o vetor $\hat{s}$ possui todas as componentes positivas, e podemos, dessa forma, obter o elipsóide tipo Choi-Goldfarb descrito no teorema (4.1), ou seja,

$$E = \{x \in \mathcal{R}^6 / \|\hat{S}(x - \hat{x})\|^2 \leq \varepsilon^2\}$$

onde

$$\hat{S} = \begin{pmatrix} 2/5 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 7/10 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3/5 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1/2 \end{pmatrix},$$

$$\hat{x} = (\bar{x}, A\bar{x} - b) = \left(\frac{1}{3}, \frac{1}{3}, 3, 3, 0, 0\right), \quad e \quad \varepsilon = \left[(\hat{x}^t \hat{s})^2 + \|\hat{S}\hat{x}\|^2\right]^{\frac{1}{2}} \simeq (73.39 + 40.54)^{\frac{1}{2}} \simeq 10.67.$$

Daí, computando-se o limite (5.8) com esses parâmetros elipsoidais, e tomando $\bar{u} = U^*$, obtemos:

$$\begin{aligned} z^* &\leq \bar{u}^t b + \varepsilon \left[\sum_{j=1}^n (\delta_j / \tilde{s}_j)^2 \right]^{\frac{1}{2}} \\ &\simeq 28\frac{5}{6} + 10.67 \times 0.79 \\ &\simeq 38.30 \end{aligned}$$

Note que esse limite é bastante impreciso em comparação com os demais já calculados. Podemos, entretanto, melhorá-lo utilizando algumas iterações do procedimento descrito no final da seção 5.2.1, e que consiste em considerar o seguinte problema:

$$\begin{aligned} \max \quad & \frac{11}{48}x_1 - \frac{11}{48}x_2 + \frac{20}{48}x_3 - \frac{20}{48}x_4 \\ \text{s.a.} \quad & 2.5x_1 + 3x_2 + 4x_3 + 5x_4 \leq UB \\ & x_i \geq 0, \quad i = 1, \dots, 4, \end{aligned} \tag{5.70}$$

Tomando-se, inicialmente, $UB = 38.30$, então, a solução de (5.70) é o ponto $(0, 0, 9.575, 0)$, o que dá um objetivo ótimo de 3.99. Com isso, por (3.3), temos $z^* \leq 28\frac{5}{6} + 3.99 \simeq 32.82$. Substituindo esse novo limite no problema (5.70), e computando a sua solução, obtemos um objetivo ótimo igual a 3.42, o que, novamente por (3.3), dá $z^* \leq 32.25$.

Esse processo pode continuar até que a função objetivo ótima de (5.70) fique inalterada, ou por um número fixo de iterações.



Observe que com essas duas iterações, obtivemos uma melhora significativa no limite calculado com utilização dos parâmetros elipsoidais. Ressaltamos, entretanto, que essa situação nem sempre se verifica na prática, ou seja, nem sempre é possível obter uma maior precisão para um limite superior qualquer utilizando o processo sugerido na seção 5.2.1.

Como o centro $\hat{x}$ do elipsóide que construímos possui algumas componentes nulas, então o resultado da proposição (4.1) não produzirá nenhuma redução no seu raio. Verifiquemos se o mesmo ocorre quando consideramos a proposição (4.2).

Para isso, suponha que a restrição $\sum_{j=1}^6 d_j x_j = p$ foi obtida da soma das restrições do problema (5.67), ou seja, considere W definido como:

$$W = \{x \in \mathcal{R}^6 / 5x_1 + 7x_2 + 8x_3 + 12x_4 + x_5 + x_6 = 64, x_j \geq 0, j = 1, \dots, 6\}$$

Então, de acordo com a proposição (4.2), temos:

$$\begin{aligned} \tilde{\varepsilon}^2(W) &= \|\hat{S}\hat{x}\|^2 + p^2 \max_{i=1, \dots, m+n} \left\{ \left(\frac{\hat{s}_i}{d_i} \right)^2 \left[1 - 2 \frac{d_i \hat{x}_i}{p} \right] \right\} \\ &\simeq 40.54 + (64)^2 \times \max \{0.006, 0.019, 0.002, -0.003, 0.36, 0.25\} \\ &\simeq 40.54 + 4096 \times 0.36 \\ &\simeq 1515.10 \end{aligned}$$

Portanto, $\tilde{\varepsilon}(W) = 38.92$ e, como $\hat{x}$ possui algumas componentes nulas, então o raio $\hat{\varepsilon}(W)$ coincide com o raio calculado para o elipsóide proposto no teorema (4.1), ou seja, $\hat{\varepsilon}(W) = 10.67$.

Tem-se, portanto, $\varepsilon(W) = \min \{38.92, 10.67\} = 10.67$, de modo que esse processo também não resultou em redução no raio do elipsóide tipo Choi-Goldfarb, e, então, o limite (5.7), calculado com esses parâmetros, se confunde com o limite que já havíamos obtido.

Considere, agora, a obtenção do elipsóide descrito na proposição (4.4).

Para tanto, tomaremos $LB = 28\frac{5}{6}$ e $UB = 38.30$, ou seja, tomaremos UB como o limite superior calculado com a utilização dos parâmetros do elipsóide tipo Choi-Goldfarb. Note, entretanto, que qualquer outro valor para UB , conhecido de antemão, pode ser utilizado nesse processo.

Suponha, ainda, $\hat{A} = [4 \ 5 \ 7 \ 10 \ 1 \ 0]$ e $b = 54$, ou seja, suponha que estamos considerando somente a primeira linha da matriz de restrições do problema (5.67), para facilitar os cálculos.

Então, de acordo com a proposição (4.4),

$$d_1 = (\hat{A}\hat{S}^{-2}\hat{A}^t)^{-1}b = \left(\frac{2275}{9}\right)^{-1} \times 54 = \frac{486}{2275}$$



e

$$d_2 = -(\hat{A}\hat{S}^{-2}\hat{A}^t)^{-1}(\hat{A}\hat{S})^{-1}e = -\left(\frac{2275}{9}\right)^{-1} \times \frac{95}{3} = -\frac{57}{455}$$

Observe que $-b^t d_2 = 54 \times \left(\frac{57}{455}\right) = \frac{3078}{455}$. Além disso, como $b^t \hat{u} = [54 \ 10] \begin{bmatrix} 3/5 \\ 1/2 \end{bmatrix} = \frac{187}{5}$ (observe que aqui devemos tomar o vetor b completo, como exige a proposição), então, temos $b^t \hat{u} - UB \leq -b^t d_2 \leq b^t \hat{u} - LB$, de modo que recaímos no caso (a), ou seja,

$$x_o = \hat{S}^{-2} \hat{A}^t (\hat{A} \hat{S}^{-2} \hat{A}^t)^{-1} b \simeq \begin{bmatrix} 5.341 \\ 1.068 \\ 3.052 \\ 0.534 \\ 0.593 \\ 0 \end{bmatrix}$$

e

$$\varepsilon^2(W) = (\hat{s}^t \hat{x})^2 - b^t (\hat{A} \hat{S}^{-2} \hat{A}^t)^{-1} b \simeq 73.39 - 54 \times \frac{9}{2275} \times 54 \simeq 61.85,$$

de modo que $\varepsilon(W) = 7.86$, o que representa um melhoramento significativo sobre o raio do elipsóide tipo Choi-Goldfarb.

Aplicando-se, agora, esses novos parâmetros elipsoidais à expressão (5.7), obtemos:

$$\begin{aligned} z^* &\leq \bar{u}^t b + \sum_{j=1}^n \delta_j \tilde{x}_j + \varepsilon \left[\sum_{j=1}^n (\delta_j / \tilde{s}_j)^2 \right]^{\frac{1}{2}} \\ &\simeq 28 \frac{5}{6} + 2.04 + 7.86 \times 0.79 \\ &\simeq 37.08 \end{aligned}$$

o que dá um limite melhor que o que havíamos obtido com o elipsóide tipo Choi-Goldfarb, mas que ainda é muito impreciso em comparação com os demais já calculados.

Podemos, entretanto, aplicar o processo iterativo da seção 5.2.1 para melhorar a sua precisão, conforme já ilustrado.

Considere, agora, a segunda linha da matriz de restrições do problema (5.67), ou seja, tome $\hat{A} = [1 \ 2 \ 1 \ 2 \ 0 \ 1]$ e $b = 10$. Então:

$$d_1 = (\hat{A} \hat{S}^{-2} \hat{A}^t)^{-1} b = \left(\frac{3389}{196}\right)^{-1} \times 10 = \frac{1960}{3389}$$

e

$$d_2 = -(\hat{A}\hat{S}^{-2}\hat{A}^t)^{-1}(\hat{A}\hat{S})^{-1}e = -\left(\frac{3389}{196}\right)^{-1} \times \frac{125}{14} = -\frac{12250}{23723}$$

Observe que $-b^t d_2 = \frac{122500}{23723}$, e, da mesma forma que o anterior, temos $b^t \hat{u} - UB \leq -b^t d_2 \leq b^t \hat{u} - LB$, de modo que recaímos, novamente, no caso (a), ou seja,

$$x_o = \hat{S}^{-2} \hat{A}^t (\hat{A} \hat{S}^{-2} \hat{A}^t)^{-1} b \simeq \begin{bmatrix} 0.313 \\ 0.498 \\ 1.018 \\ 0.125 \\ 1.385 \\ 1.995 \end{bmatrix}$$

e

$$\varepsilon^2(W) = (\hat{s}^t \hat{x})^2 - b^t (\hat{A} \hat{S}^{-2} \hat{A}^t)^{-1} b \simeq 73.39 - 10 \times \frac{169}{3389} \times 10 \simeq 68.40,$$

de modo que $\varepsilon(W) = 8.27$, que é pior que o raio calculado utilizando a primeira restrição do problema (5.67). Entretanto, utilizando esses parâmetros na expressão (5.7), obtemos:

$$\begin{aligned} z^* &\leq \bar{u}^t b + \sum_{j=1}^n \delta_j \tilde{x}_j + \varepsilon \left[\sum_{j=1}^n (\delta_j / \tilde{s}_j)^2 \right]^{\frac{1}{2}} \\ &\simeq 28 \frac{5}{6} + 0.33 + 8.27 \times 0.79 \\ &\simeq 35.70 \end{aligned}$$

que é mais preciso que o limite obtido no caso anterior, mas ainda bastante impreciso em comparação com os demais já obtidos. Entretanto, a utilização do processo sugerido ao final da seção 5.2.1 produz melhoramento significativo nesse limite, conforme já ilustrado no caso do elipsóide Choi-Goldfarb.

Evidentemente, poderíamos utilizar todas as restrições do problema (5.67), incorrendo-se, entretanto, num maior custo computacional, resultante do cálculo da inversa da matriz $\hat{A}\hat{S}^{-2}\hat{A}^t$.

5.4 Estudo Numérico

Nesta seção, consideramos os resultados de um pequeno estudo numérico, feito com o intuito de avaliar a performance das localizações elipsoidais aplicadas à análise de erro devido à agregação de variáveis em Programação Linear.

Observamos, entretanto, que não pretendemos obter conclusões genéricas a respeito do comportamento do erro agregacional calculado com a utilização de elipsóides, devido à pequena quantidade de problemas considerados, bem como a sua dimensão reduzida.

No experimento realizado, consideramos 50 problemas do tipo (2.1), com $n = 100$, divididos em 5 grupos de 10 problemas com $m = 5, 10, 15, 20, 25$ restrições.

Os problemas considerados foram gerados aleatoriamente, com coeficientes $a_{ij}, c_j, b_i > 0$ e de tal modo que a matriz de restrições e o vetor c assumiram a seguinte estrutura:

$$A = [N^1 \ N^2 \ N^3 \ N^4 \ N^5] \text{ e } c = [c^1 \ c^2 \ c^3 \ c^4 \ c^5]$$

onde $N^l \in \mathcal{R}^{m \times 20}$ e $c^l \in \mathcal{R}^{20}$, $l = 1, \dots, 5$, tal que

$$0 \leq |n_{ij}^l - n_{ik}^l| \leq 10, \quad \forall n_{ij}^l, n_{ik}^l \in N^l, \quad l = 1, \dots, 5 \quad (5.71)$$

e

$$0 \leq |c_{ij}^l - c_{ik}^l| \leq 10, \quad \forall c_{ij}^l, c_{ik}^l \in c^l, \quad l = 1, \dots, 5 \quad (5.72)$$

Observamos que essa estrutura não restringe, de modo algum, a aplicabilidade dos métodos de cálculo de limites desenvolvidos, uma vez que qualquer problema pode ser disposto em estrutura semelhante à utilizada, bastando, para isso, construir as submatrizes N^l e os subvetores c^l através da aplicação de (5.71) e (5.72) ao problema original.

Não estamos, entretanto, propondo os esquemas (5.71) e (5.72) como critério de formação de *clusters*, uma vez que existem outros critérios muito mais adequados, como, por exemplo, o algoritmo proposto na seção 2.1, com a utilização da métrica (2.3) (vide experimentos numéricos de [11]). Ao invés disso, interpretamos (5.71) e (5.72) simplesmente como o critério utilizado para gerar os problemas considerados neste experimento.

As variáveis de cada problema foram agrupadas em 5 *clusters* com 20 variáveis cada um, de modo que, de cada problema original, obtivemos um problema agregado de dimensão $m \times 5$.

Ao l -ésimo *cluster* corresponderam as 20 variáveis da l -ésima submatriz N^l , e essas variáveis foram agregadas tomando-se pesos iguais a $1/20$ para cada variável do *cluster*.

No que segue, z_j^* indicará o valor ótimo da função objetivo do j -ésimo problema original considerado no experimento.

Cada um dos problemas agregados foi resolvido, e tomamos o seu objetivo ótimo como um limite inferior para o respectivo problema original, de acordo com (2.10).

A coluna correspondente aos Limites Inferiores na Tabela Resumo, a seguir, apresenta, para cada $m = 5, 10, 15, 20, 25$, a média e o desvio-padrão dos valores percentuais dos limites



inferiores obtidos com esse processo, ou seja, para cada $m = 5, 10, 15, 20, 25$, a coluna relativa aos Limites Inferiores apresenta os valores:

$$M = 100 \times \left(\sum_{j=1}^{10} \frac{z_j^* - Z_j^*}{z_j^*} \right) / 10 \text{ e } DP = \left(\frac{\sum_{j=1}^n \left(\frac{z_j^* - Z_j^*}{z_j^*} - M \right)^2}{10} \right)^{1/2}$$

onde Z_j^* indica o valor ótimo da função objetivo do j -ésimo problema agregado, para cada $m = 5, 10, 15, 20, 25$.

Observamos que não computamos os tempos de obtenção desses limites inferiores porque estamos interessados somente na comparação dos métodos de cálculo de limites superiores.

As colunas correspondentes aos Limites Superiores, na Tabela Resumo, a seguir, apresentam, para cada $m = 5, 10, 15, 20, 25$, os valores:

$$M = 100 \times \left(\sum_{j=1}^{10} \frac{UB_j - z_j^*}{z_j^*} \right) / 10, \quad DP = \left(\frac{\sum_{j=1}^n \left(\frac{UB_j - z_j^*}{z_j^*} - M \right)^2}{10} \right)^{1/2} \text{ e } T = \left(\frac{\sum_{j=1}^{10} t_j}{10} \right) \quad (5.73)$$

correspondentes à média, desvio-padrão e tempo médio de cálculo dos limites UB_j , segundo as seguintes abordagens:

1. Limite de Zipkin - Para calcular o limite (2.17), proposto por Zipkin [1], consideramos $\tilde{\sigma} = \{\{1, 2, \dots, n\}\}$ e tomamos $\sum_{j \in \tilde{\sigma}_K} d_j x_j \leq p_k$ como cada uma das restrições dos problemas originais, de modo que foram calculados m limites do tipo (2.17) para cada problema.

A coluna correspondente ao Limite I na Tabela Resumo a seguir apresenta os valores (5.73), onde UB_j corresponde à média dos m limites de Zipkin, calculados para cada uma das restrições do problema j , $j = 1, \dots, 10$ e para cada $m = 5, 10, 15, 20, 25$, e t_j é o tempo (em segundos) gasto na obtenção de UB_j .

2. Limites de Mendelsohn e Taylor/Shetty - Para calcular os limites (2.18) e (2.21), propostos por Mendelsohn[4, 5] e Taylor/Shetty[9, 5], respectivamente, analogamente ao item anterior, consideramos $\tilde{\sigma} = \{\{1, 2, \dots, n\}\}$ e tomamos $\sum_{j \in \tilde{\sigma}_K} d_j x_j \leq p_k$ como cada uma das restrições dos problemas originais, de modo que foram calculados m limites do tipo (2.18) e (2.21) para cada problema e para cada valor de m . Para minimizar as expressões (2.18) e (2.21), utilizamos a Busca de Fibonacci[17]. Um intervalo inicial

para o processo de busca foi obtido atribuindo-se valores aleatórios para θ e Φ , até se obterem os intervalos I e $\tilde{I}$, de modo que $\theta^* \in I$ e $\Phi^* \in \tilde{I}$.

As colunas correspondentes aos Limites II e III na Tabela Resumo, a seguir, apresentam, os valores (5.73), onde UB_j corresponde à média dos m limites de Mendelsohn e Taylor/Shetty, calculados para cada uma das restrições do problema j , $j = 1, \dots, 10$ e para cada $m = 5, 10, 15, 20, 25$, e t_j é o tempo (em segundos) gasto na obtenção de UB_j .

3. Limite obtido com uma iteração do algoritmo proposto na seção 3.2. - Analogamente ao cálculo dos limites do item anterior, tomamos o conjunto W como cada uma das restrições dos problemas originais.

O tamanho do passo ρ^* foi obtido similarmente aos métodos do item anterior, com a utilização da Busca de Fibonacci.

A coluna correspondente ao Limite IV na Tabela Resumo, a seguir apresenta, os valores (5.73), onde UB_j corresponde à média dos m limites computados de acordo com o algoritmo proposto na seção 3.2 para cada uma das restrições do problema j , $j = 1, \dots, 10$ e para cada $m = 5, 10, 15, 20, 25$, e t_j é o tempo (em segundos) gasto na obtenção de UB_j .

4. Limite sob Localização Elipsoidal (1) - Para calcular o limite (5.7), consideramos o elipsóide descrito no teorema (4.1).

As soluções $\hat{x}$ e $(\hat{u}, \hat{s})$ requeridas para construir o elipsóide, foram obtidas de acordo com o item 1 da observação (4.1).

Para obter $(\hat{u}, \hat{s})$, o problema (4.8) foi agregado de modo que todas as variáveis ocupassem um único *cluster*, resultando, assim, em um problema de fácil resolução.

A coluna correspondente ao Limite V na Tabela Resumo, a seguir, apresenta os valores (5.73), onde UB_j corresponde ao limite (5.7), calculado para cada problema j , $j = 1, \dots, 10$ e para cada $m = 5, 10, 15, 20, 25$, com os parâmetros elipsoidais obtidos conforme acima, e refinados de acordo com o processo proposto ao final da seção 5.2.1, tomando-se $l_j = 0$ e $L_j = \infty$, e t_j é o tempo (em segundos) gasto na obtenção de UB_j , aí incluído o tempo gasto para a obtenção da localização elipsoidal.

5. Limite sob Localização Elipsoidal (2) - Para calcular o limite (5.7), consideramos o elipsóide descrito na proposição (4.2), onde os vetores $\hat{x}$ e $(\hat{u}, \hat{s})$ foram tomados iguais



ao item anterior.

A localização $W = \{x \in \mathcal{R}^{m+n} / \sum_{j=1}^{m+n} d_j x_j = p, d_j, p > 0, x_j \geq 0\}$ foi obtida somando-se todas as restrições dos problemas originais, no formato (4.4).

A coluna correspondente ao Limite VI, na Tabela Resumo, apresenta os valores (5.73), onde UB_j corresponde ao limite (5.7), calculado de acordo com a abordagem acima para cada problema j , $j = 1, \dots, 10$ e para cada $m = 5, 10, 15, 20, 25$, e refinado de acordo com o processo proposto ao final da seção 5.2.1, tomando-se $l_j = 0$ e $L_j = \infty$, e t_j é o tempo (em segundos) gasto na obtenção de UB_j , incluído o tempo para construção dos elipsóides.

6. Limite sob Localização Elipsoidal(3) - Tomamos o elipsóide definido na proposição (4.4) no cálculo do limite (5.7). Os valores LB e UB , requeridos por essa proposição, foram tomados, respectivamente, como sendo o valor ótimo da função objetivo do problema agregado e o limite obtido no item anterior, para cada problema.

O centro x_o do elipsóide foi calculado tomando-se, em separado, cada uma das restrições do problema original, de modo que, para cada problema e para cada $m = 5, 10, 15, 20, 25$, foram obtidas m localizações elipsoidais do tipo descrito no teorema (4.3), com centro e raio definidos pela proposição (4.4).

Em seguida, foram calculados m limites do tipo (5.7), utilizando-se os elipsóides assim obtidos, e refinados de acordo com o processo proposto ao final da seção 5.2.1, tomando-se $l_j = 0$ e $L_j = \infty$.

A coluna correspondente ao Limite VII, na Tabela Resumo, a seguir, apresenta, os valores (5.73), onde UB_j corresponde ao limite (5.7), calculado de acordo com a abordagem acima para cada problema j , $j = 1, \dots, 10$ e para cada $m = 5, 10, 15, 20, 25$, e t_j é o tempo (em segundos) gasto no cálculo de UB_j , incluindo o tempo de obtenção dos elipsóides.

Observamos, por fim, que tanto na computação da expressão (3.3), quanto na computação da expressão (5.7), utilizou-se $\bar{u} = U^*$, ou seja, tomou-se $\bar{u}$ igual à solução ótima do problema (2.7).

A seguir, apresentamos a tabela contendo o resumo dos resultados do teste.



TABELA RESUMO - Limites em % (M=média, DP=desvio-padrão, T= tempo (em segundos))

Restrições	Índices	Limites Inferiores	Limites Superiores						
			I	II	III	IV	V	VI	VII
m = 5	M	90.63	131.43	1.77	37.62	11.62	3.07	3.07	3.07
	DP	0.03	3.11	0.03	0.66	0.22	0.06	0.06	0.06
	T	*	0.04	1.77	4.09	3.57	0.45	0.70	8.51
m = 10	M	84.05	221.14	1.75	103.52	48.47	2.18	2.18	2.18
	DP	0.09	2.36	0.03	1.22	0.68	0.04	0.04	0.04
	T	*	0.04	1.78	6.62	5.41	0.70	0.75	9.33
m = 15	M	77.95	1066.51	4.72	407.71	151.58	5.11	5.11	5.11
	DP	0.16	15.09	0.07	6.55	2.39	0.08	0.08	0.08
	T	*	0.04	1.82	10.04	7.46	1.00	1.40	10.43
m = 20	M	85.68	328.47	0.68	118.26	52.15	0.72	0.72	0.72
	DP	0.13	3.07	0.02	1.08	0.43	0.02	0.02	0.02
	T	*	0.04	1.81	12.60	9.47	1.35	1.85	11.75
m = 25	M	85.87	422.76	2.39	139.00	61.62	2.54	2.54	2.54
	DP	0.06	5.96	0.05	1.18	0.56	0.06	0.06	0.06
	T	*	0.04	1.79	13.82	11.34	1.66	2.25	12.70

- Limite I - Média dos limites de Zipkin, computados por (2.17)
- Limite II - Média dos limites de Mendelsohn, computados por (2.18)
- Limite III - Média dos limites de Taylor/Shetty, computados por (2.21)
- Limite IV - Média dos limites computados com uma iteração do algoritmo da seção 3.2
- Limite V - Limite (5.7), computado com o elipsóide descrito no teorema (4.1)
- Limite VI - Limite (5.7), computado com o elipsóide descrito na proposição (4.2)
- Limite VII - Limite (5.7), computado com o elipsóide descrito na proposição (4.4)

Resultados - Com base nos dados listados na Tabela Resumo acima, podemos obter as seguintes conclusões:

1. A estrutura que utilizamos para o vetor custo e a matriz de restrições dos problemas-testes, bem como a forma como o problema agregado foi construído, forneceram limites inferiores razoavelmente bons para z^* , conforme se pode constatar da coluna referente aos limites inferiores na tabela acima.
2. Os limites superiores dispostos na Tabela Resumo, ordenados pela sua precisão, para cada $m = 5, 10, 15, 20, 25$, verificam

$$II < V = VI = VII < IV < III < I$$



Se considerarmos, entretanto, o tempo de execução, o ordenamento se torna:

$$I < V < \begin{cases} VI < II < IV < III < VII, & \text{para } m = 5, 10, 15 \\ II < VI < IV < VII < III & \text{para } m = 20, 25 \end{cases}$$

Ressaltamos, entretanto, que se trata de um comportamento médio, que deve ser interpretado cuidadosamente, em virtude dos vários parâmetros que influenciam a análise do erro agregacional.

3. Com a coleção de problemas considerados, o Limite II mostrou-se o mais preciso entre todos os demais. Isto se deve ao fato de que o método de Mendelsohn propõe a busca por vetores duais na direção do vetor U^* , e, nos problemas considerados, os problemas agregados forneceram soluções duais razoavelmente boas para o processo de busca. Observamos, entretanto, que outros experimentos têm demonstrado que essa situação não se mantém quando a solução desagregada não é boa (vide, por exemplo, os experimentos de Leisten[5] e Litvinchev-Rangel[2]).
4. Os limites V, VI e VII também se apresentaram bastante precisos, e com a vantagem de que os limites V e VI são calculados com reduzido custo computacional, em comparação com os demais.

Observamos, novamente, que esse é um comportamento médio, e não deve ser generalizado, devido ao pequeno número de problemas analisados, bem como a sua dimensão reduzida. Acreditamos, entretanto, que esse comportamento venha a se manter em situações genéricas, sempre que o problema agregado fornecer soluções desagregadas razoavelmente boas, pois isso tende a reduzir o raio dos elipsóides que abordamos, o que influi consideravelmente na obtenção de limites mais precisos.

Podemos afirmar, desse modo que a utilização de elipsóides na análise do erro devido à agregação se apresenta como uma boa alternativa computacional.

5. Os limites gerados com as localizações elipsoidais se mostraram bastante superiores, em precisão, aos correspondentes limites de Zipkin e Taylor/Shetty e ao limite obtido da aplicação do algoritmo da seção 3.2. Observe que as localizações elipsoidais utilizadas no cálculo do limite VII foram geradas por processo semelhante ao do cálculo do limite de Zipkin, pois consideramos, no cálculo do centro e do raio dos elipsóides, somente uma restrição do problema original. Além disso, as localizações elipsoidais dos limites V e VI foram obtidas com menor custo computacional, pois só se necessitou dos vetores $\hat{x}$

e $(\hat{u}, \hat{s})$, factíveis para (4.4) e (4.5), respectivamente, que podem ser obtidos facilmente com a aplicação da metodologia da agregação.

Com esta série de problemas, obtivemos os mesmos limites para os três tipos de localizações elipsoidais analisadas. Isto mostra que as técnicas de redução do raio apresentadas no Capítulo 4 não influenciaram no limite final, obtido com o processo de refinamento sugerido ao final da seção 5.2.1, o que sugere que o esquema proposto no teorema (4.1) se apresenta como a melhor alternativa para cálculo de limites sob localização elipsoidal, devido ao seu reduzido custo computacional, em comparação com os demais métodos.

Devido ao seu alto custo computacional, resultante da manipulação de matrizes de projeção, acreditamos que o esquema proposto pelo teorema (4.3) seja mais aplicável às técnicas de pontos interiores, conforme proposta de Choi-Goldfarb[10]. Entretanto, voltamos a salientar que não podemos generalizar esta conclusão, devido ao contexto sob o qual o experimento foi realizado.

6. Verificamos, ainda, que é de fundamental importância a utilização de algum esquema de refinamento do limite obtido com localizações elipsoidais. Ressalte-se, porém, que esse esquema deve ser simples e barato computacionalmente, como o esquema iterativo que propusemos ao final da seção 5.2.1 e que utilizamos com sucesso neste experimento. Observamos, entretanto, que o método sugerido nem sempre garante decréscimo do limite inicial e depende muito da qualidade do vetor U^* utilizado na obtenção do vetor $\delta = c - U^{*t}A$.
7. Por fim, observamos que não queremos ressaltar aqui a superioridade de um método sobre o outro, mesmo porque isso é impossível, já que algum método pode não funcionar bem com problemas com certas estruturas, mas pode apresentar uma performance muito boa com problemas com estruturas particulares. É o caso, por exemplo, do Limite IV, que apresentou um resultado muito bom no experimento realizado em [2], com problemas com restrições do tipo $0 \leq x_j \leq 1$.

O capítulo a seguir tece as considerações finais a respeito do trabalho.



Capítulo 6

Conclusão

A técnica da agregação apresenta um campo muito vasto para pesquisa. São inúmeras as áreas que merecem estudo, com o intuito de aprimorar essa metodologia.

Pesquisas futuras deveriam investigar, por exemplo, novas técnicas de agrupamento dos dados e de determinação de pesos para formar o problema agregado, além de novos métodos para obter soluções desagregadas mais próximas à solução ótima do problema original, sem, entretanto, aumentar o custo computacional do processo.

Focalizamos, com esse trabalho, as metodologias de análise do erro devido à agregação de variáveis em Programação Linear.

A abordagem de Zipkin [1], discutida na seção 2.3, foi estendida, no Capítulo 3, para tratar os casos em que se conhece um conjunto localização genérico, fornecendo, assim, outras alternativas para análise do erro agregacional.

Uma das consequências dessa abordagem é a possibilidade de se utilizar localizações não-lineares na computação de limites superiores para z^* . Com esse intuito, discutimos, no Capítulo 4, as localizações não-lineares tipo elipsoidais, e apresentamos várias técnicas de derivação dessas localizações para o caso de Programação Linear.

A partir dos elipsóides tipo Choi-Goldfarb [10], apresentamos vários métodos de redução do raio dessas localizações elipsoidais, com o objetivo de se obter uma maior precisão para o limite calculado com a utilização dessas localizações.

Conforme demonstrado pelo estudo numérico da seção 5.3, a aplicação das localizações elipsoidais na análise do erro agregacional se mostrou uma boa alternativa para a obtenção de limites superiores para z^* .

Além de garantir a unicidade da solução do problema de maximização em (3.3) (vide seção 5.1), as localizações elipsoidais facilitam o cálculo da expressão (3.3), sem exigir um



alto custo computacional, o que é uma grande vantagem em relação às abordagens existentes na literatura (vide [4, 5, 2]).

Acreditamos que o esquema proposto para calcular limites com a utilização dos elipsóides funcionará muito bem sempre que for possível obter boas soluções desagregadas a partir da solução ótima do problema agregado. Isso pode ser conseguido, por exemplo, com a aplicação do algoritmo de construção de *clusters* proposto na seção 2.1, que reduz bastante a diferença $z^* - Z^*$, conforme demonstrou o experimento numérico realizado por Weintraub[11].

Deixamos em aberto, entretanto, vários tópicos para pesquisas futuras, como por exemplo, a derivação de outros tipos de elipsóides, como uma alternativa aos elipsóides (4.2), que focalizamos neste texto, bem como a exploração dos elipsóides duais desenvolvidos no Capítulo 4, e a investigação de estruturas para a sua aplicação na análise do erro agregacional, além do desenvolvimento de novas técnicas para tornar mais eficiente a aplicação dos elipsóides na metodologia da agregação.

É interessante, também, investir em pesquisas que visem a redução da distância entre a solução desagregada e a solução ótima do problema original, conforme já vem sendo proposto, por exemplo, pelos esquemas de agregação/desagregação iterativa, que produzem a convergência da solução desagregada para a solução ótima do problema original (vide abordagem de [23]). Ter pontos factíveis próximos à solução ótima dos problemas (4.4) e (4.5) é importante porque, devido às relações de complementaridade, teríamos melhores valores para os raios dos elipsóides que abordamos.

Apesar de não podermos obter conclusões genéricas a respeito do comportamento do erro devido à agregação, já que uma grande variedade de parâmetros influem no processo agregacional, como, por exemplo, a dimensão dos problemas, os algoritmos de formação do problema agregado e os métodos de desagregação, acreditamos que as localizações elipsoidais se mostram uma boa opção para o cálculo de limites superiores, e merecem o enfoque de pesquisas futuras para tornar cada vez mais precisos os limites obtidos com sua utilização.



Bibliografia

- [1] P.H.Zipkin, *Bounds on the effect of aggregating variables in Linear Programming*, Ops.Res., 1980, vol. 28, pp 403-418.
- [2] I.S.Litvinchev and S.Rangel, *Localization of optimal solution and a posteriori bounds for aggregation*, Computers Ops. Res., 1999, V26, N10-11, pp 967-988.
- [3] I.S.Litvinchev, *Decomposition-Aggregation method for convex programming problems*, Optimization, 1991, vol 22, pp 47-56.
- [4] D.F.Rogers, R.D.Plante, R.T.Wong and J.R.Evans, *Aggregation and Disaggregation techniques and methodology in optimization*, Ops. Res. 39, 1991, pp 553-582.
- [5] R.Leisten, *A posteriori error bounds in Linear Programming aggregation*, Computers Ops. Res., 1997, vol 24, pp 1-16.
- [6] I.S.Litvinchev, *Bounds on the suboptimization of aggregation in convex programming*, Comp. Maths. Math. Phys. 33, 1993, pp 1007-1015.
- [7] L.S.Lasdon, *Optimization Theory for Large Systems*, Macmillan, London, 1973.
- [8] R. Mendelsohn, *Improved bounds for aggregated linear programs*, Ops. Res. 28, 1980, pp 1450-1453.
- [9] C.M.Shetty and R.W.Taylor, *Solving large-scale linear programs by aggregation*, Computers Ops. Res. 14, 1987, pp 385-393.
- [10] I.C.Choi and D.Goldfarb, *On solution containing ellipsoids in Linear Programming*, JOTA 80, 1994, pp 161-173.
- [11] A.Weintraub G.S. and M. Yadlin, *Aggregation procedures in forest manegement planning using cluster analysis*, Forest Science, 43(2), 1997, pp274-285.



- [12] I.S.Litvinchev, *On improved Goldfarb-Choi solution containing ellipsoids in Linear Programming*, Relatório Técnico TR98003-OC - DCCE - UNESP - São José do Rio Preto (SP) - 1998.
- [13] I.S.Litvinchev, *On solution-containing ellipsoids in convex optimization - Part I*, Relatório Técnico TR98004-OC- DCCE - UNESP - São José do Rio Preto (SP) - 1998.
- [14] M.J.Todd, *Improved bounds and containing ellipsoids in Karmarkar's Linear Programming Algorithm*, Maths. Oper. Res. 13, 1988, pp 650-659.
- [15] Y.Ye, *Recovering optimal basic variables in Karmarkar's polynomial algorithm for Linear Programming*, Maths. Oper. Res. 15, 1990, pp 564-572.
- [16] M.S.Bazarra, J.J.Jarvis and H.D.Sherali, *Linear Programming and Network Flows*, John Wiley & Sons, Inc., 2a. edição, 1990, 684 pp, ISBN 0-471-63681-9.
- [17] M.S.Bazaraa, H.D.Sherali and C.M.Shetty, *Nonlinear Programming*, John Wiley & Sons, Inc., 1993.
- [18] D.G.Luenberger, *Linear and Nonlinear Programming*, Addison-Wesley Publishing Company, 2a. edição, 1984, 491 pp, ISBN 0-201-15794-2.
- [19] B.Noble and J.W.Daniel, *Álgebra Linear Aplicada*, Prentice Hall do Brasil, 1986, ISBN 85-7054-022-1.
- [20] C.C.Gonzaga, *Algoritmos de Pontos Interiores para Programação Linear*, 17o. Colóquio de Matemática, IMPA, CNPq, 1989.
- [21] Alexander Schrijver, *Theory of Linear and Integer Programming*, 1a. edição, John Wiley and Sons, 1986, ISBN 0-471-90854-1.
- [22] Rui Eduardo Campelo e Nelson Maculan, *Algoritmos e Heurísticas*, Editora da Universidade Fluminense, Niterói(RJ), 227 pp, 1994.
- [23] K.Jörnsten, R.Leisten and S.Storøy, *Gradient Schemes in Iterative Aggregation Procedures for Variable-Aggregated LP-Problems*, Optimization, 1994, vol. 30, pp.251-268.



