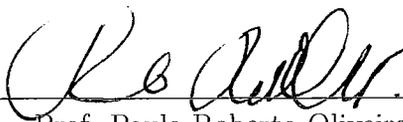


MODELAGEM E OTIMIZAÇÃO DE UM SISTEMA DE TELECOMUNICAÇÃO
SEM FIO E DE UMA CARTEIRA DE INVESTIMENTOS

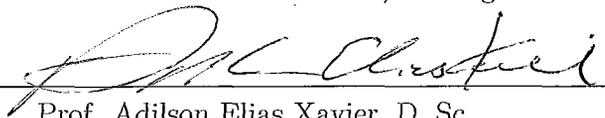
Rodrigo José Guerra Leone

TESE SUBMETIDA AO CORPO DOCENTE DA COORDENAÇÃO DOS
PROGRAMAS DE PÓS-GRADUAÇÃO DE ENGENHARIA DA
UNIVERSIDADE FEDERAL DO RIO DE JANEIRO COMO PARTE DOS
REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE DOUTOR
EM CIÊNCIAS EM ENGENHARIA DE SISTEMAS E COMPUTAÇÃO.

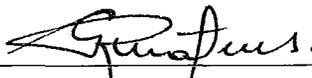
Aprovada por:



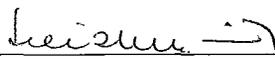
Prof. Paulo Roberto Oliveira, Dr. Ing.



Prof. Adilson Elias Xavier, D. Sc.



Prof. Geraldo Hobson Mateus, D. Sc.



Prof. Luiz Costa da Silva, PhD.



Prof. Roberto Quirino do Nascimento, D. Sc.



Prof. Sérgio Granville, PhD.

RIO DE JANEIRO, RJ - BRASIL

MARÇO DE 2004

LEONE, RODRIGO JOSÉ GUERRA

Modelagem e Otimização de um Sistema
de Telecomunicação sem fio e de uma Carteira
de Investimentos [Rio de Janeiro] 2004

XII, 191 p. 29,7 cm (COPPE/UFRJ,
D.Sc., Engenharia de Sistemas e Computação,
2004)

Tese – Universidade Federal do Rio de
Janeiro, COPPE

- 1 - Localização de ERBs
- 2 - Controle de Potência
- 3 - Carteira de Investimentos
- 4 - Programação Geométrica Signomial
- 5 - Decomposição Lagrangeana

I. COPPE/UFRJ II. Título (série)

À Andréa.

Agradecimentos

- Ao professor e amigo Paulo Roberto Oliveira, pelo apoio e confiança. Será sempre uma referência em minha vida profissional e pessoal.
- Ao professor Roberto Quirino do Nascimento, pela co-orientação e por apostar, acima de tudo, em minha capacidade desde o Mestrado.
- Aos professores Geraldo Robson Mateus, Adilson Xavier e Luiz Costa da Silva, pela atenção sempre que precisei.
- Aos professores Marcos Augusto dos Santos, Nelson Maculan Filho e Antônio Marcos Duarte Jr., pelas sugestões.
- Aos professores membros da banca examinadora.
- À Andréa, por aceitar minhas ausências e ansiedade.
- Ao Davi, simplesmente por existir.
- Aos meus pais e irmãos, por suprirem a Andréa e o Davi do carinho que nem sempre pude dar nesse período.
- Aos amigos da COPPE: Gilvan, Rosely, Geci, Verinha, Jurandir e Paulinho, por estarem sempre presentes.
- Ao amigo Sérgio Monteiro, pela ajuda nas implementações.
- Aos professores e funcionários do programa de Engenharia de Sistemas e Computação da COPPE, em especial à professora Regina Burachik, ao professor Luis Mauricio Graña Drummond, à Cláudia, Solange, Dona Gersina e Lourdes.

- Aos professores José Rodrigues Lustoza, Mamadu Lamarana Bari e Carlos Pedrosa Júnior da Fundação Visconde de Cairú, ao professor Carlos Coutinho Batalha da Faculdade de Ciências Humanas de Vitória e ao Prof. Josir Simeone Gomes do IBMEC - RJ, pela confiança e pelas oportunidades.
- À Dona Maria José e ao Herbert, pela acolhida em Belo Horizonte.
- Ao CNPq, pelo suporte financeiro.

*” ... por isso sigo caminhando, sempre em frente,
incansável, na certeza de ir até
onde Deus disser: É até aqui. ”*

José Rodrigues Lustoza

Resumo da Tese apresentada à COPPE/UFRJ como parte dos requisitos necessários para a obtenção do grau de Doutor em Ciências (D.Sc.)

MODELAGEM E OTIMIZAÇÃO DE UM SISTEMA DE TELECOMUNICAÇÃO SEM FIO E DE UMA CARTEIRA DE INVESTIMENTOS

Rodrigo José Guerra Leone

Março/2004

Orientador: Paulo Roberto Oliveira

Programa: Engenharia de Sistemas e Computação

Apresentamos dois modelos não lineares e inteiros: um para os problemas de localização de estações rádio base e de controle de potência em tecnologia CDMA e outro para o problema de otimização de uma carteira de investimentos, introduzindo uma mensuração do risco baseada em uma função utilidade. Propomos uma heurística de programação geométrica signomial aplicada a problemas 0-1, em que é introduzida uma função com a característica de transformar as soluções contínuas em soluções inteiras. Para o primeiro modelo ainda propomos uma decomposição lagrangeana, em que a resolução do problema se particiona na resolução de um problema mestre e de três subproblemas. Apresentamos os resultados computacionais obtidos para todas as metodologias referentes às simulações testadas para o modelo de telecomunicações.

Abstract of Thesis presented to COPPE/UFRJ as a partial fulfillment of the requirements for the degree of Doctor of Science (D.Sc.)

MODELING AND OPTIMIZING A NETWORK DESIGN AND PLANNING
FOR A WIRELESS SYSTEM AND AN INVESTMENT PORTFOLIO

Rodrigo José Guerra Leone

March/2004

Advisor: Paulo Roberto Oliveira

Department: Systems and Computation Engineering

We present two models: one for the network design and planning of a wireless system based on CDMA technology, in which the aim is to solve the base station location problem, the power control being considered, to meet the service demand. The other model considers the investment portfolio optimization, based on introducing the definition of risk-utility. The proposed models are nonlinear programming formulations and we propose a signomial geometric programming heuristic to obtain their optimal solutions. We also propose a Lagrangean decomposition for the first model. Computational results for some simulated problems are presented.

Índice

Introdução	1
1 Modelagem em Telecomunicação Sem Fio	5
1.1 Aspectos básicos de Sistemas de Telecomunicações Móveis Celular . . .	5
1.1.1 Conceitos Básicos	5
1.1.2 Sobre Modulação e Técnicas de Acesso	7
1.1.3 Alguns problemas e modelos existentes	8
1.2 Um novo modelo para Telecomunicação Sem Fio	20
1.2.1 Modelo Não-Linear Inteiro	21
1.2.2 Explicação do Modelo	22
1.2.3 Análise do Modelo	24
2 Modelagem em Carteira de Investimentos	26
2.1 Aspectos básicos sobre Utilidade, Risco e Retorno	26
2.1.1 Conceitos e Resultados	26
2.1.2 Modelagem e Gerenciamento do Risco	28
2.2 Proposta de Cálculo de Risco baseado em Utilidade	29
2.2.1 A Curva de Utilidade	30
2.2.2 A Distância-Utilidade e o Desvio-Utilidade	31

2.2.3	Resultados experimentais : σ versus DU	41
2.2.4	Correlação-Utilidade	45
2.3	Um novo modelo para Otimização de Carteira de Investimentos	47
2.3.1	Modelo Não-Linear Inteiro	48
2.3.2	Explicação do Modelo	49
2.3.3	Análise do Modelo	50
3	Metodologia de Programação Geométrica Signomial 0-1	52
3.1	Preliminares	52
3.2	Programação Geométrica Signomial	53
3.2.1	Função Objetivo Posinomial	54
3.2.2	Função Objetivo com Coeficientes Negativos	69
3.2.3	Técnica de Condensação	75
3.3	Referências bibliográficas	82
3.4	Metodologia de Otimização Contínua para o problema 0-1	83
3.5	Metodologia aplicada ao Problema de Telecomunicação Sem Fio	98
3.5.1	Técnica de Condensação	98
3.5.2	Heurística 0-1	103
3.5.3	Existência de Solução	111
3.6	Metodologia aplicada ao Problema de Otimização de Carteira de In- vestimentos	116
3.6.1	Técnica de Condensação	116
3.6.2	Heurística 0-1	125
3.6.3	Existência de Solução	127

4	Metodologia de Decomposição Lagrangeana	130
4.1	Preliminares	130
4.2	Referências bibliográficas	130
4.3	Decomposição Lagrangeana	133
4.3.1	Subproblema em z	136
4.3.2	Subproblema em x	136
4.3.3	Subproblema em (p, r, y)	140
4.3.4	O Problema Mestre	144
5	Um Algoritmo de Pontos Interiores Inviável	148
5.1	Algoritmo Barreira-Penalidade	148
5.1.1	A Direção de Newton perturbada	152
5.2	Método aplicado ao subproblema em (p, y, t, v)	158
6	Implementações e Resultados Computacionais	161
7	Proposta de Trabalhos Futuros	166
7.1	Quanto à modelagem	166
7.2	Quanto à implementação	167
7.3	Quanto à teoria	167
	Apêndice A	168
	Apêndice B	174
	Referências Bibliográficas	179

Introdução

Os projetos de instalação e expansão dos sistemas de comunicação sem fio requerem em geral grandes investimentos, gerando desafios técnicos significativos.

Um dos principais problemas para uma rede sem fio é onde e com que capacidade devem ser instaladas as estações rádio base (ERBs), de onde partem as transmissões. O objetivo é instalar um conjunto de estações capazes de cobrir a área desejada, atender à demanda, minimizando os custos e mantendo padrões de qualidade de serviço.

Um segundo problema se refere à alocação dos canais. O objetivo é cobrir toda a área de serviço, distribuindo as frequências de forma eficiente, atendendo à demanda e reduzindo as interferências. Nesse estudo, é necessário conhecer a tecnologia implantada nas antenas das estações. No caso da alocação de canais, as tecnologias mais usadas são a FDMA (*Frequency Division Multiplexing Access*) e a TDMA (*Time Division Multiplexing Access*) [52], [14], [89], [92].

Outro problema é o controle da potência transmitida no *link* ERB-usuário. Nesse caso, uma tecnologia utilizada é a CDMA (*Code Division Multiplexing Access*) [46], [25], [52], [14], [89], [92]. Através desse controle, podemos evitar, como veremos mais adiante, o problema da alocação de canais.

A maioria dos modelos matemáticos observa esses problemas isoladamente. Tal abordagem pode acarretar em perda de qualidade das soluções, bem como dos in-

vestimentos necessários em infra-estrutura. Existem modelos que tratam dos dois primeiros problemas em conjunto, sendo, alguns deles, de difícil implementação. No que diz respeito ao terceiro problema, não temos conhecimento de nenhum modelo que o trate conjuntamente com o problema da localização das estações rádio base, fazendo dessa lacuna uma das motivações de nosso trabalho.

A informação para a tomada de decisão sobre investimento em ativos só está completa quando são dados o retorno esperado, o risco associado a esse retorno, o intervalo de confiança para esse risco e a liquidez do investimento¹. Dessa forma, como é usual utilizarmos uma probabilidade de aproximadamente 95% para o intervalo [retorno - risco; retorno + risco], inferido por uma tendência normal de comportamento para a variável retorno, basta-nos informar o par risco/retorno. O desconhecimento desse fato pode acarretar em perdas muito além das expectativas do investidor, mas não necessariamente improváveis estatisticamente. Com os pares risco/retorno para vários ativos em mãos, quase nunca é possível ordená-los do melhor ao pior. A decisão depende do perfil do investidor: o mais agressivo opta por ativos de maior volatilidade, cujos retornos são geralmente mais altos, correndo o risco de ganhos ainda maiores ou de perdas catastróficas, enquanto que o mais conservador se satisfaz com retornos menores, de preferência com baixo risco de perda. É a lei da compensação.

No decorrer do trabalho, consideraremos válida a seguinte hipótese central sobre um investidor conservador: a decepção com a perda é mais sentida que a satisfação com o ganho. Exemplificando, um retorno de -1% quando se esperava 2% não é compensado por um retorno de 5% para a mesma expectativa. Oscilações negati-

¹A preocupação com a liquidez é mais acentuada quando o modelo abrange o longo prazo. Como, geralmente, os modelos avaliam o retorno no período imediatamente posterior ao investimento, não se faz necessário levá-la em consideração.

vas são mais influentes que oscilações positivas e, portanto, devem assumir maior ponderação.

O cálculo do risco pela variância (média da soma dos quadrados dos desvios em relação à média) tem a propriedade de penalizar variações superiores a uma unidade, porém não diferencia variações negativas de variações positivas. Assim sendo, uma variação de 3% além do esperado tem peso nove vezes maior do que uma variação positiva de 1% . O mesmo continuaria válido, caso as variações fossem negativas. Mais ainda, o peso da variação positiva de 3% é idêntico ao da variação negativa de 3%. Já para o desvio-médio, os pesos são proporcionais aos valores absolutos das variações.

Os modelos de otimização de carteira de investimentos a partir de um conjunto de ativos, geralmente, consideram o risco definido pelo desvio-padrão (raiz quadrada de variância) e, portanto, sofrem a mesma consequência apontada acima. Em nosso trabalho, propomos uma medida de risco que leva em conta a hipótese central apresentada anteriormente.

Nosso trabalho está dividido em seis capítulos. O primeiro capítulo trata da modelagem em telecomunicação sem fio: a seção 1.1 apresenta os aspectos básicos da telecomunicação móvel celular, expondo os conceitos e as tecnologias necessários para o entendimento da modelagem. A seção ainda aborda alguns problemas mais frequentes em comunicação móvel e apresenta modelos referentes a esses problemas. A seção 1.2 apresenta um novo modelo não-linear inteiro para telecomunicação sem fio, visando o suprimento da lacuna exposta acima. Na verdade, trata-se de um problema binário e fracionário-linear devido ao conjunto de restrições de qualidade de serviço (QoS). O capítulo 2 trata da modelagem em carteira de investimentos.

A seção 2.1 expõem os conceitos e resultados fundamentais sobre a teoria de risco e define novas medidas de variância e covariância, núcleo da proposta de risco baseado em utilidade, tema da seção 2.2. Finalmente, a seção 2.3, introduz o modelo não-linear inteiro para a otimização de um carteira de investimentos, levando-se em conta as medidas de risco da seção anterior. Assim como o modelo de telecomunicação, esse modelo é binário e não-linear. No capítulo 3, apresentamos uma heurística para o problema 0-1 e introduzimos a metodologia de Programação Geométrica Signomial para a resolução dos modelos, originalmente não-lineares inteiros, linearizados e relaxados para o intervalo $[0; 1]$, nas variáveis inteiras. Na busca da solução ótima, utilizamos um algoritmo de pontos interiores para o problema dual, que será introduzido no capítulo 5. No capítulo 4, aproveitamos a estrutura do modelo de telecomunicação sem fio, proposto na seção 1.2, para expor seu tratamento via Decomposição Lagrangeana. Por essa metodologia, a resolução é decomposta nas resoluções de um problema mestre e de três subproblemas escravos. Apoiados em suas respectivas características, propomos métodos de otimização para cada um deles. No capítulo 5, apresentamos o algoritmo de pontos interiores inviáveis utilizado tanto para a metodologia de programação geométrica, quanto para a de decomposição lagrangeana. O capítulo 6 traz detalhes das implementações e os resultados computacionais, para os dois modelos e para as diferentes metodologias. Apresentamos algumas propostas de trabalhos futuros no capítulo 7. Finalmente, nos apêndices encontram-se resumos das metodologias utilizadas no trabalho e alguns resultados computacionais.

Capítulo 1

Modelagem em Telecomunicação Sem Fio

1.1 Aspectos básicos de Sistemas de Telecomunicações Móveis Celular

1.1.1 Conceitos Básicos

Nesta seção, apresentaremos alguns conceitos que julgamos básicos para o entendimento do tema e algumas tecnologias utilizadas na comunicação móvel celular sem fio de forma a facilitar o entendimento do texto. As definições a seguir foram baseadas em [52] e [14].

Definição 1.1. *Uma onda eletromagnética senoidal plana é uma perturbação eletromagnética em que o campo elétrico varia em função do tempo e das coordenadas espaciais na forma*

$$\mathbf{E} = \mathbf{E}_0 \cos\left(2\pi ft - \frac{2\pi x}{\lambda} + \phi\right)$$

onde $|\mathbf{E}_0|$ é a amplitude da onda, f é a frequência em hertz (ciclos por segundo), λ é o comprimento da onda e ϕ é a fase.

Definição 1.2. *O comprimento de onda é a distância percorrida pela onda durante um período (inverso da frequência).*

Definição 1.3. *Canal é um meio de comunicação entre dois pontos.*

Definição 1.4. *Largura de banda é a diferença entre a maior e a menor frequência em um canal. Este conceito é extremamente importante, visto que a capacidade de um canal é, em parte, dependente da largura da banda.*

Definição 1.5. *Uma célula é a área geográfica atendida por um transmissor, a Estação Rádio Base (ERB). Em outras palavras, é o conjunto de todos os pontos onde o sinal da ERB é recebido dentro dos limites de ruído e de interferência aceitáveis. Idealmente, seriam áreas circulares, mas na prática podem assumir formas totalmente irregulares, pois dependem do relevo e da topografia. Por conveniência, são representadas por hexágonos.*

Definição 1.6. *A mudança automática de chamada de uma célula para outra à medida que o usuário se desloca é chamada de handoff e é monitorada por uma Central de Comunicação e Controle, na qual a ERB está conectada.*

Definição 1.7. *O ganho de uma antena é a razão entre a intensidade de radiação em uma dada direção e a intensidade de radiação de uma antena isotrópica, para uma mesma potência incidente na entrada das duas antenas.*

Definição 1.8. *A interferência co-canal é aquela devida ao uso da mesma frequência em células vizinhas. A interferência adjacente é aquela causada por canais adjacentes em uma mesma ERB ou célula.*

Definição 1.8. *O reuso de frequência é a utilização da mesma frequência para transmissão de dois sinais independentes. Depende, entre outros, da potência do sinal, das frequências usadas, do relevo, do ambiente, do tipo e da altura da antena. A distância de reuso é a distância mínima entre duas ERBs transmitindo em canais com mesma frequência sem que haja interferência mútua.*

1.1.2 Sobre Modulação e Técnicas de Acesso

A modulação é o processo pelo qual a informação a ser transmitida é convertida em uma forma conveniente a sua transmissão (geralmente, ocorre a translação da banda básica original para uma banda muito mais alta).

A implementação de um esquema de modulação eficiente e resistente aos problemas apresentados pelo canal móvel não é simples, principalmente devido ao ambiente hostil em termos de propagação em sistemas celulares.

Os sistemas celulares de primeira geração utilizam modulação analógica, enquanto que os sistemas de segunda geração empregam a modulação digital (tanto para os canais de voz, como para os canais de controle). Esse último tipo de modulação tem várias vantagens com relação ao primeiro: maior imunidade a ruído, maior facilidade e praticidade de se multiplexar as informações, maior segurança, além de poder comportar códigos de detecção e/ou eliminação de erros.

No caso da transmissão digital, o sinal original de banda base é transmitido através da modulação de uma portadora senoidal, alterando a amplitude, a frequência ou a fase dessa portadora, que assumem um conjunto de valores discretos. Temos, então, as três formas básicas de modulação de sinais digitais: o ASK (*Amplitude Shift Keying* ou Chaveamento de Amplitude), em que a frequência e a fase permanecem inalteradas, o FSK (*Frequency Shift Keying* ou Chaveamento de Frequência), em que a amplitude e a fase não são alteradas e o PSK (*Phase Shift Keying* ou Chaveamento de Fase), quando tanto a amplitude quanto a frequência não são modificadas.

Visando maior capacidade de transmissão, surgiu a idéia de multiplexação, que consiste em agregar várias informações em uma mesma largura de banda. No caso

da comunicação sem fio, destacam-se a FDM (*Frequency Division Multiplexing*), a TDM (*Time Division Multiplexing*) e a CDM (*Code Division Multiplexing*). As técnicas de acesso são processos através dos quais um determinado usuário tem acesso a um canal. Nos sistemas celulares, um canal é alocado ao usuário, temporariamente, mediante critério de demanda. Das formas de multiplexação acima descritas, resultam as técnicas de acesso FDMA (*Frequency Division Multiplexing Access*), TDMA (*Time Division Multiplexing Access*) e CDMA (*Code Division Multiplexing Access*), respectivamente. Enfatizaremos essa última técnica de acesso por ser ela a utilizada em nossa proposta.

A arquitetura CDMA disponibiliza toda a largura de banda para todos os usuários. Porém, cada conexão estação-usuário recebe um código específico, o mais aleatório e ortogonal possível aos demais, permitindo a utilização do mesmo canal ao mesmo tempo pelos usuários, assim como a interferência entre eles. O uso desses códigos e do espalhamento garante alta segurança e sigilo para o sistema, independente da criptografia. Através do controle de potência de transmissão, podemos minimizar a interferência, aumentando sua eficiência e a qualidade de serviço. Essa redução de potência também permite elevar a capacidade do sistema, tornando-o mais elástico, ao invés de, como nos demais sistemas, bloquear os usuários quando atinge o seu limite máximo. Assim, o número de canais pode ser elevado. Outra característica dessa técnica é a possibilidade de um *handoff* suave entre as células cocanal.

1.1.3 Alguns problemas e modelos existentes

Nesta subseção, apresentaremos três dos principais problemas envolvendo a comunicação móvel sem fio e alguns modelos, já existentes na literatura, que retratem

esses problemas, de forma a minimizá-los. Apresentá-los-emos separadamente, para entendermos a essência de cada um deles. O primeiro problema consiste na Localização das Estações Rádio Base (ERBs), o segundo na Alocação dos Canais e o terceiro, no Controle de Potência. Essa seção está baseada nos trabalhos [52], [92] e [78].

O Problema da Localização de ERBs

Consideremos um conjunto de locais candidatos à instalação de uma facilidade, no nosso caso, de uma estação rádio base. O problema da localização de ERB consiste em selecionar naquele conjunto um subconjunto de mínimo custo com a propriedade de cobrir toda a área em estudo, atendendo a sua demanda e explorando eficientemente o espectro de frequência. Em outras palavras, o problema pode ser entendido como a cobertura total ou a cobertura máxima, que consiste em assinalar a cada ponto da área pelo menos uma ERB capaz de oferecer um sinal com nível mínimo, permitindo a comunicação e/ou a transmissão de dados, satisfazendo os padrões de qualidade de serviço (QoS), levando em conta que o custo desse serviço está diretamente vinculado à quantidade de ERBs instaladas. Assim, a solução desse problema será obtida quando selecionarmos um subconjunto de ERBs que atinge a relação custo/QoS ótima.

Modelos de Localização de ERBs

Seja uma região dividida em um número M de pequenas quadrículas de dimensões variáveis (no nosso caso, serão aproximadas por pontos). Seja o conjunto de N ERBs candidatas à instalação distribuídas nessa região. Vamos definir os seguinte conjuntos de variáveis de decisão, que serão usadas em todo o trabalho,

tomando, como conhecida, alguma medida de sinal recebido em cada quadrícula e originário de cada ERB:

$$z_i = \begin{cases} 1, & \text{se a ERB } i \in N \text{ é localizada;} \\ 0, & \text{caso contrário;} \end{cases}$$

$$x_{ij} = \begin{cases} 1, & \text{se o ponto } j \in M \text{ da região é coberto pela ERB } i \in N; \\ 0, & \text{caso contrário;} \end{cases}$$

Um primeiro modelo de localização, baseado nesses conjuntos, pode ser definido como:

Modelo de Localização (ML1):

$$\min \sum_{i \in N} c_i z_i$$

s.a.:

$$\sum_{i \in N} b_{ij} x_{ij} \geq T \quad \forall j \in M$$

$$\sum_{i \in N} x_{ij} \geq 1 \quad \forall j \in M$$

$$x_{ij} \leq z_i \quad \forall i \in N, \forall j \in M$$

$$x_{ij}, z_i \in \{0, 1\} \quad \forall i \in N, \forall j \in M$$

onde

c_i : é o custo fixo de instalação da ERB i ;

T : valor limite para a medida do sinal capaz de viabilizar a comunicação;

b_{ij} : estimativa do sinal da ERB i no ponto j ;

Observações:

1. A função objetivo minimiza, como requerido, o custo fixo de instalação das ERBs. Notemos que o custo varia de acordo com a localização da facilidade $i \in N$. Dependendo do mercado, esses custos variam de 200 a 500 mil dólares cada.
2. O primeiro conjunto de restrições garante que o sinal recebido por cada quadrícula $j \in M$ é no mínimo igual ao valor limite capaz de estabelecer a comunicação com qualidade de serviço;
3. O segundo conjunto de restrições garante que todo ponto $j \in M$ estará coberto por pelo menos uma ERB;

4. O terceiro conjunto de restrições impõe que só participem da cobertura total da área aquelas ERBs que tenham sido instaladas;
5. Finalmente, o último conjunto de restrições determina a integrabilidade das variáveis.

Percebendo que os primeiro e segundo conjunto de restrições serão satisfeitos simultaneamente sempre que o sinal de uma ERB i for capaz de atender à viabilidade de comunicação em uma quadrícula $j \in M$, podemos simplificar o modelo ML1, selecionando as transmissões da ERB i para o ponto j que satisfaçam essa condição, gerando:

Modelo de Localização (ML2):

$$\min \sum_{i \in N} c_i z_i$$

s.a:

$$\sum_{i \in N} a_{ij} x_{ij} \geq 1 \quad \forall j \in M$$

$$z_i, x_{ij} \in \{0, 1\} \quad \forall i \in N, \forall j \in M$$

onde

$$a_{ij} = \begin{cases} 1, & \text{se } b_{ij} \geq T, \forall i, \forall j \\ 0, & \text{caso contrário} \end{cases}$$

Observemos que cada valor b_{ij} é uma constante conhecida e influencia diretamente na simplificação do modelo, pois é usada para a associação das ERBs para as quadrículas selecionadas. Esse novo modelo torna bem mais simples a implementação, continuando viável para muitas aplicações, entre elas os projetos iniciais de atendimento e às regiões de baixa demanda por comunicação.

O Problema de Alocação de Canais

Já sabemos que o custo de serviço está diretamente ligado ao número de ERBs instaladas. A redução desse número implica na diminuição de custos, mas pode acarretar também na redução da qualidade de serviço. Assim, para atender a esse objetivo, sem interferir na QoS, é preciso uma exploração eficiente do espectro de frequência. Observamos que uma dificuldade adicional é o problema da carência de frequências livres, isto é, o número de aplicações dos sistemas sem fio aumenta, enquanto que o espectro de frequência disponível para transmissão sem fio permanece constante. Assim, é preciso dividir esse espectro entre uma quantidade crescente de sistemas.

A alocação de canais consiste em distribuir entre as ERBs o conjunto de canais disponíveis no sistema, observando os níveis de interferência e distância de reuso. O problema de alocação de canais consiste em calcular essa distribuição.

Nesse contexto entretanto é indispensável o conhecimento da tecnologia utilizada. Dependendo dela, a solução do problema segue caminhos distintos: podemos resolvê-lo dividindo os canais em subconjuntos ortogonais, baseados em um fator de reuso e distribuí-los às estações observando a distância de reuso; podemos criar mapas de interferência para cada ERB, atribuindo um canal a uma ERB somente se é suficientemente distante de cada um dos canais das ERBs interferentes e podemos ainda reduzir a interferência ajustando as potências de transmissão de cada ERB.

Para um melhor entendimento dos modelos, dividimos a apresentação em duas partes. Primeiramente, tratamos de um modelo bem simples de alocação fixa de canais, apenas para a familiarização com o problema. Mais adiante, quando tratarmos em conjunto dos problemas de localização de ERBs e de alocação de canais,

apresentaremos um modelo mais elaborado.

Modelo de Alocação Fixa de Canais

Neste caso de alocação, supõe-se que os canais disponíveis são ortogonais. Isso significa que fica estabelecido um intervalo de um canal entre canais consecutivos. Assim, obtemos uma abordagem bem mais simples para o problema em detrimento de um número total reduzido de canais disponíveis no sistema.

O problema de alocação de frequências sob esse aspecto se reduz a um problema de alocação de canais às ERBs:

Sejam N e K o conjunto de ERBs localizadas e de canais, respectivamente, e consideremos o seguinte conjunto de variáveis de decisão:

$$y_{ik} = \begin{cases} 1, & \text{se o canal } k \text{ é alocada à estação } i; \\ 0, & \text{caso contrário} \end{cases}$$

Seja G^i o conjunto de estações interferentes à estação $i \in N$. Por construção, cada frequência k só pode ser alocada a uma estação G^i . Seja d_i , $i \in N$, o parâmetro de demanda de chamadas. Essa medida condiciona diretamente a alocação de canais às ERBs.

Um modelo matemático com o objetivo de maximizar o número de canais no sistema pode ser formulado como:

Modelo de Alocação de Canais (MAC):

$$\max \sum_{i \in N} \sum_{k \in K} y_{ik}$$

sujeito a:

$$\sum_{l \in G^i} y_{lk} \leq 1, \quad \forall i \in N, \forall k \in K$$

$$\sum_{k \in K} y_{ik} \geq d_i, \quad \forall i \in N$$

$$y_{ik} \in \{0, 1\}, \quad \forall i \in N, \forall k \in K$$

A função objetivo, como desejado, maximiza o número de canais disponíveis no sistema. Isso implica em melhor atendimento aos usuários. O primeiro conjunto de restrições impõe que cada canal só seja alocado a uma das ERBs interferentes entre si. O segundo conjunto de restrições garante o atendimento à demanda e finalmente o terceiro conjunto de restrições indica a integrabilidade das variáveis de decisão.

Modelo de Localização de Estações e Alocação de Canais

Nesta subseção, apresentaremos um modelo para a resolução simultânea dos problemas abordados anteriormente: a localização das estações rádio base e a alocação de canais. Esse modelo, que chamaremos de MLA1, é uma tentativa de unir um modelo básico de localização de ERB com o modelo MAC, apresentado na subseção anterior, de forma a alocar um canal a apenas uma das ERBs pertencentes ao conjunto de ERBs interferentes entre si. Existem vários outros modelos, entre eles um desenvolvido por LEE [45], cujo critério para alocação de canais consiste em agrupá-los em subgrupos não interferentes, baseado na largura da banda e na tecnologia adotada nas antenas das ERBs.

Modelo MLA1:

$$\min \sum_{i \in N} c_i z_i$$

s.a.:

$$\sum_{i \in N} a_{ij} x_{ij} \geq 1, \quad \forall j \in M$$

$$\sum_{j \in M} a_{ij} x_{ij} \leq m_i z_i, \quad \forall i \in N$$

$$\sum_{l \in G^i} y_{lk} \leq 1, \quad \forall i \in N, \forall k \in K$$

$$\sum_{k \in K} y_{ik} - \sum_{j \in M} d_j x_{ij} \geq 0, \quad \forall i \in N$$

$$\sum_{k \in K} y_{ik} \leq n_i z_i, \quad \forall i \in N$$

$$z_i, x_{ij}, y_{ik} \in \{0, 1\}, \quad \forall i \in N, \forall j \in M, \forall k \in K$$

onde

$$a_{ij} = \begin{cases} 1, & \text{se } b_{ij} \geq T, \forall i, \forall j \\ 0, & \text{caso contrário} \end{cases}$$

e b_{ij} é uma constante previamente conhecida.

Observações:

1. O primeiro conjunto de restrições garante o cobrimento de toda a área em estudo, com um sinal capaz de viabilizar a comunicação, associando a cada ponto $j \in M$ pelo menos uma ERB i ;
2. O segundo conjunto de restrições limita o número de pontos alocados a cada ERB, sendo m_i a quantidade máxima de pontos cobertos pela estação $i \in N$;

3. O terceiro conjunto de restrições impede que um canal seja alocado a mais de uma ERB do conjunto de ERBs interferentes entre si;
4. O quarto conjunto de restrições reflete a ponderação fundamental entre a área coberta por uma ERB e o maior número de canais a ela alocados. Logicamente, quanto maior a área, maior o número de canais necessários. Como estes são um recurso escasso, o modelo tentará diminuir aquela área;
5. O quinto conjunto de restrições limita o número de canais alocados a cada ERB, sendo n_i o número máximo de canais suportados pela ERB i . Além disso, o conjunto garante que só sejam alocados canais às ERBs instaladas, através da relação entre as variáveis z e y ;
6. Finalmente, o último conjunto de restrições indica a integrabilidade das variáveis de decisão.

O Problema do Controle de Potência

Nesta subseção, trataremos do controle da potência emitida pelas estações transmissoras, usado na arquitetura CDMA. Sua vantagem em relação às demais abordagens é que tal controle permite que mais sistemas dividam a mesma banda. Enfatizaremos sua importância e mostraremos o processo de montagem das restrições [89]. Apresentaremos ainda um modelo desenvolvido por SAMPATH [79] para transmissão de voz.

O critério usado para o controle de potência consiste logicamente em minimizar o total de potência transmitida. São várias as razões para tanto, entre elas:

- Menor potência transmitida significa menor espaço ocupado no espectro e conseqüentemente aumento na capacidade de usuários em uma mesma área;

- Redução no custo da frequência usada, já que cada operador precisa pagar pelo uso do espectro disponível;
- Redução no consumo de bateria. Além disso, com uma potência de transmissão mais baixa, a vida útil da bateria é maior e o equipamento poderá ser mais leve;
- Proteção ao meio ambiente, já que as ondas de rádio interagem com os organismos vivos, provocando, em casos extremos, efeitos nocivos.

No controle de potência, cada usuário especifica um número mínimo tolerável de qualidade de serviço (QoS). Geralmente, esse valor é expresso em forma de *bit error rate* (BER). Esse mínimo pode ser expresso através da relação $\frac{E_b}{N_0 + I_0}$, onde E_b é a energia por bit do sinal recebido, N_0 é a potência de ruído recebido e I_0 é a potência de sinal interferente recebido. De acordo com SAMPATH *et al.* [78], para um sistema unicelular, aquele que só contém uma ERB, tal relação pode ser dada por:

$$\left(\frac{E_b}{N_0 + I_0}\right)j = \frac{W}{r_j} \frac{h_j p_j}{\sum_{l \neq j} h_l p_l + \eta_0 W} \quad (1.1)$$

para cada quadrícula j da região, onde

W : é a largura da banda, da ordem de 1,25 MHz;

η_0 : é a densidade de ruído, considerado gaussiano, da ordem de 10^{-6} W/hertz;

h_j : é o ganho por canal;

p_j : é a variável referente à potência transmitida para a quadrícula j , limitada superiormente por 0,5 W;

r_j : é a variável referente à taxa de transmissão da quadrícula j , limitada inferiormente por 8 kbps.

Modelo de Controle de Potência (MCP):

Seja M o número de usuários (quadrículas) no sistema. Cada um deles especifica uma qualidade de serviço γ_j , $j = 1, \dots, M$, desejada. Definamos por r_j a taxa de transmissão e por p_j a potência transmitida para cada quadrícula $j \in M$ (r e p são os respectivos vetores). Sejam ainda W a largura da banda, η_0 o ruído gaussiano, h_j o ganho obtido por cada quadrícula j , P_j as potências máximas permitidas, e $R_j > 0$ as taxas de transmissão mínimas aceitáveis para cada usuário.

Assim, o modelo fica formulado por:

$$\min_{p,r} \sum_{j=1}^M p_j$$

sujeito a

$$\begin{aligned} \frac{W}{r_j} \frac{h_j p_j}{\sum_{l \neq j} h_l p_l + \eta_0 W} &\geq \gamma_j, \quad \forall j \\ 0 < p_j &\leq P_j, \quad \forall j \\ r_j &\geq R_j, \quad \forall j \end{aligned} \tag{1.2}$$

A função objetivo controla a potência transmitida. O primeiro conjunto de restrições garante uma qualidade de serviço (QoS) no mínimo igual à desejada. O segundo conjunto de restrições está de acordo com as limitações físicas da transmissão. Finalmente, o último conjunto garante que a taxa de transmissão não será menor do que a mínima tolerável. Sobre esse modelo, os autores enunciam a seguinte proposição:

Proposição 1.1. *Na solução ótima, todas as restrições de QoS e de taxa de transmissão são ativas, ou seja, um ponto (p^*, r^*) é ótimo se e só se para todas suas componentes valer:*

$$r_j^* = R_j \text{ e } p_j^* \text{ tal que } \frac{W}{R_j} \frac{h_j p_j^*}{\sum_{l \neq j} h_l p_l^* + \eta_0 W} = \gamma_j.$$

Observação: Não apresentaremos a prova para essa proposição, por ser idêntica à prova da proposição 3.2, apresentada mais adiante. SARAYDAR *et al.* [83] interpretam essa propriedade da solução ótima da seguinte forma: a percepção da qualidade em uma transmissão de voz permanece constante para um certo *signal to interference* (SIR, equação 1.1). Assim, qualquer qualidade superior a essa é a mesma para o ouvido humano, indicando que basta atingir o mínimo desejável de QoS, para alcançarmos o ótimo. Entretanto, serviços de transmissão de dados são sensíveis a mudanças no SIR, por isso requerem um controle de potência diferente do convencional para voz, no caso, com a introdução de uma função utilidade (ver capítulo 6).

1.2 Um novo modelo para Telecomunicação Sem Fio

Os problemas de localização de estações rádio base, de alocação de canais e de controle de potência são geralmente tratados independentemente, originando soluções que podem ser diferentes da solução ótima. Existem, como vimos no capítulo 2, alguns modelos que abordam, simultaneamente, os problemas de localização de ERBs e de alocação de frequências, utilizando a tecnologia TDMA. Existem também vários modelos na literatura [78], [100], [98], [30] envolvendo o controle de potência. O objetivo dessa seção é apresentar um novo modelo para telecomunicação sem fio¹, envolvendo, simultaneamente, os problema de localização de

¹O desenvolvimento deste modelo resultou da colaboração do Prof. Geraldo Robson Mateus, do Departamento de Ciências da Computação da UFMG.

estações rádio base e de controle de potência. Como relatado em [92], apesar dos serviços de telecomunicações móveis estarem aparentemente evoluindo em direção ao GSM (*Global System for Mobile communications*), a tecnologia utilizada em nosso trabalho é o CDMA, pois o modelo resultante tem características interessantes, possibilitando também o emprego de várias metodologias de otimização.

Como poderemos perceber, o modelo é resultado do acoplamento de um modelo básico de localização de ERBs e do modelo (MCP) de controle de potência desenvolvido em [78] e apresentado no capítulo anterior.

1.2.1 Modelo Não-Linear Inteiro

O tratamento em conjunto dos problemas de localização de ERBs e de controle de potência requer uma atenção especial na formulação das restrições e na escolha da função objetivo. Além de viabilizar isoladamente cada um dos problemas, o modelo deve conter uma variável que crie um elo entre eles, de modo a impedir falsas soluções ótimas. Consideremos, então, o seguinte problema de Otimização:

$$\min \sum_{i=1}^N c_i z_i + C. \sum_{i=1}^N \sum_{j=1}^M p_{ij}$$

s.a :

$$\begin{aligned} \sum_{i=1}^N a_{ij} x_{ij} &\geq 1, & \forall j \\ \sum_{j=1}^M a_{ij} x_{ij} &\leq \Lambda_i z_i, & \forall i \\ \frac{W h_j p_{ij}}{\left(\sum_{k \neq j} a_{ik} h_k p_{ik} \right) + \left(\sum_{s \neq i} \sum_{t=1}^M a_{sj} h_t p_{st} \right) + \eta_0 W} &\geq \gamma_j r_{ij}, & \forall i, \forall j \\ 0 &\leq p_{ij} \leq P_j x_{ij}, & \forall i, \forall j \\ R_j x_{ij} &\leq r_{ij}, & \forall i, \forall j \\ x_{ij}, z_i &\in \{0, 1\}, & \forall i, \forall j \end{aligned} \quad (1.3)$$

onde

$$a_{ij} = \begin{cases} 1, & \text{se a ERB } i \text{ é capaz de cobrir a quadrícula } j \\ 0, & \text{caso contrário.} \end{cases}$$

1.2.2 Explicação do Modelo

Primeiramente, dividimos a região em estudo em M quadrículas de tamanho variável, enumerado-as por $j = 1, \dots, M$. Nessa região, espalhamos N ERBs candidatas à instalação, indicando-as por $i = 1, \dots, N$. Cada ERB i tem um custo de instalação c_i e uma capacidade máxima Λ_i , de quadrículas atendidas. Cada quadrícula j possui um valor máximo P_j , para a potência de transmissão. Valores superiores a esse máximo podem danificar o equipamento do usuário, além de consumir mais

bateria e aumentar a interferência prejudicando a comunicação. Cada quadrícula define também um valor mínimo requerido γ_j , para a qualidade de serviço e um valor mínimo de taxa de transmissão R_j . Valores abaixo desse mínimo inviabilizam a comunicação. Entretanto, valores muito superiores a R_j aumentam a quantidade de potência transmitida necessária para a comunicação com qualidade, tanto para a quadrícula em questão, como para as quadrículas interferentes, aumentando, conseqüentemente, o custo. Consideramos ainda as constantes W , largura da banda, η_0 , ruído gaussiano e C , custo da potência transmitida para a ERB i . Assim como γ_j , as constantes R_j , W e η_0 seguem a forma do modelo MCP (1.2), apresentado na seção 1.1.

O objetivo do modelo é minimizar os custos de instalação das estações rádio base e da transmissão de potência entre as ERBs e as quadrículas. Para homogeneizar a função objetivo como minimizadora de custo, o somatório da potência transmitida foi multiplicado pela constante C , que pode ser entendida como o custo unitário da medida de potência.

As variáveis do modelo podem ser separadas em quatro conjuntos: z_i , variáveis de instalação, que assumem valor unitário se a ERB i é instalada, ou nulo, caso contrário; x_{ij} , variáveis de cobertura, que assumem valor unitário, se a quadrícula j é coberta pela ERB i , ou nulo, caso contrário; p_{ij} , variáveis que indicam a potência transmitida pela quadrícula j para a ERB i . Essas variáveis são contínuas, limitadas inferiormente por zero e superiormente por P_j . Observemos ainda que, quando a quadrícula j não é coberta pela ERB i , o valor de p_{ij} é igual a zero. Finalmente, o conjunto r_{ij} são as variáveis relativas à taxa de transmissão entre as quadrículas e as ERBs. Assim como p_{ij} , essas variáveis são contínuas, porém limitadas inferiormente

por R_j , taxa mínima requerida por cada quadrícula j para viabilização da comunicação. Observemos que R_j vem multiplicada por x_{ij} , fazendo com que o limite inferior para a taxa seja nulo, quando a quadrícula j não estiver sendo coberta pela ERB i . O limite superior para a taxa de transmissão está condicionado aos valores da potência de transmissão, no terceiro conjunto de restrições

O primeiro conjunto de restrições garante que cada quadrícula é coberta por pelos menos uma ERB. Isso significa que o sinal naquela quadrícula, originário de alguma ERB, é suficiente.

Pelo segundo conjunto de restrições, o somatório dos sinais originários de cada ERB i que chegam nas quadrículas por ela cobertas não pode ultrapassar a sua capacidade máxima de atendimento. Mais ainda, cada restrição considera a possibilidade da ERB em questão não ser instalada. Nesse caso, o somatório anterior é forçosamente nulo.

O terceiro conjunto de restrições garante a qualidade de serviço requerida por cada quadrícula. Além disso, cada restrição leva em conta o fato da ERB i não cobrir a quadrícula j . Nesse caso, o valor de r_{ij} seria nulo, assim como a qualidade de serviço requerida e a potência transmitida.

1.2.3 Análise do Modelo

A menos das restrições de qualidade de serviço, as demais funções do modelo são lineares. Desse modo, uma melhor compreensão do problema exige, entre outros, a análise daquelas restrições.

Na forma apresentada, cada linha (i, j) é uma função fracionária-linear em p_{ij} e linear em r_{ij} . Como tal, possui a propriedade de ser pseudoconvexa e pseudocôncava (ver [3]). Tem-se também, como consequência desses fatos, que cada restrição é

quasiconvexa, quasicôncava, estritamente quasiconvexa e estritamente quasicôncava. Segue daí a importante propriedade de que a relaxação contínua do modelo não-linear terá cada solução local, supondo que exista, sendo também global (ver [3]), em um ponto verificando as condições de KKT.

A proposição 1.1, transcrita de [78], continua válida para nosso modelo e será provada na seção 4.3.

Capítulo 2

Modelagem em Carteira de Investimentos

2.1 Aspectos básicos sobre Utilidade, Risco e Retorno

2.1.1 Conceitos e Resultados

Definição 2.1. *O prêmio de risco é o incremento, em relação à taxa de retorno livre de risco, exigido pelo investidor, para assumir um investimento com risco diferente de zero.*

Definição 2.2. *Um investidor é dito conservador quando tem grande aversão ao risco e agressivo, caso contrário. Os investidores moderados são aqueles com média aversão ao risco.*

Podemos, então, concluir que o investidor conservador exige maior prêmio de risco que o investidor agressivo para assumir o mesmo investimento de risco.

Geralmente, a influência do fator risco é uma função não-linear do prêmio exigido. Assim, a atitude do investidor em relação ao risco pode ser interpretada matematicamente como uma função de *trade-off*, ou curva de indiferença ou utilidade, entre risco e retorno.

Definição 2.3. O retorno esperado para uma carteira composta por N ativos é dado por:

$$R_{Cart} = \sum_{i=1}^N x_i R_i, \quad (2.1)$$

onde x_i é a ponderação do ativo i na carteira e R_i é o seu retorno individual.

Definição 2.4. O risco associado a uma carteira de investimentos formada por N ativos é dado por:

$$\sigma_{Cart} = \sqrt{\sum_{i=1}^N x_i^2 \sigma_i^2 + \sum_{i=1}^N \sum_{j \neq i}^N x_i \sigma_i x_j \sigma_j COR(i; j)}$$

onde x_i , com $\sum_{i=1}^N x_i = 1$, e σ_i são, respectivamente, a ponderação e desvio-padrão do ativo i e $COR(i; j)$ é a correlação linear entre os retornos dos ativos i, j .

Para melhor entendimento, consideremos uma carteira formada por dois ativos.

Pela definição, temos:

$$\sigma_{Cart} = \sqrt{x_1^2 \sigma_1^2 + x_2^2 \sigma_2^2 + 2x_1 \sigma_1 x_2 \sigma_2 COR(1; 2)}$$

onde $x_1 + x_2 = 1$.

Assim, a ponderação x_1 para uma carteira de risco mínimo, $\sigma'_{Cart} = 0$ é dada por:

$$x_1 = \frac{\sigma_2(\sigma_2 - \sigma_1 COR(1; 2))}{\sigma_1^2 + \sigma_2^2 - 2\sigma_1 \sigma_2 COR(1; 2)}$$

Observação: A limitação ao intervalo $[0; 1]$ para as ponderações x_i não é obrigatória. É apenas uma regra para a apresentação da teoria. O caso $x_i > 1$ indica a necessidade de tomar empréstimo e enquanto que o caso $x_i < 0$ implica na venda a descoberto.

Definição 2.5. Dizemos que uma carteira é diversificada quando os ativos que a compõem são escolhidos de modo a minimizar seu risco. Uma carteira é dita

alavancada quando os ativos que a compõem são tomados de modo a aumentar seu risco.

Observação: Por essa definição, fica claro que o princípio básico para montarmos uma carteira diversificada é escolher ativos de alta correlação negativa entre si. Ao contrário, para termos uma carteira alavancada, devemos escolher ativos de alta correlação positiva.

Definição 2.6. *Uma carteira é dita eficiente quando oferece o maior retorno possível para um valor específico do risco. Ou, o menor risco possível para um dado retorno. A fronteira eficiente é o conjunto formado por todas as carteiras eficientes.*¹

2.1.2 Modelagem e Gerenciamento do Risco

Esta subseção se baseia em opiniões de Duarte [19] e Damodaran [16] no intuito de evidenciar a necessidade do gerenciamento e de uma boa modelagem do risco².

Segundo Duarte "A implementação do gerenciamento de risco deve ser uma decisão de quem efetivamente detém o poder decisório na instituição.", deve-se "...buscar profissionais qualificados e experientes para essa tarefa.", "...um mau gerenciamento de risco pode levar a uma falsa sensação de segurança, o que pode ser até mesmo pior que desconhecer - e, portanto, temer - o risco de suas posições hoje." e "A confiabilidade da estimativa final obtida para o risco de uma instituição está diretamente relacionada à qualidade dos dados usados e dos procedimentos computacionais implementados."

Para Damodaran "...um bom modelo de risco e retorno:

1. Deve proporcionar uma medida de risco que se aplique a todos os ativos, não

¹Gráficos de risco e retorno para dois ativos com diferentes correlações e da fronteira eficiente para dois ou mais ativos podem ser encontrados em [10].

²Maiores aprofundamentos quanto à Teoria de Risco podem ser obtidos em [9], [10], [16] e [76].

sendo específica apenas a um deles;

2. Deve estabelecer claramente quais tipos de risco são recompensados e quais não são, assim como fornecer um fundamento lógico para essa distinção;
3. Deve proporcionar medidas de risco padronizadas, isto é, um investidor diante de uma medida de risco para um ativo específico deve ser capaz de tirar conclusões a respeito de se o ativo está acima ou abaixo da média em termos de risco;
4. Deve traduzir a medida de risco em uma "taxa de retorno" que o investidor deve exigir como compensação por assumir o risco;
5. Deve funcionar bem não apenas ao explicar retornos passados, mas também ao prever retornos futuros."

Os modelos mais comuns³, com menor ou maior grau de complexidade, por exemplo, o Modelo de Média-Variância de Markowitz, o Modelo de Precificação de Ativos de Capital (CAPM) [91], [85], [48] e a Teoria de Precificação por Arbitragem (APT) [75], satisfazem a todos esses requisitos, com a ressalva de apenas recompensarem o risco de mercado⁴.

2.2 Proposta de Cálculo de Risco baseado em Utilidade

Nosso objetivo nessa seção é propor um cálculo para a mensuração do risco, levando em conta a característica suposta comum aos investidores racionais e con-

³Uma revisão bibliográfica sobre esses modelos pode ser encontrada em [44].

⁴Segundo [16], são mudanças inesperadas em fluxos de caixa de projetos, geradas por mudanças em taxas de juros, taxas de inflação e condições econômicas que afetam todos os projetos e todas as empresas, ainda que em graus diferentes

servadores, comentada na introdução. Para tanto, introduziremos uma função utilidade, dependente dos parâmetros média e desvio-padrão, de modo a desconsiderar observações discrepantes e penalizar perdas com peso superior ao de privilegiar ganhos. Na próxima subseção, apresentaremos um exemplo simples, para retratar a necessidade da introdução de uma função utilidade com tais propriedades.

2.2.1 A Curva de Utilidade

A Teoria de Utilidade encontra-se amplamente difundida e é utilizada em diversas áreas do conhecimento, entre elas, Economia [58], [34], Telecomunicações [83], [84] e Finanças [20]. Seu conceito se baseia no princípio da saturabilidade⁵ e, por questões didáticas, é comum apresentá-lo ligado apenas à propensão ou satisfação em consumir⁶, mas pode-se, perfeitamente, estendê-lo ao desprazer de não consumir.

Por exemplo, imaginemos um jovem que costuma beber 15 garrafas de refrigerante por semana. Se em uma dada semana, ele consumir 16 garrafas, sua satisfação aumentará, devido a esse acréscimo de uma garrafa. Se na semana seguinte, ele consumir 17 garrafas, outra vez, sua satisfação aumentará, porém, podemos supor, não em mesmo grau. Em outras palavras, o segundo acréscimo de uma garrafa, apesar de interessante, não é recebido com o mesmo entusiasmo do primeiro acréscimo. Prosseguindo nessa rotina, chegaremos a um ponto em que consumir o acréscimo será indiferente para o jovem.

Se, ao invés de receber garrafas a mais, a partir da quantidade de costume, fosse-lhe tirada uma garrafa na semana, é razoável supor decepção. Se na semana seguinte,

⁵O primeiro a atentar para esse princípio foi o economista alemão Herman Heinrich Gossen: "A partir do pressuposto de que toda conduta humana tem por objetivo um máximo de satisfação, Gossen desenvolve algumas leis, das quais duas são conhecidas como Leis de Gossen. A primeira apresenta o princípio da utilidade decrescente: "A quantidade de uma mesma satisfação diminui constantemente à medida que a realizamos sem interrupção, até obter a saciedade"..." [80].

⁶Maior detalhamento e outros exemplos podem ser encontrados em [34] e [58].

o jovem apenas dispusesse de 13 garrafas, outra vez, haveria descontentamento, porém em maior grau: a variação negativa de 15, como esperado, para 14 será menos sentida que a variação de 14 para 13. Mais ainda, o acréscimo de 15 para 16 não é equivalente ao decréscimo de 15 para 14, muito menos o acréscimo de 16 para 17 não iguala o decréscimo de 14 para 13. No extremo positivo, o acréscimo de uma garrafa semanal, quase indiferente para o jovem, tem peso infinitamente menor que o decréscimo de outra garrafa em um ponto em que o consumo semanal estivesse próximo a zero.

A satisfação com o ganho marginal e a decepção com a perda marginal são chamados de utilidade marginal. A função do consumo, cuja derivada é a utilidade marginal é a função utilidade⁷.

Em relação ao mercado financeiro, podemos identificar o jovem com o investidor, a quantidade de garrafas semanais com o retorno esperado no período e a possibilidade de acréscimos e decréscimos com, respectivamente, retornos maiores ou menores que o esperado, traduzida (a possibilidade) pelo desvio-padrão. Dessa forma, variações positivas em relação ao retorno esperado seriam bem menos sentidas que oscilações negativas, cada vez em maior grau.

2.2.2 A Distância-Utilidade e o Desvio-Utilidade

Nessa subseção, apresentaremos uma proposta de cálculo de risco baseado em uma função utilidade parametrizada pela média dos retornos e pelo desvio-padrão do índice de mercado. Assim como no CAPM (ver [76]), consideraremos uma distribuição de freqüências normal para o comportamento dos retornos.

Definição 2.7. *O desvio-padrão amostral (ou populacional) de referência é um*

⁷Um resumo de várias funções utilidade, seus comportamentos e derivadas pode ser encontrado em [10] e [59].

desvio-padrão amostral (ou populacional) representativo do mercado em que está inserida a amostra.

Por exemplo, para ativos negociados na BOVESPA, o desvio-padrão de referência será o desvio-padrão das oscilações do Ibovespa.

Definição 2.8. *Uma amostra A será dita Amostra Referencial Válida se:*

1. $\bar{X} + 3\sigma > 0$ e
2. $P(Y) \geq 90\%$, onde $Y = \{X_i \in A; X_i + 3\sigma > 0, \text{ para } i = 1, \dots, n\}$,

onde n é o número de observações na amostra, X_i é a observação, \bar{X} é a média, σ é o desvio-padrão amostral de referência e $P(Y)$ é a probabilidade associada ao conjunto Y

Definição 2.9. *A Função-Utilidade definida em A é dada por*

$$U(X) = \rho \ln \left(\frac{X + 3\sigma}{\bar{X} + 3\sigma} \right), \quad (2.2)$$

onde X é uma observação, \bar{X} é a média da amostra referencial válida, σ é o desvio-padrão amostral (ou populacional) de referência e ρ é um parâmetro positivo que depende do perfil do investidor, apenas quanto ao risco.

Considerações:

- Tomamos apenas amostras referenciais válidas, pois ativos com $\bar{X} + 3\sigma \leq 0$ têm diminutas possibilidades de retorno positivo. De fato, como estamos supondo um comportamento normal para os retornos, haverá aproximadamente 0,115% de chance de um futuro retorno ser negativo (ver [101]);
- Ainda sobre amostra referencial válida, mesmo satisfazendo $\bar{X} + 3\sigma > 0$, só serão considerados os ativos tais que $X + 3\sigma > 0$, para 90% das observações (as observações complementares na amostra serão descartadas). Os ativos que não satisfazem essa premissa também têm pouca chance de originar retornos positivos, pelo mesmo motivo do item acima. Outra razão para essa consideração é a garantia de termos pelo menos 81%(= 90% × 90%) das observações em condições de avaliarmos correlação entre dois ativos.

Observações:

1. O quociente entre parênteses é sempre positivo, e o logaritmo está bem definido;
2. As propriedades desta função se justapõem ao comportamento desejado para a avaliação das observações. Dentre outras, destacamos que o valor absoluto do logaritmo para variações negativas em relação à média cresce mais rápido que o (valor absoluto) do logaritmo para variações positivas em relação à média;
3. Em [20], propõe-se uma função utilidade linear dada pela soma de uma *wealth function* linear (função riqueza) com uma função penalidade, também linear, em que essa última é aplicada apenas para retornos (ou riquezas) abaixo do

mínimo aceitável. Com essa metodologia, dá-se maior peso à variações negativas do que à variações positivas, porém mantendo-o constante.

Definição 2.10. *A Distância-Utilidade para $x, y \in A$ é o número real dado por*

$$d(x, y) = \begin{cases} d(x, \bar{X}) - d(y, \bar{X}), & \text{para } x \leq y \leq \bar{X} \text{ ou } \bar{X} \leq y \leq x \\ d(y, \bar{X}) - d(x, \bar{X}), & \text{para } y \leq x \leq \bar{X} \text{ ou } \bar{X} \leq x \leq y \\ d(x, \bar{X}) + d(y, \bar{X}), & \text{para } x \leq \bar{X} \leq y \text{ ou } y \leq \bar{X} \leq x \end{cases} \quad (2.3)$$

onde

$$\begin{aligned} d(x, \bar{X}) &= (x - \bar{X})^2 [U(x)]^4 \\ &= (x - \bar{X})^2 \ln^4 \left(\frac{X+3\sigma}{\bar{X}+3\sigma} \right), \quad \forall x \in A \end{aligned} \quad (2.4)$$

e $\rho = 1$.

Observações:

- O número real $d(x, \bar{X})$ é positivo para todo $x \in A$, com $x \neq \bar{X}$;
- Consideramos $\rho = 1$, para facilitar o entendimento, porém as definições e proposição desta seção continuam válidas para $0 < \rho \neq 1$.

Proposição 2.1. *A Distância-Utilidade satisfaz às propriedades de distância [47]:*

- d1) $d(x, x) = 0$;
- d2) Se $x \neq y$, então $d(x, y) > 0$;
- d3) $d(x, y) = d(y, x)$ e
- d4) $d(x, z) \leq d(x, y) + d(y, z)$.

Prova:

d1) $d(x, x) = d(x, \bar{X}) - d(x, \bar{X}) = 0$, pelas primeira e segunda hipóteses. Pela terceira hipótese, temos $d(x, x) = d(x, \bar{X}) + d(x, \bar{X}) = 0$, pois $x = y = \bar{X}$.

d2)

Caso 1: $\bar{X} \leq y \leq x$

$$\begin{aligned}d(x, y) &= d(x, \bar{X}) - d(y, \bar{X}) \\&= (x - \bar{X})^2 \ln^4\left(\frac{x+3\sigma}{\bar{X}+3\sigma}\right) - (y - \bar{X})^2 \ln^4\left(\frac{y+3\sigma}{\bar{X}+3\sigma}\right) \\&\geq (x - \bar{X})^2 \ln^4\left(\frac{x+3\sigma}{\bar{X}+3\sigma}\right) - (x - \bar{X})^2 \ln^4\left(\frac{y+3\sigma}{\bar{X}+3\sigma}\right) \\&= (x - \bar{X})^2 \left[\ln^4\left(\frac{x+3\sigma}{\bar{X}+3\sigma}\right) - \ln^4\left(\frac{y+3\sigma}{\bar{X}+3\sigma}\right) \right]\end{aligned}$$

Use agora o fato de que $\ln^4(t)$ é crescente para $t \geq 1$ (pois $\frac{\ln^3(t)}{t} \geq 0$) para concluir que a expressão acima é estritamente positiva se $x \neq y$.

Caso 2: $y < \bar{X} < x$

$d(x, y) = d(x, \bar{X}) + d(y, \bar{X}) > 0$, usando que $x \neq \bar{X}$ e a positividade de d .

Os casos simétricos são semelhantes.

d3) Imediata pela definição.

d4)

Caso 1: $z \leq x$ e $y \leq \bar{X}$

Temos,

$$d(x, z) = d(z, \bar{X}) - d(x, \bar{X});$$

$$d(x, y) = d(x, \bar{X}) - d(y, \bar{X});$$

$$d(y, z) = d(z, \bar{X}) - d(y, \bar{X}).$$

Daí,

$$\begin{aligned}
d(x, z) &= d(z, \overline{X}) - d(x, \overline{X}) \\
&= d(z, \overline{X}) - d(y, \overline{X}) + d(y, \overline{X}) - d(x, \overline{X}) \\
&= d(y, z) - d(x, y) \\
&\leq d(x, y) + d(y, z)
\end{aligned}$$

Caso 2: $x \leq z \leq y \leq \overline{X}$

Temos,

$$\begin{aligned}
d(x, z) &= d(x, \overline{X}) - d(z, \overline{X}); \\
d(x, y) &= d(x, \overline{X}) - d(y, \overline{X}); \\
d(y, z) &= d(z, \overline{X}) - d(y, \overline{X}).
\end{aligned}$$

Daí,

$$\begin{aligned}
d(x, z) &= d(x, \overline{X}) - d(z, \overline{X}) \\
&= d(x, \overline{X}) - d(y, \overline{X}) + d(y, \overline{X}) - d(z, \overline{X}) \\
&= d(x, y) - d(y, z) \\
&\leq d(x, y) + d(y, z)
\end{aligned}$$

Caso 3: $x \leq y \leq z \leq \overline{X}$

Temos,

$$\begin{aligned}
d(x, z) &= d(x, \overline{X}) - d(z, \overline{X}); \\
d(x, y) &= d(x, \overline{X}) - d(y, \overline{X}); \\
d(y, z) &= d(y, \overline{X}) - d(z, \overline{X}).
\end{aligned}$$

Daí,

$$\begin{aligned}
d(x, z) &= d(x, \overline{X}) - d(z, \overline{X}) \\
&= d(x, \overline{X}) - d(y, \overline{X}) + d(y, \overline{X}) - d(z, \overline{X}) \\
&= d(x, y) + d(y, z)
\end{aligned}$$

Caso 4: $x \leq y \leq \bar{X} \leq z$

Temos,

$$d(x, z) = d(x, \bar{X}) + d(z, \bar{X});$$

$$d(x, y) = d(x, \bar{X}) - d(y, \bar{X});$$

$$d(y, z) = d(y, \bar{X}) + d(z, \bar{X}).$$

Daí,

$$\begin{aligned}d(x, z) &= d(x, \bar{X}) + d(z, \bar{X}) \\ &= d(x, \bar{X}) - d(y, \bar{X}) + d(y, \bar{X}) + d(z, \bar{X}) \\ &= d(x, y) + d(y, z)\end{aligned}$$

Caso 5: $z \leq x \leq \bar{X} \leq y$

Temos,

$$d(x, z) = d(z, \bar{X}) - d(x, \bar{X});$$

$$d(x, y) = d(x, \bar{X}) + d(y, \bar{X});$$

$$d(y, z) = d(z, \bar{X}) + d(y, \bar{X}).$$

Daí,

$$\begin{aligned}d(x, z) &= d(z, \bar{X}) - d(x, \bar{X}) \\ &\leq d(z, \bar{X}) + d(x, \bar{X}) \\ &\leq d(z, \bar{X}) + d(x, \bar{X}) + d(y, \bar{X}) + d(y, \bar{X}) \\ &= d(x, y) + d(y, z)\end{aligned}$$

Caso 6: $x \leq z \leq \bar{X} \leq y$

Temos,

$$d(x, z) = d(x, \bar{X}) - d(z, \bar{X});$$

$$d(x, y) = d(x, \bar{X}) + d(y, \bar{X});$$

$$d(y, z) = d(z, \bar{X}) + d(y, \bar{X}).$$

Daí,

$$\begin{aligned}d(x, z) &= d(x, \bar{X}) - d(z, \bar{X}) \\ &\leq d(x, \bar{X}) + d(z, \bar{X}) \\ &\leq d(z, \bar{X}) + d(x, \bar{X}) + d(y, \bar{X}) + d(y, \bar{X}) \\ &= d(x, y) + d(y, z)\end{aligned}$$

Caso 7: $x \leq \bar{X} \leq z \leq y$

Temos,

$$\begin{aligned}d(x, z) &= d(x, \bar{X}) + d(z, \bar{X}); \\ d(x, y) &= d(x, \bar{X}) + d(y, \bar{X}); \\ d(y, z) &= d(y, \bar{X}) - d(z, \bar{X}).\end{aligned}$$

Daí,

$$\begin{aligned}d(x, z) &= d(x, \bar{X}) + d(z, \bar{X}) \\ &= d(x, \bar{X}) - d(y, \bar{X}) + d(y, \bar{X}) + d(z, \bar{X}) \\ &= d(x, y) - d(y, z) \\ &\leq d(x, y) + d(y, z)\end{aligned}$$

Caso 8: $x \leq \bar{X} \leq y \leq z$

Temos,

$$\begin{aligned}d(x, z) &= d(x, \bar{X}) + d(z, \bar{X}); \\ d(x, y) &= d(x, \bar{X}) + d(y, \bar{X}); \\ d(y, z) &= d(z, \bar{X}) - d(y, \bar{X}).\end{aligned}$$

Daí,

$$\begin{aligned}d(x, z) &= d(x, \bar{X}) + d(z, \bar{X}) \\ &= d(x, \bar{X}) - d(y, \bar{X}) + d(y, \bar{X}) + d(z, \bar{X}) \\ &= d(x, y) + d(y, z)\end{aligned}$$

Definição 2.11. A *Variância-Utilidade* (VAR_U) é o número real positivo dado por

$$VAR_U = \frac{1}{n} \sum_{i=1}^n d(X_i, \bar{X}), \quad (2.5)$$

onde n é a quantidade de observações para as quais a função-utilidade está bem definida, isto é, as observações em que o logaritmo existe.

Definição 2.12. O *Desvio-Utilidade* (DU) é a raiz quadrada da *Variância-Utilidade*.

Observação: Mesmo se tratando de amostras, consideramos n e não $n - 1$. A teoria para explicar essa possível diferenciação no denominador não faz parte dos objetivos desse trabalho.

Propriedade 2.1. Os valores X_i , tais que

$$\frac{\bar{X} - 3\sigma(e - 1)}{e} < X_i < \bar{X}e + 3\sigma(e - 1)$$

contribuem para um desvio-utilidade menor que o desvio-padrão, onde \bar{X} é a média da amostra referencial válida, σ é o desvio-padrão de referência e e é o neperiano.

Prova:

$$\begin{aligned}
& \frac{\bar{X}-3\sigma(e-1)}{e} < X_i < \bar{X}e + 3\sigma(e-1) \\
\iff & \bar{X} - 3\sigma e + 3\sigma < X_i e < \bar{X}e^2 + 3\sigma e(e-1) \\
\iff & \bar{X} + 3\sigma < (X_i + 3\sigma)e < (\bar{X} + 3\sigma)e^2 \\
\iff & \frac{1}{e} < \frac{X_i+3\sigma}{\bar{X}+3\sigma} < e \\
\iff & \ln\left(\frac{1}{e}\right) < \ln\left(\frac{X_i+3\sigma}{\bar{X}+3\sigma}\right) < \ln e \\
\iff & -1 < \ln\left(\frac{X_i+3\sigma}{\bar{X}+3\sigma}\right) < 1 \\
\iff & 0 < \ln^4\left(\frac{X_i+3\sigma}{\bar{X}+3\sigma}\right) < 1
\end{aligned}$$

Daí, para todo X_i satisfazendo à condição enunciada, teremos

$$\begin{aligned}
& (X_i - \bar{X})^2 \ln^4\left(\frac{X_i+3\sigma}{\bar{X}+3\sigma}\right) \leq (X_i - \bar{X})^2 \\
\implies & \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2 \ln^4\left(\frac{X_i + 3\sigma}{\bar{X} + 3\sigma}\right) \leq \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2
\end{aligned}$$

Tirando a raiz quadrada, temos que o desvio-padrão é cota superior para o desvio-utilidade.

Propriedade 2.2. Para $\bar{X} > 3\sigma\left(\frac{2-e}{e-1}\right)$, $\bar{X}e + 3\sigma(e-1)$ é sempre discrepante pelo critério do Escore Z^8 . Para $\bar{X} > \frac{3\sigma}{e-1}$, $\frac{\bar{X}-3\sigma(e-1)}{e}$ é discrepante.

Prova: Basta comparar $\bar{X}e + 3\sigma(e-1)$ com $\bar{X} + 3\sigma$ e $\frac{\bar{X}-3\sigma(e-1)}{e}$ com $\bar{X} - 3\sigma$.

Observação: Como buscamos ativos, cujos retornos médios são positivos, ou pouco negativos, espera-se $\bar{X} > 3\sigma\frac{2-e}{e-1}$ discrepante, na grande maioria das amostras.

Comentários ao modelo proposto: verificação das condições de Damodaran [16] (subseção 2.1.2)

⁸Para uma amostra com comportamento normal, o critério do Escore Z diz que uma observação será discrepante se estiver afastada mais do que três desvios-padrão da média, já que sua chance de ocorrência se restringe a 0,24% (ver [101]).

A medida de risco obtida via desvio-utilidade, claramente, apenas leva em conta o risco de mercado, porém se aplica a qualquer ativo, desde que se consiga um desvio-padrão de referência, satisfazendo, portanto aos requisitos (1) e (2) da subseção 2.1.2. Como a unidade de desvio-utilidade é a unidade original dos retornos, podemos entender a informação resultante como uma "taxa de retorno" cobrada para se assumir risco, com o adicional de se considerar as utilidades marginais de perdas e ganhos. Como a distribuição de frequências das oscilações para a metodologia proposta é induzida pela distribuição de frequências para a metodologia usual (aquela que supõe o risco em termos de desvio-padrão), supor que esta última tem tendência normal induz a inferência que nos permite avaliar as probabilidades para futuras oscilações via desvio-utilidade. Assim, os requisitos (4) e (5) podem ser considerados como satisfeitos. Para completarmos os requisitos de um bom modelo, satisfazendo o item (3), definimos o risco médio como o desvio-utilidade do mercado no qual estão inseridos os ativos em estudo. No caso de ativos negociados na BOVESPA, as medidas de risco estariam padronizadas pelo desvio-utilidade do índice Ibovespa para o período em questão.

2.2.3 Resultados experimentais : σ versus DU

Analizamos as variações percentuais diárias de 03/01/2002 a 27/09/2002 (Fonte: Economática) de 67 empresas cujas ações são negociadas na BOVESPA, considerando o desvio-padrão do índice Ibovespa para o período em questão ($\sigma = 1,9989\%a.d.$) como desvio-padrão de referência. Os resultados, em ordem crescente de risco, para cada uma das metodologias, estão apresentados na tabela a seguir:

Desvio-Padrão	Desvio-Utilidade
1 Banespa PN 0,6537%	1 Banespa PN 0,0955%
2 Unipar PNB 1,6697%	2 Sadia SA PN 0,1350%
3 Confab PN 1,7168%	3 Fosfertil PN 0,1520%
4 Sadia SA PN 1,7684%	4 Tele Sudeste Celula PN 0,1826%
5 Perdigão PN 1,7751%	5 Perdigão PN 0,3346%
6 Tele Sudeste Celula PN 1,8041%	6 Ripasa PN 0,4899%
7 Fosfertil PN 1,8518%	7 Ultrapar PN 0,5014%
8 Caemi Metal PN 1,8690%	8 Unipar PNB 0,5609%
9 Embraco PN 1,8733%	9 Petroquisa PN 0,6294%
10 Marcopolo PN 1,9261%	10 Unibanco PN 0,8050%
11 Ambev PN 1,9517%	11 Confab PN 0,8425%
12 Magnesita PNA 1,9699%	12 Magnesita PNA 0,9966%
13 Gerdau Met PN 1,9940%	13 Forjas Taurus PN 1,0793%
14 Suzano PN 1,9962%	14 Embraco PN 1,3075%
15 Vale do Rio Doce PNA 2,0002%	15 Caemi Metal PN 1,3338%
16 Duratex PN 2,0121%	16 Ambev PN 1,4052%
17 Ultrapar PN 2,0471%	17 Lojas Americanas PN 1,4983%

18 Forjas Taurus PN 2,1244%	18 Vale do Rio Doce PNA 1,5602%
19 Souza Cruz ON 2,1245%	19 Suzano PN 1,6695%
20 Ripasa PN 2,1376%	20 Brasil Telecom PN 1,8188%
21 Votorantim C P PN 2,2213%	21 Ferbasa PN 2,0446%
22 Petrobrás PN 2,2459%	22 Marcopolo PN 2,1467%
23 Embraer PN 2,2801%	23 Coteminas PN 2,2244%
24 Itausa PN 2,3529%	24 Belgo Mineira PN 2,3232%
25 Ferbasa PN 2,3777%	25 Varig PN 2,7684%
26 Santista Têxtil PN 2,3870%	26 Duratex PN 3,1708%
27 Belgo Mineira PN 2,4061%	27 Votorantim C P PN 3,3270%
28 Coteminas PN 2,4266%	28 Itaubanco PN 3,6648%
29 Tractebel PNB 2,4568%	29 Itausa PN 3,6741%
30 Gerdau PN 2,4901%	30 Telemar PN 3,8502%
31 Pão de Açúcar PN 2,4957%	31 Souza Cruz ON 4,0169%
32 Bradesco PN 2,5481%	32 Emae PN 4,2870%
33 Sid Tubarão PN 2,5668%	33 Sid Nacional ON 4,3390%
34 Itaubanco PN 2,6454%	34 Sanepar PN 4,3514%
35 Lojas Americanas PN 2,6870%	35 Telefonica BDR 4,4362%
36 Unibanco PN 2,7084%	36 Petrobrás PN 4,4694%
37 Sabesp ON 2,7122%	37 Polialden PN 4,6685%
38 Telemar PN 2,7639%	38 Bradesco PN 4,7571%
39 Telefonica BDR 2,7760%	39 Gerdau Met PN 4,8545%
40 Brasil Telecom PN 2,7833%	40 Sabesp ON 4,9614%
41 Aracruz PNB 2,8023%	41 Embraer PN 5,1321%
42 Polialden PN 2,8941%	42 Light ON 5,5652%
43 Copel PNB 2,9045%	43 Cemig PN 5,6142%
44 Cemig PN 2,9384%	44 Trikem PN 5,6193%
45 Light ON 3,0443%	45 Copel PNB 6,2859%
46 Usiminas PNA 3,0476%	46 Pão de Açúcar PN 6,3317%
47 Sanepar PN 3,0562%	47 Tele Centroeste Cel PN 6,6301%
48 Eletrobrás PNB 3,0744%	48 Sid Tubarão PN 6,6434%
49 Bombril PN 3,1053%	49 Usiminas PNA 6,6455%
50 Petroquisa PN 3,1094%	50 Klabin PN 7,5690%
51 Cosipa PN 3,1469%	51 Ipiranga Pet PN 8,0784%
52 Ipiranga Pet PN 3,2339%	52 Telebrás RCTB ADR 8,1385%
53 Emae PN 3,2699%	53 Aracruz PNB 8,2813%
54 Trikem PN 3,4117%	54 Tractebel PNB 9,1407%
55 Sid Nacional ON 3,4145%	55 Cosipa PN 9,3479%
56 Tele Celular Sul PN 3,4497%	56 Gerdau PN 9,8575%
57 Tele Nordeste Celul PN 3,4664%	57 Telemig Celul Part PN 10,1719%
58 Cesp PN 3,5065%	58 Eletrobrás PNB 10,3858%

59 Telemig Celul Part PN 3,5348%	59 Eletropaulo Metro po PN 10,6552%
60 Telebrás RCTB ADR 3,5791%	60 Santista Têxtil PN 10,7680%
61 Klabin PN 3,6317%	61 Cesp PN 11,3531%
62 Tele Centroeste Cel PN 3,7489%	62 Tele Celular Sul PN 11,6175%
63 Eletropaulo Metro po PN 4,6381%	63 Bombril PN 12,1316%
64 Rhodia-Ster ON 4,9046%	64 Tele Nordeste Celul PN 15,3073%
65 Varig PN 5,3138%	65 Embratel Part PN obs.
66 Embratel Part PN 5,7501%	66 Net PN obs.
67 Net PN 6,3457%	67 Rhodia-Ster ON obs.

(Tabela 2.1: Desvios padrão e utilidade ao dia.)

Observação: As amostras dos ativos Embratel Part PN, Net PN e Rhodia-Ster ON atenderam à condição (1) da definição 2.8, mas não satisfizeram à condição (2), não constituindo, portanto, uma Amostra Referencial Válida para o período em estudo: no caso da Embratel Part PN, apenas 87,50% das observações satisfaziam a segunda condição, enquanto que para a Net PN e para a Rhodia-Ster ON, apenas 81,77% e 85,94%, respectivamente.

Para uma melhor visualização das mudanças ocorridas nas posições, podemos agrupar os resultados da tabela anterior em classes definidas por variações na classificação:

Variação na classificação (n ^o de posições)	Ocorrências
[-35; -21]	3
[-20; -5]	21
[-4; 1]	6
0	3
[1; 4]	11
[5; 20]	14
[21; 35]	5
Mais que 35	1
TOTAL	64

Observação: O ativo que sofreu maior variação positiva na classificação foi o Varig PN: passou da 65^a para 25^a posição. O ativo Santista Têxtil PN sofreu o maior

efeito contrário: caiu da 26^a para 60^a posição.

2.2.4 Correlação-Utilidade

Definição 2.13. A *Covariância-Utilidade* entre duas amostras referenciais válidas A e B é dada por

$$COV_U(A; B) = \frac{1}{n} \sum_{i=1}^n \left(X_i^A - \overline{X^A} \right) \ln^2 \left(\frac{X_i^A + 3\sigma}{\overline{X^A} + 3\sigma} \right) \times \left(X_i^B - \overline{X^B} \right) \ln^2 \left(\frac{X_i^B + 3\sigma}{\overline{X^B} + 3\sigma} \right), \quad (2.6)$$

onde $X_i^A \in A$ é uma observação e $\overline{X^A}$ é a média da amostra referencial válida A , $X_i^B \in B$ é uma observação e $\overline{X^B}$ é a média da amostra referencial válida B , σ é o desvio-padrão amostral (ou populacional) de referência e n é o número de produtos bem definidos (para os quais a função utilidade é definida em ambas as amostras).

Observação: Como apenas consideramos amostras referenciais válidas, temos garantido um valor para n no mínimo igual a 81% do número de observações.

Definição 2.14. A *Correlação-Utilidade* entre duas amostras referenciais válidas A e B é dada por

$$COR_U(A; B) = \frac{COV_U(A; B)}{DU_A \times DU_B} \quad (2.7)$$

onde $COV_U(A; B)$ é a covariância-utilidade entre as amostras referenciais válidas A e B e DU_A e DU_B são os seus respectivos desvios-utilidade.

Proposição 2.2. A *Correlação-Utilidade* é um coeficiente limitado inferiormente por -1 e superiormente por 1 .

Prova: Conseqüência imediata da desigualdade de Cauchy-Schwartz.

Definição 2.15. A *Variância-Utilidade* de uma carteira formada por dois ativos A e B é dada por:

$$VAR_U = p_A^2 DU_A^2 + p_B^2 DU_B^2 + 2p_A p_B DU_A DU_B COR_U(A; B) \quad (2.8)$$

e seu Desvio-Utilidade por:

$$DU = \sqrt{VAR_U}, \quad (2.9)$$

onde p_A , p_B , DU_A e DU_B são, respectivamente, as ponderações e os desvios-utilidade de A e B e $COR_U(A;B)$ é a correlação-utilidade entre esses ativos.

Para melhor visualizarmos as diferenças, calculamos a correlação usual e a correlação-utilidade para os ativos Belgo-Mineira PN e Embraer PN, que, respectivamente, subiu 3 posições (Tabela 2.1: coluna 1, linha 27 para coluna 2, linha 24) e caiu 19 posições (Tabela 2.1: coluna 1, linha 23 para coluna 2, linha 41) devido à mudança de metodologia para cálculo de risco. Em seguida, calculamos o par risco/retorno de uma carteira formada por esses dois ativos, para diferentes ponderações, segundo as duas metodologias. Os resultados estão expostos a seguir

	Belgo Mineira PN	Embraer PN
Retorno Esperado diário (histórico)	0,4183%	0,1240%
Desvio-Padrão diário (tabela 2.1)	2,4061%	2,2801%
Correlação		0,0157
Desvio-Utilidade diário (def. 2.12)	2,3232%	5,1321%
Correlação-Utilidade (def. 2.14)		0,0100

Observação: O retorno esperado (diário) é a média dos retornos diários na amostra referencial válida coletada. O desvio-padrão e a correlação são calculados da maneira usual (ver [101]). O desvio-utilidade e a correlação-utilidade são calculados de acordo com a definições apresentadas anteriormente.

p_i Belgo-Mineira PN	p_i Embraer PN	Retorno Carteira	σ Carteira	DU Carteira
0%	100%	0,1240%	2,2801	5,1321%
10%	90%	0,1535%	2,0699%	4,6271%
20%	80%	0,1829%	1,8938%	4,1365%
30%	70%	0,2123%	1,7620%	3,6663%
40%	60%	0,2417%	1,6850%	3,2253%
50%	50%	0,2712%	1,6703%	2,8273%
60%	40%	0,3006%	1,7197%	2,4929%
70%	30%	0,3300%	1,8278%	2,2506%
80%	20%	0,3595%	1,9851%	2,1321%
90%	10%	0,3889%	2,1810%	2,1579%
100%	0%	0,4183%	2,4061%	2,3232%

Observação: Os retornos, os desvios-padrão e os desvios-utilidade da carteira são calculados pelas fórmulas apresentadas nas definições 2.3, 2.4 e 2.15, respectivamente.

A conclusão imediata dessa tabela é que a ponderação ótima (em termos de risco) para uma carteira será diferente para cada uma das abordagens: a carteira de mínimo risco medido pelo desvio-padrão dificilmente conterá os ativos nos mesmos percentuais que aquela baseada em desvio-utilidade.

2.3 Um novo modelo para Otimização de Carteira de Investimentos

O objetivo dessa seção é apresentar um modelo para uma carteira de investimentos que maximize a utilidade, levando em conta o retorno, o risco e a maior aversão a desvios negativos do que a desvios positivos. Veremos que o proposto é um modelo comum para otimização de carteiras acrescido da componente desvio-utilidade definida na seção anterior.

2.3.1 Modelo Não-Linear Inteiro

A montagem de uma carteira de investimentos, geralmente, considera um conjunto de ativos para os quais já se conhece o retorno médio esperado, o risco associado a esse retorno e a correlação entre cada um deles. De acordo com Ross [76], esses dados podem, simplesmente, traduzir a expectativa do indivíduo, baseada em uma série histórica de retornos, ou, podem ser resultado de uma análise detalhada das perspectivas de uma empresa. Assim, a otimização de uma carteira se resume a definir quais ativos a comporão e quais as suas ponderações.

Consideremos o seguinte problema de Otimização⁹:

$$\max f(R_{Cart}, DU_{Cart})$$

s.a:

$$\sum_{i=1}^N x_i y_i \leq C_0 \quad (2.10)$$

$$b_i y_i \leq x_i \leq C_0 y_i \quad \forall i$$

$$x_i \in \mathbb{R}_+; y_i \in \{0, 1\} \quad \forall i$$

onde

- R_{Cart} e DU_{Cart} são funções de $\mathbb{R}^n \rightarrow \mathbb{R}$ que representam respectivamente o retorno e o desvio-utilidade da carteira de investimentos;
- f é uma função utilidade dada por:

$$f(u, v) = \rho_1 u - \rho_2 v^2 \quad (2.11)$$

⁹O modelo apresentado resultou da colaboração do Prof. Antônio Marcos Duarte Jr., do IBMEC-RJ.

onde ρ_1 e ρ_2 , com $\rho_1 + \rho_2 = 1$ são parâmetros dados. Elevamos a segunda variável ao quadrado para trabalharmos com a variância-utilidade da carteira, o que não altera o objetivo do modelo.

- $b_i < C_0$ são dados.

Usando as definições 2.1, 2.8 e 2.9, temos o seguinte problema:

$$\begin{aligned} & \max \left[\rho_1 \sum_{i=1}^N R_i x_i \right. \\ & \left. - \rho_2 \left(\sum_{i=1}^N x_i^2 DU_i^2 + \sum_{i=1}^N \sum_{j \neq i} x_i DU_i x_j DU_j COR_U(i, j) \right) \right] \\ \text{s. a:} & \end{aligned} \tag{2.12}$$

$$\sum_{i=1}^N x_i y_i \leq C_0$$

$$b_i y_i \leq x_i \leq C_0 y_i \quad \forall i$$

$$x_i \in \mathbb{R}_+; y_i \in \{0, 1\} \quad \forall i$$

onde R_i , DU_i e $COR_U(i, j)$ são dados previamente calculados, obtidos de amostras de retornos passados dos ativos e N é o número de ativos.

2.3.2 Explicação do Modelo

O problema considera N ativos para os quais são conhecidos as taxas de retorno esperadas, R_i , os desvios-utilidade, DU_i , as correlações-utilidade, $COR_U(i, j)$, entre cada par de ativos, a participação mínima, b_i , em termos monetários, do ativo, caso este faça parte da carteira, o capital total disponível para investimento, C_0 , e os coeficientes ρ_1 e ρ_2 .

As variáveis do problema podem ser separadas em dois grupos: x_i , variáveis de ponderação, e y_i , variáveis de alocação. As variáveis x_i são contínuas, podendo assumir qualquer valor não-negativo. Elas indicam a participação, em relação ao capital disponível, de cada ativo na carteira. As variáveis y_i são binárias, assumindo valor unitário, caso o ativo componha a carteira, e nulo, caso contrário.

O objetivo do modelo é maximizar a utilidade da carteira, considerando, para tanto, variações no par risco/retorno, conseqüências das diferentes ponderações de cada ativo que a compõe. Os coeficientes ρ_1 e ρ_2 indicarão o perfil do investidor. Quanto mais próximo de um estiver o parâmetro ρ_1 , mais agressivo será o investidor, pois priorizará a maximização do retorno, em detrimento da minimização do risco. Em contrapartida, quanto mais próximo de zero estiver esse parâmetro, mais conservador será o investidor, pois preferirá correr menos riscos, mesmo que o retorno se estabeleça em níveis mais baixos. No caso de termos $\rho_1 = \rho_2 = 0,5$, diremos que o investidor tem perfil moderado.

A primeira restrição (Disponibilidade de Capital) impede que o capital utilizado ultrapasse a disponibilidade. As demais restrições delimitam inferiormente a participação de cada ativo. Apenas terão participação nula, aqueles ativos que não compuserem a carteira. A maior participação de um ativo é, obviamente, o total de capital disponível.

2.3.3 Análise do Modelo

De uma maneira geral, quando um modelo tem por objetivo a maximização do retorno, independente do risco, a disponibilidade de capital é uma restrição de igualdade. No contexto financeiro, a razão para a igualdade é a certeza de que todo o capital disponível deve ser utilizado na solução ótima, a menos que os retornos

individuais de todos os ativos fossem negativos. Em outras palavras, a restrição em questão é ativa no ótimo.

Como nosso modelo objetiva a maximização da utilidade, que depende do perfil do investidor quanto ao par risco/retorno, não podemos concluir que a restrição continuará ativa no ótimo. Esse fato, em conjunto com a relaxação contínua da variável y , determinam uma região viável não convexa para o problema, o que indica possível dificuldade na resolução do problema inteiro.

Como $y_i \in \{0, 1\}, \forall i$, temos, do segundo conjunto de restrições, que $0 \leq x_i \leq C_0, \forall i$ (já que $b_i \geq 0, \forall i$). Assim, o conjunto de soluções viáveis é limitado e não vazio (a origem é uma solução viável). Portanto, a continuidade da função objetivo garante a existência de uma solução.

Capítulo 3

Metodologia de Programação Geométrica Signomial 0-1

3.1 Preliminares

Nosso objetivo, neste capítulo, é usar a Programação Geométrica Signomial¹. Para obtermos soluções 0-1 nas variáveis inteiras, a partir da solução contínua, utilizamos uma função baseada na distribuição gama, que será apresentada na seção 3.4. Essa função permitirá que as relaxações contínuas das variáveis inteiras (binárias) assumam valores próximos de 0 ou de 1. Nessa mesma seção, apresentaremos o algoritmo Barreira-Penalidade, utilizado para resolver os sucessivos problemas posinomiais. Na seção 3.2, apresentamos algumas idéias centrais, através de definições e resultados essenciais, visando um melhor entendimento da metodologia de programação geométrica signomial. Na seção 3.3 constam alguns trabalhos que consideramos representativos em métodos de pontos interiores, em otimização combinatoria e em programação geométrica.

Os problemas de programação geométrica (PG) podem ser divididos em dois grupos: PG Posinomial e PG Signomial. O primeiro pode ser visto através da Teoria

¹Este capítulo resulta de colaboração com o Prof. Roberto Quirino do Nascimento, da UFPB.

de Dualidade como um problema convexo com restrições lineares. No apêndice A, consta uma apresentação dessa família de problemas.

3.2 Programação Geométrica Signomial

Definição 3.1. *Um problema de programação geométrica é chamado de Problema de Programação Geométrica Signomial [21], quando um ou mais termos posinomiais (ver definição no apêndice A) no problema primal são negativos.*

A ocorrência de termos posinomiais com coeficientes positivos e negativos implica na necessidade de uma versão expandida para o problema dual [21].

Consideremos o seguinte problema primal

$$\min g_0(t) = \sum_{i \in J[0]} \sigma_{0i} c_{0i} \prod_{j=1}^m t_j^{a_{0ij}}$$

s.a.

$$g_k(t) = \sum_{i \in J[k]} \sigma_{ki} c_{ki} \prod_{j=1}^m t_j^{a_{kij}} \leq \sigma_k, \quad k = 1, \dots, p \quad (3.1)$$

$$t_j > 0, \quad j = 1, \dots, m$$

onde

$$c_{ki} > 0, \quad k = 0, 1, \dots, p, \quad i \in J[k]$$

$$\sigma_k = \pm 1, \quad k = 1, \dots, p$$

$$\sigma_{ki} = \pm 1, \quad k = 0, \dots, p, \quad i \in J[k]$$

$$J[k] = \{m_k, m_{k+1}, \dots, n_k\}, \quad k = 0, 1, \dots, p$$

$$m_0 = 1, \quad m_1 = n_0 + 1, \quad m_2 = n_1 + 1, \quad \dots, \quad m_p = n_{p-1} + 1, \quad n_p = n,$$

Observemos que a formulação do problema acima introduz $J[k] + p$ variáveis adicionais em relação ao problema posinomial (ver Apêndice A). As $J[k]$ variáveis, definidas por σ_{ki} , $k = 0, 1, \dots, p$, $i \in J[k]$, são introduzidas de modo a absorver o sinal de cada termo, enquanto que as p variáveis σ_k , $k = 1, 2, \dots, p$, são usadas para generalizar (para ± 1) o lado direito da restrição. Essas variáveis são conhecidas por funções sinal.

Essa versão expandida do problema posinomial introduz uma notação indispensável para identificar cada componente no problema: de modo a evitar confusão, cada símbolo é identificado a uma equação, um termo e uma variável de decisão.

Com base no que vale para problemas posinomiais, nosso objetivo, nesse momento, é explorar sua forma especial, para obter um problema equivalente com restrições lineares, que é mais fácil de se resolver. O desenvolvimento do processo se dará em duas fases: primeiro, o caso em que todas as funções sinal na função objetivo são positivas. Em seguida, abordaremos o caso geral. Nosso texto está baseado em [7].

3.2.1 Função Objetivo Posinomial

Desde que todas as funções sinal no objetivo são positivas, podemos desconsiderar os fatores σ_{0i} , obtendo do problema 3.1:

$$\min g_0(t) = \sum_{i \in J[0]} c_{0i} \prod_{j=1}^m t_j^{a_{0ij}}$$

s.a.

$$g_k(t) = \sum_{i \in J[k]} \sigma_{ki} c_{ki} \prod_{j=1}^m t_j^{a_{kij}} \leq \sigma_k, \quad k = 1, \dots, p \quad (3.2)$$

$$t_j > 0, \quad j = 1, \dots, m$$

onde

$$c_{ki} > 0, \quad k = 0, 1, \dots, p, \quad i \in J[k]$$

$$\sigma_k = \pm 1, \quad k = 1, \dots, p$$

$$\sigma_{ki} = \pm 1, \quad k = 1, \dots, p, \quad i \in J[k]$$

$$J[k] = \{m_k, m_{k+1}, \dots, n_k\}, \quad k = 0, 1, \dots, p$$

$$m_0 = 1, \quad m_1 = n_0 + 1, \quad m_2 = n_1 + 1, \quad \dots, \quad m_p = n_{p-1} + 1, \quad n_p = n,$$

Observemos que, como $\sigma_k = \pm 1$, as restrições podem ser escritas como

$$g_k(t) = g_k = \sigma_k \left[1 - \sigma_k \sum_{i \in J[k]} \sigma_{ki} c_{ki} \prod_{j=1}^m t_j^{a_{kij}} \right] \geq 0, \quad k = 1, \dots, p. \quad (3.3)$$

Sejam

$$\omega_{0i} = \frac{c_{0i} \prod_{j=1}^m t_j^{a_{0ij}}}{g_0(t)}, \quad i = 1, 2, \dots, J[0] \quad (3.4)$$

os pesos para cada termo na função objetivo. Definamos os pesos para os termos das restrições como seus próprios valores absolutos.

$$\omega_{ki} = c_{ki} \prod_{j=1}^m t_j^{a_{kij}}, \quad k = 1, \dots, p, \quad i = 1, \dots, J[k]. \quad (3.5)$$

Observemos que, no caso da restrição ser ativa no ótimo, a soma das contribuições (pesos) de cada termo (multiplicado por σ_{ki}) deve ser igual a σ_k .

Por conveniência, denotemos $g_0(t)$ por t_0 . Usando as equações 3.3, 3.4 e 3.5, o problema original pode ser escrito como:

$$\min t_0$$

s.a.

$$t_0 \omega_{0i} = c_{0i} \prod_{j=1}^m t_j^{a_{0ij}}, \quad i = 1, 2, \dots, J[0]$$

$$g_k = \sigma_k \left[1 - \sigma_k \sum_{i \in J[k]} \sigma_{ki} c_{ki} \omega_{ki} \right] \geq 0, \quad k = 1, \dots, p \quad (3.6)$$

$$\omega_{ki} - c_{ki} \prod_{j=1}^m t_j^{a_{kij}} = 0$$

$$1 - \sum_{i=1}^{J[0]} \omega_{0i} = 0$$

Definindo $t_j = e^{u_j}$, $j = 0, 1, \dots, m$, o problema 3.6 se transforma em

$$\min e^{u_0}$$

s.a.

$$e^{u_0} \frac{\omega_{0i}}{c_{0i}} = \exp \left[\prod_{j=1}^m a_{0ij} u_j \right], \quad i = 1, 2, \dots, J[0]$$

$$g_k = \sigma_k \left[1 - \sigma_k \sum_{i \in J[k]} \sigma_{ki} c_{ki} \omega_{ki} \right] \geq 0, \quad k = 1, \dots, p \quad (3.7)$$

$$\frac{\omega_{ki}}{c_{ki}} = \exp \left[\prod_{j=1}^m a_{kij} u_j \right]$$

$$1 - \sum_{i=1}^{J[0]} \omega_{0i} = 0$$

Tomando o logaritmo natural, temos

$$\min u_0$$

s.a.

$$\prod_{j=1}^m a_{0ij} u_j - u_0 = \ln \left[\frac{\omega_{0i}}{c_{0i}} \right], \quad i = 1, 2, \dots, J[0]$$

$$g_k = \sigma_k \left[1 - \sigma_k \sum_{i \in J[k]} \sigma_{ki} c_{ki} \omega_{ki} \right] \geq 0, \quad k = 1, \dots, p$$

$$\prod_{j=1}^m a_{kij} u_j = \ln \left[\frac{\omega_{ki}}{c_{ki}} \right]$$

$$\sum_{i=1}^{J[0]} \omega_{0i} = 1$$

e reordenando as restrições,

$$\min u_0$$

s. a.

$$\sum_{i=1}^{J[0]} \omega_{0i} = 1$$

$$\prod_{j=1}^m a_{0ij} u_j - u_0 = \ln \left[\frac{\omega_{0i}}{c_{0i}} \right], \quad i = 1, 2, \dots, J[0] \quad (3.8)$$

$$\prod_{j=1}^m a_{kij} u_j = \ln \left[\frac{\omega_{ki}}{c_{ki}} \right]$$

$$g_k = \sigma_k \left[1 - \sigma_k \sum_{i \in J[k]} \sigma_{ki} c_{ki} \omega_{ki} \right] \geq 0, \quad k = 1, \dots, p$$

Como, por definição, os pesos são não negativos, as variáveis u_0, u_1, \dots, u_m são irrestritas com relação a sinal.

Sejam $\lambda_0, \delta_{0i}, \delta_{ki}$ e λ_i os multiplicadores de Lagrange para as restrições na ordem em que elas se apresentam. O Lagrangeano do problema 3.8 é dado por:

$$\begin{aligned} L(u, \omega, \lambda, \delta) = & u_0 - \lambda_0 \left[1 - \sum_{i=1}^{J[0]} \omega_{0i} \right] - \sum_{i=1}^{J[0]} \delta_{0i} \left[\ln \frac{\omega_{0i}}{c_{0i}} + u_0 - \sum_{j=1}^m a_{0ij} u_j \right] \\ & - \sum_{k=1}^p \sum_{i=1}^{J[k]} \delta_{ki} \left[\ln \frac{\omega_{ki}}{c_{ki}} - \sum_{j=1}^m a_{kij} u_j \right] - \sum_{k=1}^p \lambda_k \sigma_k \left(1 - \sigma_k \sum_{i=1}^{J[k]} \sigma_{ki} \omega_{ki} \right) \end{aligned} \quad (3.9)$$

As condições necessárias para um ponto estacionário primal são:

$$\frac{\delta L}{\delta u_0} = 1 - \sum_{i=1}^{J[0]} \delta_{0i} = 0, \quad (3.10)$$

$$\frac{\delta L}{\delta \omega_{0i}} = \lambda_0 - \frac{\delta_{0i}}{\omega_{0i}} = 0, \quad i = 1, \dots, J[0]. \quad (3.11)$$

$$\lambda_k g_k(t) = 0, \quad k = 1, \dots, p, \quad (3.12)$$

$$g_k(t) \geq 0$$

$$\lambda_k > 0$$

Das equações 3.10 e 3.11 em conjunto com 3.4, temos:

$$\lambda_0 = 1 \quad (3.13)$$

$$\delta_{0i} = \omega_{0i}, \quad i = 1, 2, \dots, J[0]. \quad (3.14)$$

A equação 3.14 nos garante um fato interessante: os multiplicadores de Lagrange δ_{0i} para os termos da função objetivo são simplesmente os pesos correspondentes ω_{0i} .

Como $g_0(t)$ é posinomial, os δ_{0i} são pesos em seu sentido usual, isto é:

$$0 < \delta_{0i} \leq 1$$

$$\sum_{i=1}^{J[0]} \delta_{0i} = 1$$

As derivadas com respeito aos outros pesos ω_{ki} , $k \neq 0$ são dadas por:

$$\frac{\delta L}{\delta \omega_{ki}} = -\frac{\delta_{ki}}{\omega_{ki}} + \lambda_k \sigma_k^2 \sigma_{ki} = 0$$

Portanto,

$$\delta_{ki} = \lambda_k \omega_{ki} \sigma_{ki} \quad (3.15)$$

já que $\sigma_k^2 = 1$.

Somando os dois lados da equação 3.15 em i e multiplicando por σ_k , encontramos:

$$\sigma_k \sum_{i=1}^{J[k]} \delta_{ki} = \lambda_k \sigma_k \sum_{i=1}^{J[k]} \sigma_{ki} \omega_{ki} = \lambda_k \quad (3.16)$$

pois os termos $\sigma_{ki} \omega_{ki}$ devem somar σ_k para cada restrição ativa. Além disso, se $\lambda_k \neq 0$, usando 3.15 e 3.16, podemos efetuar a divisão, obtendo

$$\omega_{ki} = \frac{\delta_{ki}}{\lambda_k \sigma_{ki}} = \delta_{ki} \left[\sigma_k \sigma_{ki} \sum_{i=1}^{J[k]} \delta_{ki} \right]^{-1}, \quad k = 1, \dots, p \quad i = 1, \dots, J[k] \quad (3.17)$$

de forma que os pesos são obtidos a partir dos multiplicadores de Lagrange δ_{ki} , calculados pela diferenciação com respeito a $u_j (j \neq 0)$:

$$\frac{\delta L}{\delta u_j} = \sum_{k=0}^p \sum_{i=1}^{J[k]} \delta_{ki} a_{kij} = 0, \quad j = 1, 2, \dots, m \quad (3.18)$$

com a hipótese de que

$$\sigma_{01}, \dots, \sigma_{0J[0]} = 1 \quad (\text{as funções sinal são positivas para } k = 0) \quad (3.19)$$

O problema dual pode ser encontrado expressando-se o Lagrangeano (3.9) em função de u_j , δ_{ki} , $g_k(t)$, usando o fato de que $\lambda_0 = 1$ (3.13):

$$\begin{aligned}
L(\delta, u, v, g) = & \sum_{k=0}^p \sum_{i=1}^{J[k]} \delta_{ki} \left[\ln \left(\frac{c_{0ki}}{\omega_{ki}} \right) \right] + (u_0 - 1) \left(1 - \sum_{i=1}^{J[k]} \delta_{0i} \right) \\
& + \sum_{j=1}^m u_j \sum_{k=0}^p \sum_{i=1}^{J[k]} a_{kij} \delta_{ki} - \sum_{k=1}^p \lambda_k g_k(t)
\end{aligned} \tag{3.20}$$

ou, equivalentemente,

$$\begin{aligned}
L(\delta, u, v, g) = & \sum_{k=0}^p \sum_{i=1}^{J[k]} \delta_{ki} \left[\ln \left(\frac{c_{0ki}}{\omega_{ki}} \right) \right] + (u_0 - 1) \left(1 - \sum_{i=1}^{J[k]} \delta_{0i} \right) \\
& + \sum_{j=1}^m u_j \sum_{k=0}^p \sum_{i=1}^{J[k]} a_{kij} \delta_{ki}
\end{aligned} \tag{3.21}$$

pois $\lambda_k g_k(t) = 0$, para $k = 1, \dots, p$.

De acordo com [95], o lagrangeano é idêntico ao lagrangeano de outro problema de otimização cuja função objetivo é

$$z(\delta) = \sum_{k=0}^p \sum_{i=1}^{J[k]} \delta_{ki} \left[\ln \left(\frac{c_{0ki}}{\omega_{ki}} \right) \right] \tag{3.22}$$

e as restrições são

$$\sum_{i=1}^{J[0]} \delta_{0i} = 1 \tag{3.23}$$

$$\sum_{k=0}^p \sum_{i=1}^{J[k]} a_{kij} \delta_{ki} = 0 \quad j = 1, \dots, m \tag{3.24}$$

onde

$$\omega_{ki} = \delta_{ki} \left[\sigma_k \sigma_{ki} \sum_{j=1}^{J[k]} \delta_{kj} \right]^{-1}$$

dado por 3.17.

Observemos que, para essa nova formulação, as variáveis δ_{ki} são as variáveis de decisão e $(u_0 - 1)$ e u_1, u_2, \dots, u_m são os multiplicadores de Lagrange. Assim, o problema equivalente tem a seguinte função objetivo:

$$\max z(\delta) = \sum_{k=0}^p \sum_{i=1}^{J[k]} \delta_{ki} \left[\ln \left(\frac{c_{0ki}}{\omega_{ki}} \right) \right] \quad (3.25)$$

e as restrições

$$\sum_{i=1}^{J[0]} \delta_{0i} = 1 \quad (3.26)$$

$$\sum_{k=0}^p \sum_{i=1}^{J[k]} a_{kij} \delta_{ki} = 0 \quad j = 1, \dots, m \quad (3.27)$$

onde

$$\omega_{ki} = \delta_{ki} \left[\sigma_k \sigma_{ki} \sum_{i=1}^{J[k]} \right]^{-1}$$

dado por 3.17.

Faz-se necessário, nesse momento, caracterizarmos as variáveis δ_{ki} . Usando 3.17, $\delta_{ki} = \omega_{ki} \sigma_{ki} \lambda_k$, dado por 3.16 e observando que $\omega_{ki} > 0$ e $\lambda_k \geq 0$, é imediato que δ_{ki}

determine o sinal de σ_{ki} . Assim, se $\sigma_{ki} = 1$, teremos $\delta_{ki} \geq 0$, caso contrário, teremos $\delta_{ki} \leq 0$.

Definamos então as novas variáveis duais

$$v_{ki} = \sigma_{ki} \delta_{ki} \geq 0 \quad (3.28)$$

Substituindo essas variáveis em 3.25, 3.26 e 3.27, temos

$$\max z(v) = \sum_{k=0}^p \sum_{i=1}^{J[k]} \sigma_{ki} v_{ki} \left[\ln \left(\frac{c_{ki} \sigma_k \sum_i v_{ki} \sigma_{ki}}{v_{ki}} \right) \right]$$

s.a.

$$\sum_{i=1}^{J[0]} v_{0i} = 1$$

$$\sum_{k=0}^p \sum_{i=1}^{J[k]} v_{ki} \sigma_{ki} a_{ki,j} = 0 \quad j = 1, 2, \dots, m$$

$$v_{ki} \geq 0 \quad i = 1, \dots, J[k]$$

Assim, temos $m + 1$ restrições lineares em variáveis tantas quanto o número de termos n na função objetivo e em todas as restrições signomiais

$$n = \sum_{k=0}^p J[k].$$

Quando não há grau de dificuldade, isto é $n - (m + 1) = 0$, a solução é única e, uma vez conhecidos os valores ótimos v_{ki}^* das variáveis duais, os pesos ótimos ω_{ki}^* podem ser calculados diretamente de 3.17 e da relação $v_{ki} = \sigma_{ki} \delta_{ki}$.

Quando há grau de dificuldade, entretanto, é necessário mais análise para encontrar esses pesos ótimos.

Por conveniência, é usual trabalharmos com a exponencial de $z(v)$, chamada de função dual e denotada por $d(v)$. Assim, definimos o problema dual de programação geométrica da seguinte forma:

$$\max d(v) = \prod_{k=0}^p \prod_{i=1}^{J[k]} \left(\frac{c_{ki} v_{k0}}{v_{ki}} \right) \sigma_{ki} v_{ki}$$

s.a.

$$\sum_{i=1}^{J[0]} v_{0i} = 1$$

$$\sum_{k=0}^p \sum_{i=1}^{J[k]} v_{ki} \sigma_{ki} a_{kij} = 0 \quad j = 1, 2, \dots, m$$

(3.29)

$$v_{k0} = \sigma_k \sum_i \sigma_{ki} v_{ki}$$

$$v_{ki} \geq 0 \quad i = 1, \dots, J[k]$$

O sinal de v_{k0} é facilmente estabelecido, pois de 3.17 temos

$$\begin{aligned}
\omega_{ki} &= \frac{\delta_{ki}}{\lambda_k \sigma_{ki}} \\
&= \frac{\sigma_{ki} \delta_{ki}}{\lambda_k \sigma_{ki}^2} \\
&= \frac{\sigma_{ki} \delta_{ki}}{\lambda_k \sigma_{ki}^2} \\
&= \frac{v_{ki}}{\lambda_k} \\
&= \frac{v_{ki}}{\sigma_k \sum_j \delta_{kj}} \\
&= \frac{v_{ki}}{\sigma_k \sum_j \sigma_{kj}^2 \delta_{kj}} \\
&= \frac{v_{ki}}{\sigma_k \sum_j \sigma_{kj} v_{kj}} \\
&= \frac{v_{ki}}{v_{k0}}
\end{aligned}$$

Assim, $v_{k0} \geq 0$ desde que ω_{ki} é positivo (por construção) e v_{ki} é não negativo (por definição). Observemos que as variáveis v_{ki} estão relacionadas com os multiplicadores de Lagrange λ_k pela equação 3.15 e $v_{ki} = \sigma_{ki} \delta_{ki}$.

Podemos, nesse ponto, obter importantes relações entre as variáveis duais e as variáveis primais originais em 3.7. De 3.17, temos

$$\omega_{ki} = \delta_{ki} \left[\sigma_k \sigma_{ki} \sum_{i=1}^{J[k]} \delta_{ki} \right]^{-1}$$

Daí,

$$\frac{v_k}{\sigma_k \sum_i \sigma_{ki} v_{ki}} = \frac{v_{ki}}{v_{k0}}$$

e

$$\omega_{0i} = v_{0i}.$$

As variáveis primais podem ser descobertas a partir das equações

$$c_{0i} \prod_{j=1}^m t_j^{a_{0ij}} = v_{0i} g_0(t) \quad i = 1, 2, \dots, J[0] \quad (3.30)$$

$$c_{ki} \prod_{j=1}^m t_j^{a_{kij}} = v_{ki} = \frac{v_{ki}}{v_{k0}} \quad k = 1, \dots, p; \quad i = 1, \dots, J[0] \quad (3.31)$$

Ainda não determinamos quando o ponto crítico do problema será um ponto de máximo, um ponto de mínimo ou um outro ponto especial para a função dual ou seu logaritmo. O que estabelecemos foi que para qualquer conjunto de variáveis duais v fazendo de u_0 um mínimo local, teremos

$$z(v^0) = u_0^0 = u_0$$

De fato, no mínimo local u_0^0 , todos os termos das restrições no lagrangeano 3.21 são iguais a zero e, para as variáveis δ_0 correspondentes (que determinam v^0), os termos das restrições duais desaparecem. Quando não há grau de dificuldade, podemos obter a solução mínima resolvendo um conjunto de restrições lineares duais (3.29) e calculando $g_0(t^0)$ a partir da relação $z(v^0) = u_0^0$ e da definição $u^0 = \ln g_0$.

Então, como $z(v^0) = \ln d(v^0)$, temos

$$g_0(t^0) = d(v^0).$$

A não negatividade de λ_k e v_{ki} , em conjunto com a complementaridade requerida em 3.12, completam as condições necessárias de KKT para um ponto estacionário. Esse ponto estacionário será um mínimo para $g_0(t)$ apenas com a hipótese inicial de

que $g_0(t)$ possui um mínimo local. Nesse caso, $g_0(t^0) = g_0(t^*)$ é o mínimo global, pois só pode haver um ponto estacionário quando não há grau de dificuldade.

Os resultados desta subseção se direcionam para três possibilidades distintas:

Possibilidade 1: quando todas as funções sinal são positivas, o mínimo de $g_0(t)$ é encontrado via maximização de $d(v)$. É o caso do problema de programação geométrica posinomial.

Possibilidade 2: quando uma ou mais funções sinal são negativas, então o ponto estacionário para o problema dual de programação geométrica também é estacionário para o problema primal. Por enquanto, diremos simplesmente isso para garantir que esse ponto estacionário é um mínimo local. Outras condições de maior ordem podem ser encontradas em [7].

Possibilidade 3: quando não há grau de dificuldade, precisamos apenas das restrições duais para obtermos a solução dual v^* , que, por sua vez, nos leva a t^* . Esse ponto será o ponto de mínimo global do problema primal se o problema primal possuir uma solução mínima.

A seguir, apresentamos um exemplo ilustrativo do que foi estudado:

Exemplo [68]

$$\min g_0(t) = t_1$$

s.a.

$$g_1(t) = 5t_1^{-1}t_2 - t_1^{-1}t_3^2t_4^4 \leq 1$$

$$g_2(t) = \frac{5}{2}t_2t_3^{-2} + \frac{3}{2}t_3^{-1}t_4 \leq 1$$

Para esse problema temos (veja problema 3.1):

$$\begin{aligned} \sigma_{01} &= 1; & \sigma_{21} &= -1; & \sigma_2 &= 1; & c_{12} &= 1; \\ \sigma_{11} &= 1; & \sigma_{22} &= 1; & c_{01} &= 1; & c_{21} &= \frac{5}{2}; \\ \sigma_{12} &= -1; & \sigma_1 &= 1; & c_{11} &= 5; & c_{22} &= \frac{3}{2}. \end{aligned}$$

Como não há grau de dificuldade, o conjunto de restrições duais é suficiente para determinarmos as variáveis duais (obtidas pela relação 3.17):

$$\begin{aligned} v_{01} &= 1 \\ v_{01} - v_{11} + v_{12} &= 0 \\ v_{11} - v_{21} &= 0 \\ -2v_{12} + 2v_{21} - v_{22} &= 0 \\ -4v_{12} + v_{22} &= 0 \end{aligned}$$

A solução única desse sistema é dada por

$$\begin{aligned} v_{01}^* &= 1 & v_{11}^* &= 1,5 & v_{12}^* &= 0,5 \\ v_{21}^* &= 1,5 & v_{22}^* &= 2 \end{aligned}$$

e

$$\begin{aligned} v_{10}^* &= 1 \geq 0 \\ v_{20}^* &= 0,5 \geq 0 \end{aligned}$$

pois $v_{k0} = \sigma_k \sum_i \sigma_{ki} v_{ki}$ como no problema 3.29.

Usando novamente 3.29, temos que a solução dual é dada por $d(v^*) = 0,796 = g_0^*(t)$. Usando 3.30 e 3.31, obtemos

$$\begin{aligned} t_1^* &= 0,796; & t_2^* &= 0,239; \\ t_3^* &= 0,446; & t_4^* &= 1,15. \end{aligned}$$

As análises complementares necessárias para estabelecermos que tais pontos são mínimos globais podem ser encontradas em [7] (pp.195-197).

3.2.2 Função Objetivo com Coeficientes Negativos

Para resolvermos problemas nos quais os termos da função objetivo podem ser negativos, precisamos conhecer de antemão quando o ótimo $g_0^*(t)$ é positivo ou negativo. Não sendo possível prever o sinal, a própria resolução computacional leva a sua determinação.

Suponhamos o problema de programação geométrica signomial dado como em 3.2:

$$\min g_0(t) = \sum_{i=1}^{J[0]} \sigma_{0i} c_{0i} \prod_{j=1}^m t_j^{a_{0ij}}$$

s. a.

$$g_k(t) = \sum_{i=1}^{J[k]} \sigma_{ki} c_{ki} \prod_{j=1}^m t_j^{a_{kij}} \leq \sigma_k \quad k = 1, \dots, p$$

$$t_j > 0 \quad j = 1, \dots, m$$

CASO 1: Custo Ótimo Positivo

Suponhamos que o valor da função objetivo signomial é maior que zero no ótimo ($g_0^*(t) > 0$). Nesse caso, podemos reformular o problema da seguinte maneira:

$$\min t_0$$

s. a.

$$\bar{g}_0(t) = \sum_{i=1}^{J[0]} \sigma_{0i} c_{0i} t_0^{-1} \prod_{j=1}^m t_j^{a_{0ij}} \leq 1 \quad (3.32)$$

$$g_k(t) = \sum_{i=1}^{J[k]} \sigma_{ki} c_{ki} \prod_{j=1}^m t_j^{a_{kij}} \leq \sigma_k \quad (3.33)$$

$$t_j > 0, \quad j = 0, 1, \dots, m$$

onde $\bar{g}_0(t) = t_0^{-1}g_0(t)$.

Observemos que essa nova formulação é a mesma de 3.2, pois o problema foi reduzido a um problema com função objetivo posinomial e restrições signomiais. Como foram acrescentados uma variável e um termo, o grau de dificuldade permaneceu inalterado.

As restrições duais, obtidas como em 3.29, associadas à variável t_0 são:

$$\begin{aligned} v_{00} &= 1 \\ v_{00} - \sigma_{01}v_{01} - \sigma_{02}v_{02} - \dots - \sigma_{0J[0]}v_{0J[0]} &= 0 \end{aligned} \tag{3.34}$$

de forma que combinando-as, temos

$$\sigma_{01}v_{01} + \sigma_{02}v_{02} + \dots + \sigma_{0J[0]}v_{0J[0]} = 1 \tag{3.35}$$

A menos das funções sinal σ_{0i} , a equação 3.35 é igual à primeira restrição do problema dual 3.29. Assim sendo, na prática, não é necessário acrescentar a restrição

$$\bar{g}_0(t) \leq 1$$

nem mesmo trocar a função objetivo por t_0 . Basta-nos trocar a restrição

$$\sum_{i=1}^{J[0]} v_{0i} = 1$$

por

$$\sum_{i=1}^{J[0]} \sigma_{0i}v_{0i} = 1$$

em 3.29. A função objetivo e as demais restrições permanecem exatamente as mesmas de 3.29.

CASO 2: Custo Ótimo Negativo

Suponhamos agora um valor esperado negativo para o custo ótimo. Para esse caso, devemos fazer:

$$t_0 \leq g_0(t) \tag{3.36}$$

Donde

$$t_0^{-1}g_0(t) \geq 1$$

ou

$$-t_0^{-1}g_0(t) \leq -1 \tag{3.37}$$

Observemos que a equação 3.37 tem a mesma forma da equação 3.32, exceto pelas $\sigma_{0i}, i = 1, \dots, J[0]$ e pelo sinal negativo em seu lado direito.

O objetivo continua sendo minimizar t_0 e as restrições ainda levam às relações

$$v_{00} = 1$$

$$v_{00} + (-1)^2 \sum_{i=1}^{J[0]} \sigma_{0i} v_{0i} = 0$$

que implica em

$$\sum_{i=1}^{J[0]} \sigma_{0i} v_{0i} = -v_{00} = -1$$

Como no caso 1, não é necessário reformular o problema 3.7 como minimização de t_0 . Basta-nos observar que a restrição

$$\sum_{i=1}^{J[0]} v_{0i} = 1$$

em 3.29 é dada por

$$\sum_{i=1}^{J[0]} \sigma_{0i} v_{0i} = -v_{00} = -1. \quad (3.38)$$

As demais restrições de 3.29 permanecem as mesmas. Entretanto, observemos que o valor de $g_0^*(t)$ no ótimo não é $d^*(v)$ dado em 3.29, pois

$$g_0^*(t) = -(t_0^*)^{-1}. \quad (3.39)$$

Como devemos ter a igualdade entre os ótimos primal e dual, esse último deve ser ajustado de modo a assegurar a congruência:

$$d(v) = -1 \left[\prod_{k=0}^p \prod_{i=1}^{J[k]} \frac{c_{ki} v_{k0}^{\sigma_{ki} v_{ki}}}{v_{ki}} \right]^{-1} \quad (3.40)$$

Tanto o caso 1 como o caso 2 podem ser resumidos em uma única formulação com a introdução de uma nova função sinal σ_0 , sinal de $g_0^*(t)$. Assim, a restrição

$$\sum_{i=1}^{J[0]} v_{0i} = 1$$

passa a ser dada por

$$\sum_{i=1}^{J[0]} \sigma_{0i} v_{0i} = \sigma_0 (= \pm 1). \quad (3.41)$$

O valor de σ_0 é geralmente conhecido para a maioria dos problemas. Mudanças em seu sinal apenas ocasionam troca de sinal para todas as variáveis duais v_{ki} . Uma vez determinadas σ_0 e essas variáveis duais, a função dual pode ser calculada de

$$d(v) = \sigma_0 \left[\prod_{k=0}^p \prod_{i=1}^{J[k]} \frac{C_{ki} v_{k0}}{v_{ki}} \right]^{\sigma_0}. \quad (3.42)$$

Essa função dual nos leva à função primal, que pode assumir valores negativos em todos os casos previamente discutidos.

Exemplo

$$\min g_0(t) = -2t_1 t_2 t_3^4 t_4^{-1} + t_2^{-1} t_3^{-1} + 5t_1^{\frac{1}{2}} t_4$$

s.a.

$$\bar{g}_1(t) = t_4^{\frac{1}{2}} - t_2^{\frac{1}{3}} t_3 \geq 1$$

$$t_1, t_2, t_3, t_4 > 0$$

Primeiramente, observemos que a restrição deve ser escrita como

$$g_1(t) = -t_4^{\frac{1}{2}} + t_2^{\frac{1}{3}} t_3 \leq -1$$

Das equações 3.32, 3.33 e 3.41, temos

$$\sigma_{01} = -1; \quad \sigma_{02} = 1; \quad \sigma_{03} = 1;$$

$$\sigma_{11} = -1; \quad \sigma_{12} = 1; \quad \sigma_1 = -1;$$

$$c_{01} = 2; \quad c_{02} = 1; \quad c_{03} = 5;$$

$$c_{11} = 1; \quad c_{12} = 1.$$

O valor de σ_0 será tomado como $\sigma_0 = 1$. Esse problema de minimização tem 5 termos e 4 variáveis primais. Assim, não há grau de dificuldade. Pelas condições de normalidade 3.41 e de ortogonalidade

$$\sum_{k=0}^p \sum_{i=1}^{J[k]} \sigma_{ki} a_{kij} v_{ki} = 0, \quad j = 1, \dots, m, \quad (3.43)$$

as restrições duais são dadas por:

$$-v_{01} + v_{02} + v_{03} = 1$$

$$-v_{01} + \frac{1}{2}v_{03} = 0$$

$$-4v_{01} - v_{02} + v_{12} = 0$$

$$v_{01} + v_{03} - \frac{1}{2}v_{11} = 0$$

$$-v_{01} - v_{02} + \frac{1}{3}v_{12} = 0$$

Resolvendo o sistema anterior, obtemos

$$v_{01}^* = \frac{2}{3}; \quad v_{02}^* = \frac{1}{3}; \quad v_{03}^* = \frac{4}{3};$$

$$v_{11}^* = 4; \quad v_{12}^* = 3.$$

Observemos que o sinal de σ_0 foi escolhido corretamente, dado a positividade das variáveis duais ótimas. Por 3.29, temos ainda que:

$$v_{10}^* = 1.$$

A equação 3.41 deve ser satisfeita no ótimo:

$$\sigma_0 = -\frac{2}{3} + \frac{1}{3} + \frac{4}{3} = 1.$$

O valor ótimo dual é dado por 3.42:

$$d(v^*) = 38,322 = g_0^*(t)$$

As variáveis primais podem ser facilmente calculadas pelas equações 3.30 e 3.31:

$$2t_1 t_2 t_3^4 t_4^{-1} = 25,548$$

$$t_1^{-1} t_3^{-1} = 12,774$$

$$5t_1^{\frac{1}{2}} t_4 = 51,096$$

$$t_4^{\frac{1}{2}} = 4$$

obtendo-se:

$$t_1^* = 0,4083 \quad t_2^* = 0,00422$$

$$t_3^* = 18,558 \quad t_4^* = 16$$

3.2.3 Técnica de Condensação

Segundo [7], existem três maneiras básicas de se abordar um problema de programação geométrica signomial: a primeira consiste em desenvolver um algoritmo que resolva diretamente o problema primal [71], [88]; pela segunda abordagem, aproveitamos as vantagens da estrutura única do dual do problema de programação

geométrica [5], [41], [69], [8]. Finalmente, a terceira abordagem utiliza uma estrutura híbrida: modifica e reorganiza o problema primal para simplificar sua estrutura e assim utilizar a formulação dual para resolver o primal modificado.

Apresentaremos a seguir uma técnica, sugerida por R. J. Duffin [21], chamada de condensação, que tem, como princípio básico, transformar o problema signomial na iteração corrente, em um problema posinomial aproximado, "condensado", através da desigualdade a seguir:

Lema 3.1 (Desigualdade entre as médias aritmética e geométrica):

Consideremos v_i , escalares positivos e α_i pesos positivos tais que

$$\sum_{i=1}^n \alpha_i = 1$$

Então

$$\sum_{i=1}^n \alpha_i v_i \geq \prod_{i=1}^n (v_i)^{\alpha_i}$$

ou, fazendo $u_i = \alpha_i v_i$,

$$\sum_{i=1}^n u_i \geq \prod_{i=1}^n \left(\frac{u_i}{\alpha_i} \right)^{\alpha_i}. \quad (3.44)$$

As variáveis u_i são escalares positivos e a desigualdade vale como igualdade se e somente se

$$u_i = \alpha_i \sum_{j=1}^n u_j, \quad \text{para } i = 1, 2, \dots, n. \quad (3.45)$$

Assumamos que $g_0^*(t) > 0$. Introduzindo uma nova variável t_0 e adicionando a restrição $t_0 \geq g_0(t)$, podemos escrever o problema 3.1 da seguinte forma geral:

$$\min t_0$$

s.a.

$$g_k(t) \leq \sigma_k, \quad k = 0, \dots, p$$

$$t_k > 0$$

Vamos mostrar que qualquer problema de programação geométrica pode ser escrito simplesmente em termos de funções posinomiais. Consideremos a k -ésima restrição $g_k(t)$:

$$g_k(t) = P_k(t) - Q_k(t),$$

onde

$$P_k(t) = \sum_{i \in J[k]^+} \sigma_{ki} u_i(t) \quad = \text{coleção dos termos positivos}$$

$$Q_k(t) = \sum_{i \in J[k]^-} |\sigma_{ki}| u_i(t) \quad = \text{coleção dos módulos dos termos negativos}$$

onde

$$u_i(t) = c_{ki} \prod_{j=1}^m t_j^{a_{kij}}$$

$$J[k]^+ = \{i; \sigma_{ki} = +1, i \in J[k]\}$$

$$J[k]^- = \{i; \sigma_{ki} = -1, i \in J[k]\}$$

CASO 1: $\sigma_k = 1$

Com essa consideração, temos a seguinte restrição:

$$P_k(t) - Q_k(t) \leq 1$$

que é equivalente a

$$\frac{P_k(t)}{Q_k^+(t)} \leq 1$$

onde $Q_k^+(t) = 1 + Q_k(t)$.

CASO 2: $\sigma_k = -1$

Temos que:

$$P_k(t) - Q_k(t) \leq -1$$

que é equivalente a

$$\frac{P_k^+(t)}{Q_k(t)} \leq 1$$

onde $P_k^+(t) = 1 + P_k(t)$.

Observemos que P_k , P_k^+ , Q_k , Q_k^+ são funções posinomiais. Assim, o raciocínio é semelhante para ambos os casos. Sem perda de generalidade, consideremos o segundo caso.

Identifiquemos os termos $Q_i(t)$ no problema anterior com os u_i termos da desigualdade 3.44 e 3.45. Temos obviamente que $Q_i(t) > 0$. Consideremos $Q_k(t)$ dado por:

$$Q_k(t) = \sum_i Q_{ki}(t) = \sum_i u_i(t) \tag{3.46}$$

onde $Q_{ki}(t) = |\sigma_{ki}|u_i(t) = u_i(t)$.

No ponto $t = \bar{t}$, definamos os pesos do i -ésimo termo no denominador da k -ésima restrição por

$$\delta_{ki} = \frac{Q_{ki}(\bar{t})}{\sum_i Q_{ki}(\bar{t})} = \frac{Q_{ki}(\bar{t})}{Q_k(\bar{t})} \quad (3.47)$$

Por uma aplicação direta da desigualdade 3.44, temos o seguinte posinômio condensado:

$$Q_k(t) \geq Q_k(t, \bar{t}) = \prod_i \left[\frac{Q_{ki}(t)}{\delta_{ki}} \right]^{\delta_{ki}} \quad (3.48)$$

ou

$$Q_k(t, \bar{t}) = \prod_i \left[\frac{Q_{ki}(t) Q_i(\bar{t})}{Q_{ki}(\bar{t})} \right]^{\frac{Q_{ki}(\bar{t})}{Q_k(\bar{t})}} \quad (3.49)$$

Daí

$$Q_k(t) \geq Q_k(t, \bar{t}) = \prod_i \left[\frac{c_{ki} \prod_j t_j^{a_{kij}} Q_{ki}(\bar{t})}{c_{ki} \prod_j \bar{t}_j^{a_{kij}}} \right]^{\frac{Q_{ki}(\bar{t})}{Q_k(\bar{t})}} \quad (3.50)$$

que se reduz a

$$Q_k(t) \geq Q_k(t, \bar{t}) = Q_k(\bar{t}) \prod_j \left(\frac{t_j}{\bar{t}_j} \right)^{\sum_j a_{kij} Q_{ki}(\bar{t})} \quad (3.51)$$

Observemos que na equação acima, $Q_k(\bar{t})$, \bar{t}_j e $\sum_j a_{kij} Q_{ki}(\bar{t})$ são constantes conhecidas, uma vez especificado o vetor $\bar{t} = (\bar{t}_1, \bar{t}_2, \dots, \bar{t}_m)$.

Supondo que \bar{t} é escolhido viável, temos

$$\frac{P_k(\bar{t})}{Q_k(\bar{t})} \leq 1, \quad \forall k \quad (3.52)$$

Então, qualquer solução viável para

$$\frac{P_k(t)}{Q_k(t, \bar{t})} \leq 1 \quad (3.53)$$

também será viável para

$$\frac{P_k(t)}{Q_k(t)} \leq 1 \quad (3.54)$$

No geral, não vale a volta dessa afirmação. É por isso que devemos escolher \bar{t} viável para o problema original. Em outras palavras, estamos constatando que o espaço solução para as restrições originais está totalmente contido no espaço solução para as restrições condensadas.

Resumindo, suponhamos o seguinte problema condensado:

$$\begin{aligned} & \min t_0 \\ & \text{s. a.} \\ & \frac{P_i(t)}{Q_i(t, \bar{t})} \leq 1 \quad i = 1, \dots, p \end{aligned} \quad (3.55)$$

Segundo [7], esse problema possui as seguintes propriedades teóricas e computacionais:

1. É um problema posinomial, dado que $P_i(t)$ e $Q_i(t, \bar{t})$ são posinômios;
2. Qualquer vetor t satisfazendo as restrições de 3.55 também satisfará às restrições do problema original, pois

$$\begin{aligned}
P_i(t) &< Q_i(t) \\
Q_i(t, \bar{t}) &\leq Q_i(t)
\end{aligned}
\tag{3.56}$$

Dessa forma, a região viável resultante das restrições posinomiais condensadas é um subconjunto da região viável obtida pelas restrições signomiais originais. Assim, qualquer solução ótima t^* do problema condensado será viável (não necessariamente ótima) para o problema original.

Observação: Pela definição 3.1, podemos entender a Programação Geométrica Signomial como uma extensão da Programação Geométrica Posinomial. Os problemas posinomiais são equivalentes a problemas de otimização convexa com restrições lineares e são resolvidos tanto em sua forma primal, como em sua forma dual. Em relação aos problemas de programação geométrica signomial, os algoritmos primais mais utilizados são aqueles que resolvem as sucessivas condensações. Assim, a busca da solução desses problemas se processa iterativamente através de uma seqüência de problemas posinomiais, cujas soluções convergem para pelo menos um ponto estacionário do problema signomial (ver [7]). Esse ponto estacionário pode ser entendido como o primeiro ponto que transforma a desigualdade 3.44 em igualdade a menos de uma tolerância, isto é, o ponto será dito estacionário se, para um ε dado, tivermos

$$\sum_{i=1}^n u_i - \prod_{i=1}^n \left(\frac{u_i}{\alpha_i} \right)^{\alpha_i} \leq \varepsilon$$

A diferença fundamental entre as diversas versões dos algoritmos se encontra na forma das aproximações posinomiais para os termos signomiais, veja [2], [22], [36], [67], [68], [71].

3.3 Referências bibliográficas

Nesta seção, apresentaremos uma coletânea dos trabalhos nas áreas afins à metodologia abordada nesse capítulo. Nosso propósito, com essa revisão bibliográfica, é mostrar que não temos conhecimento de outro trabalho que resolva um problema 0-1 sob a formulação de programação geométrica signomial. Não estamos, porém, considerando a extensa produção em Programação Semidefinida e suas aplicações à Otimização Contínua.

KARMARKAR [38] descreve um algoritmo de pontos interiores aproximado para resolver problemas de viabilidade $\{0, 1\}$, sustentando que vários problemas de combinatória podem ser resolvidos eficientemente através de algoritmos baseados na aproximação proposta. Os resultados obtidos provêm da aplicação do algoritmo em problemas de satisfiabilidade [37], [86] e de *set covering* [39].

Segundo WARNERS *et al.* [93], a função potencial utilizada em [38] não se mostra adequada para problemas combinatórios de grande porte. Uma das razões são as matrizes completamente densas, encontradas na resolução dos sistemas lineares envolvidos. Apesar de poder-se utilizar as propriedades para matrizes esparsas, é mais vantajoso usar uma função potencial que produza diretamente sistemas lineares esparsos. Tal função potencial é justamente a modificação proposta no trabalho em questão. Segundo os autores, a performance obtida pela função potencial modificada se mostra pelo menos tão boa quanto a da original.

MITCHELL *et al.* [55] apresentam uma revisão dos métodos de pontos interiores para a programação combinatória, discutindo, em detalhe, sua relação com técnicas para a programação linear e *network*, métodos *branch-and-bound* e *branch-and-cut*, minimização de funções potenciais não convexas, técnicas para obtenção de cotas

inferiores e relaxação em programação semidefinida.

YANG e BRICKER [97] investigam o uso de algoritmos *path-following* em problemas de programação geométrica signomial, enfatizando a qualidade da solução encontrada, em detrimento do tempo de CPU. Essa qualidade é medida em função do valor inicial da tolerância, que decresce durante o processo otimizador e serve para comparação com o valor do parâmetro da barreira logarítmica, como critério de parada para o algoritmo. Com essa estratégia, o algoritmo se aproxima da solução global, mesmo que lentamente, evitando convergir para ótimos locais nas primeiras iterações.

CHOI e BRICKER [13] trabalham com problemas semi-discretos de programação geométrica em que algumas variáveis devem ser arredondadas para tornar a solução viável. Nesse sentido, apresentam uma técnica para estimar as penalidades sofridas por esses arredondamentos.

NASCIMENTO e ANDRADE [63] resolvem um problema de otimização de um sistema hidrotérmico através de uma seqüência de problemas convexos com restrições lineares. Os autores apresentam o modelo com a formulação de um problema de programação geométrica e propõem um algoritmo de pontos interiores inviáveis para resolver a seqüência de problemas convexos citada, fazendo uso de uma função mérito do tipo barreira/penalidade para eliminar os resíduos primal e dual e encontrar soluções aproximadas próximas à trajetória central.

3.4 Metodologia de Otimização Contínua para o problema 0-1

Nessa seção, apresentamos uma proposta para a resolução de um problema misto via pontos interiores. A diferença principal com relação aos métodos de pon-

tos interiores para programação combinatória, apresentados na seção anterior, é o tratamento de problemas 0-1 a partir de técnicas de programação geométrica signomial. Não temos conhecimento de outro trabalho unindo essas linhas de estudo. Para tanto, é acrescido, à função objetivo, um termo signomial, cujo coeficiente cresce à medida que a solução se aproxima do ótimo, fazendo com que as iterações se concentrem próximas do 0 ou do 1. Para essa proposta, não temos resultados teóricos, apenas resultados computacionais (ver capítulo 5 e Apêndice B).

Consideremos a função

$$\varpi = x^n e^{-\left(\frac{x}{\beta}\right)^m}, \quad (3.57)$$

motivada pela expressão da distribuição Gama de parâmetros α e β [77], dada por:

$$f(x) = \begin{cases} \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-\frac{x}{\beta}} & \text{se } x > 0 \\ 0 & \text{se } x = 0 \end{cases}$$

onde $\alpha, \beta > 0$, e é o neperiano e

$$\Gamma(\alpha) = \int_0^\infty e^{-y} \cdot y^{\alpha-1} dy$$

Tomemos

$$\beta = \rho_0 = \frac{1+e}{2}$$

Relembremos o clássico limite (ver, por exemplo, [21], pp.101)

$$e^n = \lim_{M \rightarrow \infty} \left(1 + \frac{n}{M}\right)^M$$

que implica em

$$e^{-n} = \lim_{M \rightarrow \infty} \left(1 + \frac{n}{M}\right)^{-M}$$

Então, ϖ pode ser aproximada por

$$x^n \left(1 + \frac{x^m}{\rho_0^m M} \right)^{-M},$$

para $M > 0$ suficientemente grande e $m > 0$.

Assim, podemos definir a função $f : \mathbb{R} \rightarrow \mathbb{R}$, dada por

$$f(x) = f_{n,m,M}(x) = x^n \left(1 + \frac{x^m}{\rho_0^m M} \right)^{-M}$$

Consideremos agora $h : \mathbb{R} \rightarrow \mathbb{R}$, dada por

$$h(x) = \frac{f(x)}{f(1)}$$

Temos naturalmente que $h(1) = 1$.

Vamos encontrar o valor de n para que $h(e) = e$.

Queremos

$$h(e) = \frac{e^n \left(1 + \frac{e^m}{\rho_0^m M} \right)^{-M}}{\left(1 + \frac{1}{\rho_0^m M} \right)^{-M}} = e$$

Assim,

$$\begin{aligned} e^{n-1} &= \left(1 + \frac{e^m}{\rho_0^m M} \right)^M \cdot \left(1 + \frac{1}{\rho_0^m M} \right)^{-M} \\ \implies n &= 1 + M \ln \left(1 + \frac{e^m}{\rho_0^m M} \right) - M \ln \left(1 + \frac{1}{\rho_0^m M} \right) \end{aligned} \tag{3.58}$$

Para esse valor de n , vale:

$$g(1) = g(e) = 1,$$

onde

$$g(x) = \frac{h(x)}{x} \quad (3.59)$$

Proposição 3.1. *Para $m > 0$ e $M > 0$, a função g possui um único ponto crítico, x^* , no intervalo $(1, e)$. Mais precisamente, o ponto crítico é um ponto de máximo local no intervalo $[1, e]$. Assim, minimizar g nesse intervalo significa encontrar $x = 1$ ou $x = e$.*

Demonstração:

Consideremos a função real g dada por 3.59:

$$g(x) = f(1)^{-1} x^{n-1} \cdot \left(1 + \frac{x^m}{\rho_0^m M} \right)^{-M}$$

A fim de simplificar a apresentação, denotemos

$$D_{m,M}(x) = 1 + \frac{x^m}{\rho_0^m M}$$

Assim, temos que:

$$\begin{aligned} f(1)g'(x) &= (n-1)x^{n-2}(D_{m,M}(x))^{-M} + x^{n-1}(-M)(D_{m,M}(x))^{-M-1} \frac{m}{\rho_0^m M} x^{m-1} \\ &= x^{n-2}(D_{m,M}(x))^{-M} \left[(n-1) - \frac{m}{\rho_0^m} x^m (D_{m,M}(x))^{-1} \right] \\ &= x^{n-2}(D_{m,M}(x))^{-M} \left[M \cdot \ln \frac{1 + \frac{e^m}{\rho_0^m M}}{1 + \frac{1}{\rho_0^m M}} - \frac{m}{\rho_0^m} x^m (D_{m,M}(x))^{-1} \right] \end{aligned}$$

em que a última igualdade é conseqüência da expressão de n , dada por 3.58.

O ponto crítico dessa função é tal que

$$f(1)g'(x) = 0 \implies \begin{cases} x^{n-2} = 0 \\ (D_{m,M}(x))^{-M} = 0 \\ M \cdot \ln \frac{1 + \frac{e^m}{\rho_0^m M}}{1 + \frac{1}{\rho_0^m M}} - \frac{m}{\rho_0^m} x^m (D_{m,M}(x))^{-1} = 0 \end{cases}$$

Mas,

$$x^{n-2} = 0 \implies \begin{cases} x = 0, \text{ para } n > 2 \\ x = \pm\infty, \text{ para } n < 2 \\ \text{N\~{a}o h\~{a} solu\~{c}\~{a}o para } n = 2 \end{cases}$$

$$(D_{m,M}(x))^{-M} = 0 \implies x = \pm\infty$$

para $M > 0$.

Para que $M \cdot \ln \frac{1 + \frac{e^m}{\rho_0^m M}}{1 + \frac{1}{\rho_0^m M}} - \frac{m}{\rho_0^m} x^m (D_{m,M}(x))^{-1} = 0$, devemos ter, como $M > 0$:

$$\begin{aligned} \left(1 + \frac{x^m}{\rho_0^m M}\right) \cdot \ln \frac{1 + \frac{e^m}{\rho_0^m M}}{1 + \frac{1}{\rho_0^m M}} - \frac{m}{\rho_0^m M} x^m &= 0 \\ \implies \frac{1}{\rho_0^m M} \left[m - \ln \frac{1 + \frac{e^m}{\rho_0^m M}}{1 + \frac{1}{\rho_0^m M}} \right] x^m &= \ln \frac{1 + \frac{e^m}{\rho_0^m M}}{1 + \frac{1}{\rho_0^m M}} \end{aligned}$$

Definindo

$$B_{m,M} = m - \ln \frac{1 + \frac{e^m}{\rho_0^m M}}{1 + \frac{1}{\rho_0^m M}}$$

devemos ter

$$\frac{B_{m,M}}{\rho_0^m M} x^m = \ln \frac{1 + \frac{e^m}{\rho_0^m M}}{1 + \frac{1}{\rho_0^m M}}$$

Como $m, M, \rho > 0$, se $B_{m,M} < 0$, n\~{a}o existe raiz $x \in (1, e)$ como desejado e se $B_{m,M} = 0$, n\~{a}o existe raiz real, usando, para ambos os casos, que o logaritmo \u00e9 positivo.

Para $B_{m,M} > 0$:

$$x^m = \frac{\rho_0^m M \ln \frac{1 + \frac{e^m}{\rho_0^m M}}{1 + \frac{1}{\rho_0^m M}}}{B_{m,M}}$$

$$\implies \bar{x} = \rho_0 \left[\frac{M \ln \frac{1 + \frac{e^m}{\rho_0^m M}}{1 + \frac{1}{\rho_0^m M}}}{B_{m,M}} \right]^{\frac{1}{m}}$$

Mostremos que $B_{m,M} > 0$:

De fato, supondo

$$B_{m,M} = m - \ln \frac{1 + \frac{e^m}{\rho_0^m M}}{1 + \frac{1}{\rho_0^m M}} > 0,$$

temos

$$m > \ln \frac{1 + \frac{e^m}{\rho_0^m M}}{1 + \frac{1}{\rho_0^m M}}$$

$$\iff e^m > \frac{1 + \frac{e^m}{\rho_0^m M}}{1 + \frac{1}{\rho_0^m M}}$$

$$\iff e^m \left(1 + \frac{1}{\rho_0^m M} \right) > 1 + \frac{e^m}{\rho_0^m M}$$

$$\iff e^m > 1$$

$$\iff m > 0$$

Para que $\bar{x} \in (1, e)$, devemos ter, equivalentemente:

$$\left(\frac{1}{\rho_0} \right)^m < \frac{M \cdot \left[\ln \left(1 + \frac{e^m}{\rho_0^m M} \right) - \ln \left(1 + \frac{1}{\rho_0^m M} \right) \right]}{m - \left[\ln \left(1 + \frac{e^m}{\rho_0^m M} \right) - \ln \left(1 + \frac{1}{\rho_0^m M} \right) \right]} < \left(\frac{e}{\rho_0} \right)^m,$$

ou seja,

$$(i.) \left[M + \left(\frac{1}{\rho_0} \right)^m \right] \left[\ln \left(1 + \frac{e^m}{\rho_0^m M} \right) - \ln \left(1 + \frac{1}{\rho_0^m M} \right) \right] > m \left(\frac{1}{\rho_0} \right)^m$$

e

$$(ii.) \left[M + \left(\frac{e}{\rho_0} \right)^m \right] \left[\ln \left(1 + \frac{e^m}{\rho_0^m M} \right) - \ln \left(1 + \frac{1}{\rho_0^m M} \right) \right] < m \left(\frac{e}{\rho_0} \right)^m$$

Para simplificar a apresentação, fixemos a variável m e denotemos

$$u_1(M) = M + \left(\frac{1}{\rho_0} \right)^m$$

$$u_e(M) = M + \left(\frac{e}{\rho_0} \right)^m$$

$$v(M) = \ln \left(1 + \frac{e^m}{\rho_0^m M} \right) - \ln \left(1 + \frac{1}{\rho_0^m M} \right)$$

$$w_1(M) = u_1(M)v(M)$$

$$w_e(M) = u_e(M)v(M)$$

Temos que

$$u_1(0) = \left(\frac{1}{\rho_0} \right)^m$$

$$u_e(0) = \left(\frac{e}{\rho_0} \right)^m$$

$$v(0) := \lim_{M \rightarrow 0_+} v(M) = \lim_{M \rightarrow 0_+} \left[\ln \left(1 + \frac{e^m}{\rho_0^m M} \right) - \ln \left(1 + \frac{1}{\rho_0^m M} \right) \right] = m$$

essa última obtida pela regra de L'Hôpital. Assim,

$$\lim_{M \rightarrow 0_+} w_1(M) = \frac{m}{\rho_0^m}$$

$$\lim_{M \rightarrow 0_+} w_e(M) = \frac{me^m}{\rho_0^m}$$

isto é, as expressões (i.) e (ii.) valem com a igualdade no limite $M \rightarrow 0$, para m fixo.

Calculando as derivadas, temos

$$u'_1(M) = 1$$

$$u'_e(M) = 1$$

$$v'(M) = \frac{1}{M} \left[\frac{-e^m}{\rho_0^m M + e^m} + \frac{1}{\rho_0^m M + 1} \right] = \frac{\rho_0^m (1 - e^m)}{(\rho_0^m M + e^m)(\rho_0^m M + 1)}$$

e, conseqüentemente,

$$w'_1(M) = \ln \left(1 + \frac{e^m}{\rho_0^m M} \right) - \ln \left(1 + \frac{1}{\rho_0^m M} \right) + \left[M + \left(\frac{1}{\rho_0} \right)^m \right] \frac{\rho_0^m (1 - e^m)}{(\rho_0^m M + e^m)(\rho_0^m M + 1)}$$

$$= \ln \left(1 + \frac{e^m}{\rho_0^m M} \right) - \ln \left(1 + \frac{1}{\rho_0^m M} \right) + \frac{1 - e^m}{\rho_0^m M + e^m}$$

$$w'_e(M) = \ln \left(1 + \frac{e^m}{\rho_0^m M} \right) - \ln \left(1 + \frac{1}{\rho_0^m M} \right) + \left[M + \left(\frac{e}{\rho_0} \right)^m \right] \frac{\rho_0^m (1 - e^m)}{(\rho_0^m M + e^m)(\rho_0^m M + 1)}$$

$$= \ln \left(1 + \frac{e^m}{\rho_0^m M} \right) - \ln \left(1 + \frac{1}{\rho_0^m M} \right) + \frac{1 - e^m}{\rho_0^m M + 1}$$

Da expressão para w'_1 , temos:

$$\begin{aligned} w''_1(M) &= \frac{\rho_0^m (1 - e^m)}{(\rho_0^m M + e^m)(\rho_0^m M + 1)} - \frac{(1 - e^m) \rho_0^m}{(\rho_0^m M + e^m)^2} \\ &= \frac{(\rho_0^m M + e^m) \rho_0^m (1 - e^m) - (\rho_0^m M + 1) \rho_0^m (1 - e^m)}{(\rho_0^m M + e^m)^2 (\rho_0^m M + 1)} \\ &= \frac{-\rho_0^m}{\rho_0^m M + 1} \cdot \frac{(e^m - 1)^2}{(\rho_0^m M + e^m)^2} < 0 \end{aligned}$$

Além disso, temos,

I)

$$\begin{aligned}
\lim_{M \rightarrow \infty} w_1(M) &= \lim_{M \rightarrow \infty} \left[M + \frac{1}{\rho_0^m} \right] \ln \frac{1 + \frac{e^m}{\rho_0^m M}}{1 + \frac{1}{\rho_0^m M}} \\
&= \lim_{M \rightarrow \infty} \frac{\frac{\rho_0^m(1-e^m)}{(\rho_0^m M + e^m)(\rho_0^m M + 1)}}{-\frac{1}{(M + \frac{1}{\rho_0^m})^2}} \\
&= \lim_{M \rightarrow \infty} -\frac{(1-e^m)(\rho_0^m M + 1)}{\rho_0^m(\rho_0^m M + e^m)} \\
&= \frac{e^m - 1}{\rho_0^m} \\
&> \frac{m}{\rho_0^m} = \lim_{M \rightarrow 0_+} w_1(M)
\end{aligned}$$

onde a terceira igualdade é obtida pela regra de L'Hôpital;

II)

$$\lim_{M \rightarrow 0_+} w'_1(M) = m + \frac{1 - e^m}{e^m} = m - 1 + \frac{1}{e^m} > 0$$

para $m > 0$ e

$$\lim_{M \rightarrow \infty} w'_1(M) = 0$$

Conclui-se que a desigualdade (i.) é verificada para todo para (m, M) positivo.

Analogamente, a partir da expressão para w'_e , temos:

$$w''_e(M) = \frac{(\rho_0^m M + 1)\rho_0^m(1 - e^m) - (\rho_0^m M + e^m)\rho_0^m(1 - e^m)}{(\rho_0^m M + e^m)(\rho_0^m M + 1)^2} = \frac{\rho_0^m}{\rho_0^m M + e^m} \cdot \frac{(e^m - 1)^2}{(\rho_0^m M + 1)^2}$$

que possui sinal positivo, pois $\frac{\rho_0^m}{\rho_0^m M + e^m} > 0$.

Mais ainda, como

I)

$$\begin{aligned}
 \lim_{M \rightarrow \infty} w_e(M) &= \lim_{M \rightarrow \infty} \left[M + \frac{e^m}{\rho_0^m} \right] \ln \frac{1 + \frac{e^m}{\rho_0^m M}}{1 + \frac{1}{\rho_0^m} M} \\
 &= \lim_{M \rightarrow \infty} \frac{\frac{\rho_0^m (1 - e^m)}{(\rho_0^m M + e^m)(\rho_0^m M + 1)}}{\frac{1}{(M + \frac{1}{\rho_0^m})^2}} \\
 &= \frac{e^m - 1}{\rho_0^m} \\
 &< \frac{me^m}{\rho_0^m} = \lim_{M \rightarrow 0_+} w_e(M)
 \end{aligned}$$

II)

$$\lim_{M \rightarrow 0_+} w'_e(M) = m + \left(\frac{e}{\rho_0}\right)^m \left(\frac{\rho_0}{e}\right)^m (1 - e^m) = m + 1 - e^m < 0$$

para $m > 0$ e

$$\lim_{M \rightarrow \infty} w'_e(M) = 0$$

a desigualdade (ii.) é verificada para todo para (m, M) positivo.

Portanto, \bar{x} é o ponto de máximo do intervalo $(1, e)$.

Os gráficos a seguir ilustram a conclusão sobre (i.) e (ii.), onde $\rho_0 = \frac{1+e}{2}$. Para facilitar o entendimento, passamos os termos do lado direito das expressões para o lado esquerdo.

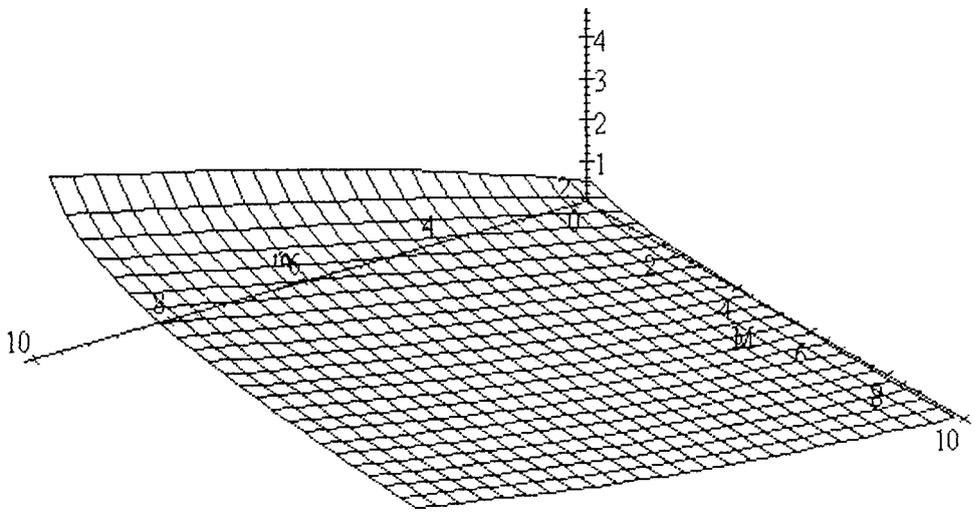


Gráfico 3.1 – Ilustração da positividade da expressão (i.)

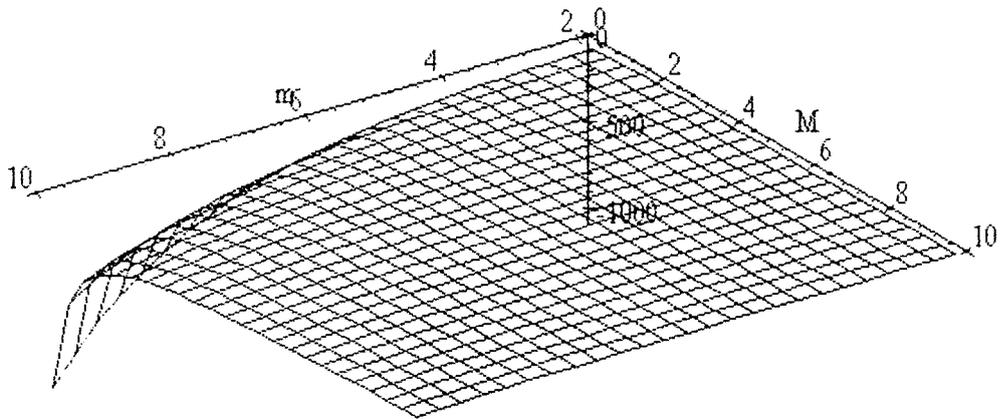


Gráfico 3.2 – Ilustração da negatividade da expressão (ii.)

Estendendo para o \mathbb{R}^N , consideremos $g : \mathbb{R}^N \rightarrow \mathbb{R}^N$, dada por

$$g(x) = (g_1(x_1), g_2(x_2), \dots, g_N(x_N))$$

onde $g_i : \mathbb{R} \rightarrow \mathbb{R}$. Podemos formular um problema de programação geométrica da seguinte forma:

$$\min \prod_{k=1}^N g_k(x_k)$$

s.a.

$$x_k \in [1, e], \quad \forall k$$

em que a solução ótima será um ponto extremo do intervalo $[1, e]$, para cada $k \in 1, \dots, N$.

Explicitando o valor de g^2 , temos

$$\min \prod_{k=1}^N x_k^{n-1} \cdot \left(1 + \frac{x_k^m}{\rho_0^m M} \right)^{-M}$$

s.a.

(3.60)

$$x_k \in [1, e], \quad \forall k$$

Em princípio, esse não é um problema de programação geométrica devido ao termo entre parênteses. No entanto, podemos transformá-lo em um problema de programação geométrica signomial da seguinte forma. A vantagem de tratar problemas 0-1 sob a formulação de programação geométrica é a fácil adaptação da função 3.57 como um termo signomial.

²Tanto na apresentação da metodologia, quanto no desenvolvimento dos cálculos, desconsideramos o denominador $f(1)$.

$$\min \prod_{k=1}^N x_k^{n-1} y_k^{-M}$$

s. a.

(3.61)

$$x_k \in [1, e] \quad \forall k$$

$$-\frac{x_k^m}{\rho_0^m M} + y_k \leq 1$$

pois o problema, com a substituição de $\left(1 + \frac{x_k^m}{\rho_0^m M}\right)$ por y_k :

$$\min \prod_{k=1}^N x_k^{n-1} y_k^{-M}$$

s. a.

(3.62)

$$x_k \in [1, e], \quad \forall k$$

$$y_k = 1 + \frac{x_k^m}{\rho_0^m M}$$

é equivalente a 3.61. De fato, mostremos que o multiplicador da restrição em questão é positivo:

$$\begin{aligned} L(x, y, \mu, \theta, \lambda) = & \prod_{k=1}^N x_k^{n-1} y_k^{-M} + \sum_{k=1}^N \mu_k (x_k - e) + \sum_{k=1}^N \theta_k (1 - x_k) \\ & + \sum_{k=1}^N \lambda_k \left(-\frac{x_k^m}{\rho_0^m M} + y_k - 1 \right) \end{aligned}$$

é a função de Lagrange associada a 3.61. As condições de ortogonalidade explicam, em particular, que

$$-M x_j^{n-1} y_j^{-M-1} \prod_{k \neq j} x_k^{n-1} y_k^{-M} + \lambda_j = 0$$

ou

$$\lambda_j = M x_j^{n-1} y_j^{-M-1} \prod_{k \neq j} x_k^{n-1} y_k^{-M}.$$

Claramente, como $x_k \geq 1$, $\forall k$, a hipótese de absurdo de que $\lambda_j = 0$ levaria a que alguma componente de y fosse ilimitada. Entretanto, temos, do segundo conjunto de restrições,

$$y_k = 1 + \frac{x_k^m}{\rho_0^m M} \leq 1 + \frac{e_k^m}{\rho_0^m M},$$

e, portanto, $\lambda_j > 0$, para todo j .

Usando a técnica da condensação, vamos encontrar as restrições posinomiais condensadas para cada restrição signomial de 3.61:

$$-\frac{x_k^m}{\rho_0^m M} + y_k \leq 1 \quad \longrightarrow \quad \frac{y_k}{1 + \frac{x_k^m}{\rho_0^m M}} \leq 1$$

Substituindo o denominador conforme técnica de condensação apresentada (ver equações 3.54 e 3.53), temos que se valer

$$y_k \left(\frac{1}{\left(\frac{1}{\bar{\delta}}\right)} \right)^{\bar{\delta}} \cdot \left(\frac{\frac{x_k^m}{\rho_0^m M}}{\frac{1}{\delta_k}} \right)^{\delta_k} \leq 1 \quad (3.63)$$

que pode ser escrita como

$$(\bar{\delta})^{\bar{\delta}} (\delta_k)^{\delta_k} \left(\frac{x_k^m}{\rho_0^m M} \right)^{-\delta_k} y_k \leq 1$$

também valerá

$$\frac{y_k}{1 + \frac{x_k^m}{\rho_0^m M}} \leq 1$$

Assim, podemos escrever o seguinte problema condensado:

$$\min \prod_{k=1}^N x_k^{n-1} \cdot y_k^{-M}$$

s. a.

$$x_k^{-1} \leq 1 \quad \forall k \quad (3.64)$$

$$\frac{1}{e} x_k \leq 1 \quad \forall k$$

$$(\bar{\delta})^{\bar{\delta}} (\delta_k)^{\delta_k} \left(\frac{x_k^m}{\rho_0^m M} \right)^{-\delta_k} y_k \leq 1 \quad \forall k$$

3.5 Metodologia aplicada ao Problema de Telecomunicação Sem Fio

3.5.1 Técnica de Condensação

Consideremos o modelo 1.3:

$$\min \sum_{i=1}^N c_i \bar{z}_i + C. \sum_{i=1}^N \sum_{j=1}^M \bar{p}_{ij}$$

s.a.

$$\sum_{i=1}^N a_{ij} \bar{x}_{ij} \geq 1$$

$$\sum_{j=1}^M a_{ij} \bar{x}_{ij} \leq \Lambda_i \bar{z}_i$$

(3.65)

$$W h_j \bar{p}_{ij} - \left(\sum_{k \neq j} a_{ik} h_k \bar{p}_{ik} + \sum_{s \neq i} \sum_{t=1}^M a_{sj} h_t \bar{p}_{st} \right) \gamma_j \bar{r}_{ij} - \eta_0 W \gamma_j \bar{r}_{ij} \geq 0$$

$$0 \leq \bar{p}_{ij} \leq P_j \bar{x}_{ij}$$

$$R_j \bar{x}_{ij} \leq \bar{r}_{ij}$$

$$\bar{x}_{ij}, \bar{z}_i \in \{0, 1\}.$$

Definamos $\forall i, \forall j$,

$$z_i = e^{\bar{z}_i},$$

$$x_{ij} = e^{\bar{x}_{ij}},$$

$$p_{ij} = e^{\bar{p}_{ij}},$$

$$r_{ij} = e^{\bar{r}_{ij}}.$$

Relaxando as variáveis inteiras no intervalo $[0; 1]$ e usando propriedades elementares do logaritmo, o problema 3.65 pode ser escrito como

$$\min \ln \left(\prod_i z_i^{c_i} \cdot \prod_i \prod_j p_{ij}^C \right)$$

s.a.

$$\begin{aligned} \ln z_i^{-\Lambda_i} \prod_j x_{ij}^{a_{ij}} &\leq \ln 1 && \forall i \\ \ln \prod_i x_{ij}^{-a_{ij}} &\leq \ln e^{-1} && \forall j \\ \ln z_i^{-1} &\leq \ln 1 && \forall i \\ \ln z_i &\leq \ln e && \forall i \\ \ln x_{ij}^{-1} &\leq \ln 1 && \forall i, \forall j \\ \ln x_{ij} &\leq \ln e && \forall i, \forall j \\ \ln p_{ij} \cdot x_{ij}^{-P_j} &\leq \ln 1 && \forall i, \forall j \\ \ln p_{ij}^{-1} &\leq \ln 1 && \forall i, \forall j \\ \ln x_{ij}^{R_j} \cdot r_{ij}^{-1} &\leq \ln 1 && \forall i, \forall j \end{aligned} \tag{3.66}$$

$$\ln r_{ij}^{\gamma_j} \cdot \left[\ln \left(e^{\eta_0 W} \cdot \prod_{k \neq j} p_{ik}^{a_{ik} h_k} \cdot \prod_{s \neq i} \prod_{t=1}^M p_{st}^{a_{sj} h_t} \right) \right] + \ln p_{ij}^{-W h_j} \leq 0 \quad \forall i, \forall j$$

Como o logaritmo é uma função crescente, o problema pode ser reescrito como:

$$\min \prod_i z_i^{c_i} \cdot \prod_i \prod_j p_{ij}^C$$

S.a.

$$\begin{aligned} z_i^{-\Lambda_i} \prod_j x_{ij}^{a_{ij}} &\leq 1 && \forall i \\ e. \prod_i x_{ij}^{-a_{ij}} &\leq 1 && \forall j \\ z_i^{-1} &\leq 1 && \forall i \\ \frac{1}{e} z_i &\leq 1 && \forall i \\ x_{ij}^{-1} &\leq 1 && \forall i, \forall j \\ \frac{1}{e} x_{ij} &\leq 1 && \forall i, \forall j \\ p_{ij} \cdot x_{ij}^{-P_j} &\leq 1 && \forall i, \forall j \\ p_{ij}^{-1} &\leq 1 && \forall i, \forall j \\ x_{ij}^{R_j} \cdot r_{ij}^{-1} &\leq 1 && \forall i, \forall j \end{aligned} \tag{3.67}$$

$$\ln r_{ij}^{\gamma_j} \cdot \left[\ln \left(e^{\eta_0 W} \cdot \prod_{k \neq j} p_{ik}^{a_{ik} h_k} \cdot \prod_{s \neq i} \prod_{t=1}^M p_{st}^{a_{sj} h_t} \right) \right] + \ln p_{ij}^{-W h_j} \leq 0 \quad \forall i, \forall j$$

Observemos que apenas mantivemos o último conjunto de restrições sob a forma logarítmica, como em 3.66. Como suas restrições são não-lineares, escrevê-lo na forma de um problema de programação geométrica (ver (PPG) no apêndice A) exige mais trabalho. Necessitamos da aproximação

$$\ln x = \varepsilon^{-1} (x^\varepsilon - 1)$$

para ε suficientemente pequeno³. Fazendo

$$\theta_{ij} = e^{\eta_0 W} \cdot \prod_{k \neq j} p_{ik}^{a_{ik} h_k} \cdot \prod_{s \neq i} \prod_{t=1}^M p_{st}^{a_{st} h_t}$$

podemos reescrever o conjunto de restrições em questão como

$$\begin{aligned} & \varepsilon^{-2} \left(r_{ij}^{\varepsilon \gamma_j} - 1 \right) \left(\theta_{ij}^\varepsilon - 1 \right) + \varepsilon^{-1} \left(p_{ij}^{-\varepsilon W h_j} - 1 \right) \leq 0 \\ \implies & \varepsilon^{-2} r_{ij}^{\varepsilon \gamma_j} \theta_{ij}^\varepsilon - \varepsilon^{-2} r_{ij}^{\varepsilon \gamma_j} - \varepsilon^{-2} \theta_{ij}^\varepsilon + \varepsilon^{-2} + \varepsilon^{-1} p_{ij}^{-\varepsilon W h_j} - \varepsilon^{-1} \leq 0 \\ \implies & \varepsilon^{-1} r_{ij}^{\varepsilon \gamma_j} \theta_{ij}^\varepsilon + \varepsilon^{-1} + p_{ij}^{-\varepsilon W h_j} \leq \varepsilon^{-1} r_{ij}^{\varepsilon \gamma_j} + \varepsilon^{-1} \theta_{ij}^\varepsilon + 1 \\ \implies & \frac{\varepsilon^{-1} r_{ij}^{\varepsilon \gamma_j} \theta_{ij}^\varepsilon + \varepsilon^{-1} + p_{ij}^{-\varepsilon W h_j}}{\varepsilon^{-1} r_{ij}^{\varepsilon \gamma_j} + \varepsilon^{-1} \theta_{ij}^\varepsilon + 1} \leq 1 \end{aligned} \quad (3.68)$$

Pela desigualdade entre as médias aritmética e geométrica 3.44, temos que

$$\varepsilon^{-1} r_{ij}^{\varepsilon \gamma_j} + \varepsilon^{-1} \theta_{ij}^\varepsilon + 1 \geq \left(\frac{1}{\overline{\alpha_{ij}}} \varepsilon^{-1} r_{ij}^{\varepsilon \gamma_j} \right)^{\overline{\alpha_{ij}}} \left(\frac{1}{\overline{\beta_{ij}}} \varepsilon^{-1} \theta_{ij}^\varepsilon \right)^{\overline{\beta_{ij}}} \left(\frac{1}{\overline{\delta_{ij}}} \right)^{\overline{\delta_{ij}}} \quad (3.69)$$

para $\overline{\alpha_{ij}}, \overline{\beta_{ij}}, \overline{\delta_{ij}}$ não negativos tais que

$$\overline{\alpha_{ij}} + \overline{\beta_{ij}} + \overline{\delta_{ij}} = 1 \quad \forall i, \forall j.$$

Mais ainda, pela condição 3.45, se

$$\overline{\alpha_{ij}} = \frac{\varepsilon^{-1} r_{ij}^{\varepsilon \gamma_j}}{\varepsilon^{-1} r_{ij}^{\varepsilon \gamma_j} + \varepsilon^{-1} \theta_{ij}^\varepsilon + 1}, \quad (3.70)$$

$$\overline{\beta_{ij}} = \frac{\varepsilon^{-1} \theta_{ij}^\varepsilon}{\varepsilon^{-1} r_{ij}^{\varepsilon \gamma_j} + \varepsilon^{-1} \theta_{ij}^\varepsilon + 1}, \quad (3.71)$$

³A demonstração dessa aproximação é imediata com o uso da regra de L'Hôpital.

$$\overline{\delta_{ij}} = \frac{1}{\varepsilon^{-1}r_{ij}^{\varepsilon\gamma_j} + \varepsilon^{-1}\theta_{ij}^{\varepsilon} + 1}, \quad (3.72)$$

a inequação 3.69 valerá com a igualdade, isto é

$$\varepsilon^{-1}r_{ij}^{\varepsilon\gamma_j} + \varepsilon^{-1}\theta_{ij}^{\varepsilon} + 1 = \left(\frac{1}{\overline{\alpha_{ij}}}\varepsilon^{-1}r_{ij}^{\varepsilon\gamma_j}\right)^{\overline{\alpha_{ij}}} \left(\frac{1}{\overline{\beta_{ij}}}\varepsilon^{-1}\theta_{ij}^{\varepsilon}\right)^{\overline{\beta_{ij}}} \left(\frac{1}{\overline{\delta_{ij}}}\right)^{\overline{\delta_{ij}}}$$

Portanto, para $\overline{\alpha_{ij}}$, $\overline{\beta_{ij}}$ e $\overline{\delta_{ij}}$ satisfazendo 3.70, 3.71 e 3.72, respectivamente, se a desigualdade

$$\frac{\varepsilon^{-1}r_{ij}^{\varepsilon\gamma_j}\theta_{ij}^{\varepsilon} + \varepsilon^{-1} + p_{ij}^{-\varepsilon Wh_j}}{\left(\frac{1}{\overline{\alpha_{ij}}}\varepsilon^{-1}r_{ij}^{\varepsilon\gamma_j}\right)^{\overline{\alpha_{ij}}} \left(\frac{1}{\overline{\beta_{ij}}}\varepsilon^{-1}\theta_{ij}^{\varepsilon}\right)^{\overline{\beta_{ij}}} \left(\frac{1}{\overline{\delta_{ij}}}\right)^{\overline{\delta_{ij}}}} \leq 1 \quad \forall i, \forall j \quad (3.73)$$

for verdadeira, teremos o conjunto de restrições 3.68 satisfeito (ver Técnica de Condensação na seção 3.2).

A inequação 3.73 é equivalente, $\forall i, \forall j$, a

$$(\varepsilon^{-1}r_{ij}^{\varepsilon\gamma_j}\theta_{ij}^{\varepsilon} + \varepsilon^{-1} + p_{ij}^{-\varepsilon Wh_j}).K.r_{ij}^{-\varepsilon\overline{\alpha_{ij}}\gamma_j}.\theta_{ij}^{-\varepsilon\overline{\beta_{ij}}} \leq 1$$

\Leftrightarrow

$$K.\varepsilon^{-1}.\theta_{ij}^{\varepsilon(1-\overline{\beta_{ij}})} .r_{ij}^{\varepsilon(1-\overline{\alpha_{ij}})\gamma_j} + K.\varepsilon^{-1}.\theta_{ij}^{-\varepsilon\overline{\beta_{ij}}} .r_{ij}^{-\varepsilon\overline{\alpha_{ij}}\gamma_j} + K.p_{ij}^{-\varepsilon Wh_j} .\theta_{ij}^{-\varepsilon\overline{\beta_{ij}}} .r_{ij}^{-\varepsilon\overline{\alpha_{ij}}\gamma_j} \leq 1$$

onde

$$K = \varepsilon^{(\overline{\alpha_{ij}}+\overline{\beta_{ij}})} .\overline{\alpha_{ij}}^{\overline{\alpha_{ij}}} .\overline{\beta_{ij}}^{\overline{\beta_{ij}}} .\overline{\delta_{ij}}$$

Portanto, o modelo 3.67 pode ser aproximado por:

$$\min \prod_i z_i^{c_i} \cdot \prod_i \prod_j p_{ij}^C$$

s.a.

$$\begin{aligned} z_i^{-\Lambda_i} \prod_j x_{ij}^{a_{ij}} &\leq 1 && \forall i \\ e \cdot \prod_i x_{ij}^{-a_{ij}} &\leq 1 && \forall j \\ z_i^{-1} &\leq 1 && \forall i \\ \frac{1}{e} z_i &\leq 1 && \forall i \\ x_{ij}^{-1} &\leq 1 && \forall i, \forall j \\ \frac{1}{e} x_{ij} &\leq 1 && \forall i, \forall j \\ p_{ij} \cdot x_{ij}^{-P_j} &\leq 1 && \forall i, \forall j \\ p_{ij}^{-1} &\leq 1 && \forall i, \forall j \\ x_{ij}^{R_j} \cdot r_{ij}^{-1} &\leq 1 && \forall i, \forall j \\ K \cdot \varepsilon^{-1} \cdot \theta_{ij}^{\varepsilon(1-\overline{\beta}_{ij})} \cdot r_{ij}^{\varepsilon(1-\overline{\alpha}_{ij})\gamma_j} + K \cdot \varepsilon^{-1} \cdot \theta_{ij}^{-\varepsilon\overline{\beta}_{ij}} \cdot r_{ij}^{-\varepsilon\overline{\alpha}_{ij}\gamma_j} \\ &+ K \cdot p_{ij}^{-\varepsilon W h_j} \cdot \theta_{ij}^{-\varepsilon\overline{\beta}_{ij}} \cdot r_{ij}^{-\varepsilon\overline{\alpha}_{ij}\gamma_j} &\leq 1 && \forall i, \forall j \end{aligned} \tag{3.74}$$

com $\theta_{ij} = e^{\eta_0 W} \cdot \prod_{k \neq j} p_{ik}^{a_{ik} h_k} \cdot \prod_{s \neq i} \prod_{t=1}^M p_{st}^{a_{sj} h_t}$

3.5.2 Heurística 0-1

Objetivando a integralidade das variáveis z e x , aplicamos a heurística desenvolvida na seção 3.4 (ver 3.61) ao problema 3.74. Obtemos, assim, o seguinte problema de programação geométrica signomial, onde $\alpha \geq 1$:

$$\min \prod_i z_i^{c_i} \cdot \prod_i \prod_j p_{ij}^C + \alpha \cdot f(1)^{-\overline{N}-\overline{NM}} \prod_i z_i^{n-1} \cdot y_i^{-M} \cdot \prod_i \prod_j x_{ij}^{n-1} \cdot y_{ij}^{-M}$$

s.a.

$$z_i^{-\Lambda_i} \prod_j x_{ij}^{a_{ij}} \leq 1 \quad \forall i$$

$$e \cdot \prod_i x_{ij}^{-a_{ij}} \leq 1 \quad \forall j$$

$$z_i^{-1} \leq 1 \quad \forall i$$

$$\frac{1}{e} z_i \leq 1 \quad \forall i$$

$$x_{ij}^{-1} \leq 1 \quad \forall i, \forall j$$

$$\frac{1}{e} x_{ij} \leq 1 \quad \forall i, \forall j$$

$$p_{ij} \cdot x_{ij}^{-P_j} \leq 1 \quad \forall i, \forall j$$

$$p_{ij}^{-1} \leq 1 \quad \forall i, \forall j$$

$$x_{ij}^{R_j} \cdot r_{ij}^{-1} \leq 1 \quad \forall i, \forall j$$

$$K \cdot \varepsilon^{-1} \cdot \theta_{ij}^{\varepsilon(1-\overline{\beta_{ij}})} \cdot r_{ij}^{\varepsilon(1-\overline{\alpha_{ij}})\gamma_j} + K \cdot \varepsilon^{-1} \cdot \theta_{ij}^{-\varepsilon\overline{\beta_{ij}}} \cdot r_{ij}^{-\varepsilon\overline{\alpha_{ij}}\gamma_j}$$

$$+ K \cdot p_{ij}^{-\varepsilon W h_j} \cdot \theta_{ij}^{-\varepsilon\overline{\beta_{ij}}} \cdot r_{ij}^{-\varepsilon\overline{\alpha_{ij}}\gamma_j} \leq 1 \quad \forall i, \forall j$$

$$-\frac{z_i^{m_z}}{\rho_0^{m_z} M} + y_i \leq 1 \quad \forall i$$

$$-\frac{x_{ij}^{m_x}}{\rho_0^{m_x} M} + y_{ij} \leq 1 \quad \forall i, \forall j$$

(3.75)

Observação: O parâmetro α , de valor inicial igual a 1, cresce à medida que a solução posinomial se aproxima da solução signomial (ver algoritmo ASIG), de modo a aumentar a ponderação da função inserida e alcançar a tolerância desejada para a solução.

Podemos ainda favorecer uma das variáveis inteiras. Como se trata de uma heurística, em que os valores do parâmetro m são testados computacionalmente, na busca do melhor deles, podemos definir um valor de m para cada variável, dando maior peso àquela que atingiu valores mais afastados de 0-1. Dessa forma, o segundo termo da função objetivo do problema anterior se divide em dois termos, para obtermos o seguinte problema signomial:

$$\min \prod_i z_i^{c_i} \cdot \prod_i \prod_j p_{ij}^C + \alpha \cdot f(1)_z^{-\bar{N}} \prod_i z_i^{n-1} \cdot y_i^{-M} + \alpha \cdot f(1)_x^{-\bar{NM}} \prod_i \prod_j x_{ij}^{n-1} \cdot y_{ij}^{-M}$$

s.a.

$$z_i^{-\Lambda_i} \prod_j x_{ij}^{a_{ij}} \leq 1 \quad \forall i$$

$$e. \prod_i x_{ij}^{-a_{ij}} \leq 1 \quad \forall j$$

$$z_i^{-1} \leq 1 \quad \forall i$$

$$\frac{1}{e} z_i \leq 1 \quad \forall i$$

$$x_{ij}^{-1} \leq 1 \quad \forall i, \forall j$$

$$\frac{1}{e} x_{ij} \leq 1 \quad \forall i, \forall j$$

$$p_{ij} \cdot x_{ij}^{-P_j} \leq 1 \quad \forall i, \forall j$$

$$p_{ij}^{-1} \leq 1 \quad \forall i, \forall j$$

$$x_{ij}^{R_j} \cdot r_{ij}^{-1} \leq 1 \quad \forall i, \forall j$$

$$K \cdot \varepsilon^{-1} \cdot \theta_{ij}^{\varepsilon(1-\bar{\beta}_{ij})} \cdot r_{ij}^{\varepsilon(1-\bar{\alpha}_{ij})\gamma_j} + K \cdot \varepsilon^{-1} \cdot \theta_{ij}^{-\varepsilon\bar{\beta}_{ij}} \cdot r_{ij}^{-\varepsilon\bar{\alpha}_{ij}\gamma_j}$$

$$+ K \cdot p_{ij}^{-\varepsilon W h_j} \cdot \theta_{ij}^{-\varepsilon\bar{\beta}_{ij}} \cdot r_{ij}^{-\varepsilon\bar{\alpha}_{ij}\gamma_j} \leq 1 \quad \forall i, \forall j$$

$$\begin{aligned}
-\frac{z_i^{m_z}}{\rho_0^{m_z M}} + y_i &\leq 1 \quad \forall i \\
-\frac{x_{ij}^{m_x}}{\rho_0^{m_x M}} + y_{ij} &\leq 1 \quad \forall i, \forall j
\end{aligned} \tag{3.76}$$

onde m_z e m_x são os diferentes valores para o parâmetro m .

Nesse ponto, precisamos condensar novamente o problema, desenvolvendo as duas últimas desigualdades. De 3.64, temos, para cada iteração signomial, o seguinte problema primal de programação geométrica posinomial:

$$\min \prod_i z_i^{c_i} \cdot \prod_i \prod_j p_{ij}^C + \alpha \cdot f(1)_z^{-\bar{N}} \prod_i z_i^{n-1} \cdot y_i^{-M} + \alpha \cdot f(1)_x^{-\bar{NM}} \prod_i \prod_j x_{ij}^{n-1} \cdot y_{ij}^{-M}$$

s.a.

$$\begin{aligned}
z_i^{-\Lambda_i} \prod_j x_{ij}^{a_{ij}} &\leq 1 && \forall i \\
e \cdot \prod_i x_{ij}^{-a_{ij}} &\leq 1 && \forall j \\
z_i^{-1} &\leq 1 && \forall i \\
\frac{1}{e} z_i &\leq 1 && \forall i \\
x_{ij}^{-1} &\leq 1 && \forall i, \forall j \\
\frac{1}{e} x_{ij} &\leq 1 && \forall i, \forall j \\
p_{ij} \cdot x_{ij}^{-P_j} &\leq 1 && \forall i, \forall j \\
p_{ij}^{-1} &\leq 1 && \forall i, \forall j \\
x_{ij}^{R_j} \cdot r_{ij}^{-1} &\leq 1 && \forall i, \forall j
\end{aligned}$$

$$\begin{aligned}
& K.\varepsilon^{-1}.\theta_{ij}^{\varepsilon(1-\overline{\beta}_{ij})}.r_{ij}^{\varepsilon(1-\overline{\alpha}_{ij})\gamma_j} + K.\varepsilon^{-1}.\theta_{ij}^{-\varepsilon\overline{\beta}_{ij}}.r_{ij}^{-\varepsilon\overline{\alpha}_{ij}\gamma_j} \\
& + K.p_{ij}^{-\varepsilon Wh_j}.\theta_{ij}^{-\varepsilon\overline{\beta}_{ij}}.r_{ij}^{-\varepsilon\overline{\alpha}_{ij}\gamma_j} \leq 1 \quad \forall i, \forall j \\
& (\overline{\delta})^{\overline{\delta}}(\delta_i)^{\delta_i} \left(\frac{z_i^{m_z}}{\rho_0^{m_z} M} \right)^{-\delta_i} y_i \leq 1 \quad \forall i \\
& (\overline{\delta})^{\overline{\delta}}(\delta_{ij})^{\delta_{ij}} \left(\frac{x_{ij}^{m_x}}{\rho_0^{m_x} M} \right)^{-\delta_{ij}} y_{ij} \leq 1 \quad \forall i, \forall j
\end{aligned} \tag{3.77}$$

onde m_z e m_x são como em 3.76 e $\overline{\alpha}_{ij}$, $\overline{\beta}_{ij}$, $\overline{\delta}_{ij}$ (embutido em K), δ_i e δ_{ij} são os coeficientes da condensação.

Consideremos

$$g_k(t) = \sum_{h \in J[k]} (\text{coef})_h \prod_{l=1}^{3\overline{N}+4\overline{NM}} t_l^{A_{hl}}, \quad k = 0, 1, \dots, 5\overline{N} + \overline{M} + 10\overline{NM} \tag{3.78}$$

onde $(\text{coef}) \in \mathbb{R}^{3+5\overline{N}+\overline{M}+8\overline{NM}}$ é dado por

$$\left[\begin{array}{c}
1 \\
\alpha f_z(1)^{-\overline{N}} \\
\alpha f_x(1)^{-\overline{N}} \\
(1)_{\overline{N} \times 1} \\
(e)_{\overline{M} \times 1} \\
(1)_{\overline{N} \times 1} \\
\left(\frac{1}{e}\right)_{\overline{N} \times 1} \\
(1)_{\overline{NM} \times 1} \\
\left(\frac{1}{e}\right)_{\overline{NM} \times 1} \\
(1)_{\overline{NM} \times 1} \\
(1)_{\overline{NM} \times 1} \\
(1)_{\overline{NM} \times 1} \\
(K.\varepsilon^{-1})_{\overline{NM} \times 1} \\
(K.\varepsilon^{-1})_{\overline{NM} \times 1} \\
(K)_{\overline{NM} \times 1} \\
((\overline{\delta})^{\overline{\delta}}(\delta_i)^{\delta_i}(\rho_0^{m_z} M)^{\delta_i})_{\overline{N} \times 1} \\
((\overline{\delta})^{\overline{\delta}}(\delta_{ij})^{\delta_{ij}}(\rho_0^{m_x} M)^{\delta_{ij}})_{\overline{NM} \times 1}
\end{array} \right]$$

$$J[0] = \{1, 2, 3\}, \quad J[k] = \{k + 3\}, \quad k = 1, \dots, 5\overline{N} + \overline{M} + 10\overline{NM} \quad (3.79)$$

e a matriz $A_{hl} \in \mathbb{R}^{3+5\overline{N}+\overline{M}+10\overline{NM}} \times \mathbb{R}^{3\overline{N}+4\overline{NM}}$ dos expoentes é dada por

$$\begin{bmatrix} (c_i)_{1 \times \overline{N}} & (0)_{1 \times \overline{NM}} & (C)_{1 \times \overline{NM}} & (0)_{1 \times \overline{N} + \overline{NM}} & (0)_{1 \times \overline{N} + \overline{NM}} \\ (n-1)_{1 \times \overline{N}} & (0)_{1 \times \overline{NM}} & (0)_{1 \times \overline{NM}} & (0)_{1 \times \overline{N} + \overline{NM}} & (-(M) \ (0))_{1 \times \overline{N} + \overline{NM}} \\ (0)_{1 \times \overline{N}} & (n-1)_{1 \times \overline{NM}} & (0)_{1 \times \overline{NM}} & (0)_{1 \times \overline{N} + \overline{NM}} & ((0) \ - (M))_{1 \times \overline{N} + \overline{NM}} \\ (\Delta_i)_{\overline{N} \times \overline{N}} & (A_1)_{\overline{N} \times \overline{NM}} & (0)_{\overline{N} \times \overline{NM}} & (0)_{\overline{N} \times \overline{NM}} & (0)_{\overline{N} \times \overline{N} + \overline{NM}} \\ (0)_{\overline{M} \times \overline{N}} & -(A_2)_{\overline{M} \times \overline{NM}} & (0)_{\overline{M} \times \overline{NM}} & (0)_{\overline{M} \times \overline{NM}} & (0)_{\overline{M} \times \overline{N} + \overline{NM}} \\ -(I)_{\overline{N} \times \overline{N}} & (0)_{\overline{N} \times \overline{NM}} & (0)_{\overline{N} \times \overline{NM}} & (0)_{\overline{N} \times \overline{NM}} & (0)_{\overline{N} \times \overline{N} + \overline{NM}} \\ (I)_{\overline{N} \times \overline{N}} & (0)_{\overline{N} \times \overline{NM}} & (0)_{\overline{N} \times \overline{NM}} & (0)_{\overline{N} \times \overline{NM}} & (0)_{\overline{N} \times \overline{N} + \overline{NM}} \\ (0)_{\overline{NM} \times \overline{N}} & -(I)_{\overline{NM} \times \overline{NM}} & (0)_{\overline{NM} \times \overline{NM}} & (0)_{\overline{NM} \times \overline{NM}} & (0)_{\overline{NM} \times \overline{N} + \overline{NM}} \\ (0)_{\overline{NM} \times \overline{N}} & (I)_{\overline{NM} \times \overline{NM}} & (0)_{\overline{NM} \times \overline{NM}} & (0)_{\overline{NM} \times \overline{NM}} & (0)_{\overline{NM} \times \overline{N} + \overline{NM}} \\ (0)_{\overline{NM} \times \overline{N}} & -(P_j)_{\overline{NM} \times \overline{NM}} & (I)_{\overline{NM} \times \overline{NM}} & (0)_{\overline{NM} \times \overline{NM}} & (0)_{\overline{NM} \times \overline{N} + \overline{NM}} \\ (0)_{\overline{NM} \times \overline{N}} & (0)_{\overline{NM} \times \overline{NM}} & -(I)_{\overline{NM} \times \overline{NM}} & (0)_{\overline{NM} \times \overline{NM}} & (0)_{\overline{NM} \times \overline{N} + \overline{NM}} \\ (0)_{\overline{NM} \times \overline{N}} & (R_j)_{\overline{NM} \times \overline{NM}} & -(I)_{\overline{NM} \times \overline{NM}} & (0)_{\overline{NM} \times \overline{NM}} & (0)_{\overline{NM} \times \overline{N} + \overline{NM}} \\ (0)_{\overline{NM} \times \overline{N}} & (0)_{\overline{NM} \times \overline{NM}} & (E_p^1)_{\overline{NM} \times \overline{NM}} & (E_r^1)_{\overline{NM} \times \overline{NM}} & (0)_{\overline{NM} \times \overline{N} + \overline{NM}} \\ (0)_{\overline{NM} \times \overline{N}} & (0)_{\overline{NM} \times \overline{NM}} & (E_p^2)_{\overline{NM} \times \overline{NM}} & (E_r^2)_{\overline{NM} \times \overline{NM}} & (0)_{\overline{NM} \times \overline{N} + \overline{NM}} \\ (0)_{\overline{NM} \times \overline{N}} & (0)_{\overline{NM} \times \overline{NM}} & (E_p^3)_{\overline{NM} \times \overline{NM}} & (E_r^3)_{\overline{NM} \times \overline{NM}} & (0)_{\overline{NM} \times \overline{N} + \overline{NM}} \\ -(\delta_i)_{\overline{N} \times \overline{N}} & (0)_{\overline{N} \times \overline{NM}} & (0)_{\overline{N} \times \overline{NM}} & (0)_{\overline{N} \times \overline{NM}} & ((I) \ (0))_{\overline{N} \times \overline{N} + \overline{NM}} \\ (0)_{\overline{NM} \times \overline{N}} & -(\delta_{ij})_{\overline{NM} \times \overline{NM}} & (0)_{\overline{NM} \times \overline{NM}} & (0)_{\overline{NM} \times \overline{NM}} & ((0) \ (I))_{\overline{N} \times \overline{N} + \overline{NM}} \end{bmatrix}$$

onde

(0) é a matriz nula nas dimensões indicadas;

(I) é a matriz identidade nas dimensões indicadas;

(c_i) é o vetor dos custos de instalação das ERBs;

$(C)_{1 \times \overline{NM}}$ é o vetor do custo da potência de transmissão, constante para todas as quadrículas;

$(n-1)$ é o vetor de coordenadas constantes e iguais a $n-1$ nas dimensões indicadas;

(M) é o vetor de coordenadas constantes e iguais a M nas dimensões indicadas;

(Δ_i) é a matriz diagonal das capacidades das ERBs;

(A_1) é dada por

$$\begin{bmatrix} (a_{1j}) & & & & \\ & (a_{2j}) & & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & (a_{1j}) \end{bmatrix}$$

(A_2) é dada por

$$\left[\text{diag}(a_{1j}) \quad \text{diag}(a_{2j}) \quad \cdots \quad \text{diag}(a_{1j}) \right]$$

onde $\text{diag}(a_{ij})$ é a matriz diagonal, cujos elementos são os elementos de (a_{ij}) ;

(P_j) e (R_j) são matrizes cujas diagonais são os vetores P_j e R_j , repetidos N vezes;

(E_p^k) e (E_r^k) são as matrizes dos expoentes das variáveis p e r , respectivamente, para cada um dos três termos k , no conjunto de restrições de QoS condensadas, dadas por:

$$E_p^1 = \begin{bmatrix} 0 & \varepsilon(1 - \overline{\beta_{12}}) & \cdots & \varepsilon(1 - \overline{\beta_{1M}}) \\ \varepsilon(1 - \overline{\beta_{21}}) & 0 & \cdots & \varepsilon(1 - \overline{\beta_{2M}}) \\ \vdots & & \ddots & \\ \varepsilon(1 - \overline{\beta_{N1}}) & \varepsilon(1 - \overline{\beta_{N2}}) & \cdots & 0 \end{bmatrix}$$

$$E_r^1 = \begin{bmatrix} 0 & \varepsilon(1 - \overline{\alpha_{12}})\gamma_2 & \cdots & \varepsilon(1 - \overline{\alpha_{1M}})\gamma_M \\ \varepsilon(1 - \overline{\alpha_{21}})\gamma_1 & 0 & \cdots & \varepsilon(1 - \overline{\alpha_{2M}})\gamma_M \\ \vdots & & \ddots & \\ \varepsilon(1 - \overline{\alpha_{N1}})\gamma_1 & \varepsilon(1 - \overline{\alpha_{N2}})\gamma_2 & \cdots & 0 \end{bmatrix}$$

$$E_p^2 = \begin{bmatrix} 0 & -\varepsilon\overline{\beta_{12}} & \cdots & -\varepsilon\overline{\beta_{1M}} \\ -\varepsilon\overline{\beta_{21}} & 0 & \cdots & -\varepsilon\overline{\beta_{2M}} \\ \vdots & & \ddots & \\ -\varepsilon\overline{\beta_{N1}} & -\varepsilon\overline{\beta_{N2}} & \cdots & 0 \end{bmatrix}$$

$$E_r^2 = \begin{bmatrix} 0 & -\varepsilon\overline{\alpha_{12}}\gamma_2 & \cdots & -\varepsilon\overline{\alpha_{1M}}\gamma_M \\ -\varepsilon\overline{\alpha_{21}}\gamma_1 & 0 & \cdots & -\varepsilon\overline{\alpha_{2M}}\gamma_M \\ \vdots & & \ddots & \\ -\varepsilon\overline{\alpha_{NM}}\gamma_M & -\varepsilon\overline{\alpha_{NM}}\gamma_M & \cdots & 0 \end{bmatrix}$$

$$E_p^3 = \begin{bmatrix} -\varepsilon W h_1 & -\varepsilon\overline{\beta_{12}} & \cdots & -\varepsilon\overline{\beta_{1M}} \\ -\varepsilon\overline{\beta_{21}} & -\varepsilon W h_2 & \cdots & -\varepsilon\overline{\beta_{2M}} \\ \vdots & & \ddots & \\ -\varepsilon\overline{\beta_{N1}} & -\varepsilon\overline{\beta_{N2}} & \cdots & -\varepsilon W h_M \end{bmatrix}$$

e $E_r^3 = E_r^2$.

Assim, podemos escrever 3.77 na forma de (PPG) (ver apêndice):

$$\min g_0(t)$$

s.a.

$$g_k(t) \leq 1, \quad k = 1, \dots, 5\overline{N} + \overline{M} + 8\overline{NM} \quad (3.80)$$

$$t \in \mathbb{R}_{++}^{3\overline{N}+4\overline{NM}}$$

O problema dual de programação geométrica associado é dado por:

$$\max u(x)$$

s.a.

$$\sum_{h \in J[0]} x_h = 1 \tag{3.81}$$

$$\sum_{h=1}^{3+5\bar{N}+\bar{M}+8\bar{N}\bar{M}} A_{lh} x_h = 0$$

$$x \in \mathbb{R}_{+}^{3+5\bar{N}+\bar{M}+8\bar{N}\bar{M}}$$

onde

$$u(x) = \prod_{h=1}^{3+5\bar{N}+\bar{M}+8\bar{N}\bar{M}} \left(\frac{(\text{coef})_h}{x_h} \right)^{x_h} \prod_{k=1}^{5\bar{N}+\bar{M}+8\bar{N}\bar{M}} \lambda_k^{\lambda_k}$$

$$\lambda_k = \sum_{h \in J[k]} x_h$$

$J[k]$ é dado por 3.79

3.5.3 Existência de Solução

Nesta parte do capítulo, provaremos a existência de uma solução ótima para o problema 3.80. Para tanto, faremos as seguintes hipóteses, considerando os valores da tabela 6.1 (essas hipóteses foram levadas em conta na simulação):

- O número de ERBs é no mínimo igual a 2 e no máximo igual a 8, isto é, $2 \leq \bar{N} \leq 8$;
- O número de quadrículas é no mínimo igual a 2 e no máximo igual a 24, isto é, $2 \leq \bar{M} \leq 24$;
- Cada ERB deve cobrir, com sinal mínimo para a comunicação, pelo menos 60% e no máximo 70% das quadrículas, isto é, $60\%M \leq \sum_j a_{ij} \leq 70\%\bar{M}$;

- Cada ERB deve ter capacidade de cobrir pelo menos 90% das quadrículas em que seu sinal chega com o valor mínimo desejado para comunicação, isto é, $\Delta_i \geq 90\% \sum_i a_{ij}$;
- Cada quadrícula deve receber sinal mínimo para comunicação de pelo menos 70% do total das ERBs, isto é, $\sum_i a_{ij} \geq 70\% \overline{N}$;
- $P_{min} = \min_j \{P_j\}$, $P_{max} = \max_j \{P_j\}$,
 $R_{min} = \min_j \{R_j\}$, $R_{max} = \max_j \{R_j\}$,
 $h_{min} = \min_j \{h_j\}$, $h_{max} = \max_j \{h_j\}$,
 $\gamma_{min} = \min_j \{\gamma_j\}$, $\gamma_{max} = \max_j \{\gamma_j\}$.

Proposição 3.2. *O problema 3.80 é superconsistente (ver apêndice A).*

Prova: De acordo com a definição A.1, devemos mostrar um vetor das variáveis que satisfaça à condição de Slater. Tomemos o ponto

$$t = (z_i^0, x_{ij}^0, p_{ij}^0, r_{ij}^0, y_i^0, y_{ij}^0) \in \mathbb{R}^{2\overline{N}+4\overline{NM}} \quad (3.82)$$

onde

$$\begin{aligned} z_i^0 &= \frac{10}{9} e^{\frac{4}{5}} & \forall i \\ x_{ij}^0 &= e^{\frac{4}{5}} & \forall i, \forall j \\ p_{ij}^0 &= e^{\frac{P_{min}}{NM-1}} & \forall i, \forall j \\ r_{ij}^0 &= e^{R_{max}} & \forall i, \forall j \\ y_i^0 &= y_{ij}^0 = 1 & \forall i, \forall j \end{aligned}$$

Vamos verificar que esse ponto é interior viável para todos os conjuntos de restrições de 3.77:

1. Capacidade das ERBs:

$$\begin{aligned}
 z_i^{-\Lambda_i} \prod_j x_{ij}^{a_{ij}} &= \left(\frac{10}{9}e^{\frac{4}{5}}\right)^{-\Lambda_i} \prod_j \left(e^{\frac{4}{5}}\right)^{a_{ij}} \\
 &\leq \left(\frac{10}{9}e^{\frac{4}{5}}\right)^{-\frac{9}{10}\frac{6}{10}M} \left(e^{\frac{4}{5}}\right)^{\frac{7}{10}M} \\
 &= \left(\frac{10}{9}\right)^{-\frac{27}{50}M} \left(e^{\frac{4}{5}}\right)^{\left(\frac{7}{10}-\frac{27}{50}\right)M} \\
 &= \left(\frac{9}{10}\right)^{\frac{50}{27}M} \left(e^{\frac{4}{5}}\right)^{\frac{4}{25}M} \\
 &= \left[\left(\frac{9}{10}\right)^{\frac{50}{27}} e^{\frac{16}{125}}\right]^M \\
 &< (1)^M \\
 &= 1
 \end{aligned}$$

2. Cobertura das quadrículas:

$$\begin{aligned}
 e. \prod_i x_{ij}^{a_{ij}} &= e. \prod_i \left(e^{\frac{4}{5}}\right)^{-a_{ij}} \\
 &\leq e. \left(e^{\frac{4}{5}}\right)^{-\frac{7}{10}N} \\
 &\leq e. \left(e^{\frac{4}{5}}\right)^{-\frac{7}{5}} \\
 &= e.e^{-\frac{28}{25}} \\
 &= e^{-\frac{3}{25}} \\
 &< 1
 \end{aligned}$$

3. $z_i^0, x_{ij}^0 \in (1, e)$ são facilmente verificados;

4. Potência de transmissão máxima:

$$\begin{aligned}
 p_{ij} \cdot x_{ij}^{-P_j} &= e^{\frac{P_{min}}{NM-1}} \cdot \left(e^{\frac{4}{5}}\right)^{-P_j} \\
 &\leq e^{\frac{P_{min}}{NM-1}} \cdot e^{-\frac{4}{5}P_{min}} \\
 &= e^{\frac{P_{min}}{NM-1} - \frac{4}{5}P_{min}} \\
 &= e^{\frac{P_{min}}{4-1} - \frac{4}{5}P_{min}} \\
 &= e^{-\frac{7}{15}P_{min}} \\
 &< 1
 \end{aligned}$$

5. Como $P_{min} > 0$, obtemos $(p_{ij}^0)^{-1} < 1$;

6. Taxa de transmissão mínima:

$$\begin{aligned}
 x_{ij}^{R_j} \cdot r_{ij}^{-1} &= \left(e^{\frac{4}{5}}\right)^{R_j} \cdot \left(e^{R_{max}}\right)^{-1} \\
 &\leq \left(e^{\frac{4}{5}}\right)^{R_{max}} \cdot e^{-R_{max}} \\
 &= e^{-\frac{1}{5}R_{max}} \\
 &< 1
 \end{aligned}$$

7. Qualidade de serviço:

Como os coeficientes da desigualdade entre as médias aritmética e geométrica são tomados de modo a garantir a igualdade, podemos verificar que t é interior

viável, para ε suficientemente pequeno, usando a restrição original:

$$\begin{aligned}
& \frac{Wh_j p_{ij}}{\sum_{k \neq j} a_{ik} h_k p_{ik} + \sum_{s \neq i} s \sum_{t=1}^{\bar{M}} a_{sj} h_t p_{it} + \eta_0 W} - \gamma_j r_{ij} \\
&= \frac{Wh_j \frac{P_{min}}{NM-1}}{\sum_{k \neq j} a_{ik} h_k \frac{P_{min}}{NM-1} + \sum_{s \neq i} \sum_{t=1}^{\bar{M}} a_{sj} h_t \frac{P_{min}}{NM-1} + \eta_0 W} - \gamma_j R_{max} \\
&\geq \frac{Wh_j}{(NM-1)h_{max} + \eta_0 W} - \gamma_j R_{max} \\
&\geq \frac{Wh_{min}}{(NM-1)h_{max} + \eta_0 W} - \gamma_j R_{max} \\
&= \frac{1250.0,23}{(192-1)0,26+10^{-6}.1250} - 0,7.8, 1 \\
&> 0
\end{aligned}$$

8. Como $z_i^0, m_z, \rho_0, M > 0$ e $y_i^0 = 1$, temos imediatamente que

$$-\frac{(z_i^0)^{m_z}}{\rho_0^{m_z} M} + y_i^0 \leq 1$$

O mesmo vale para o conjunto de restrições envolvendo x .

Portanto, por definição, o problema primal 3.80 é superconsistente.

Observação: Para a existência de uma solução ótima (Teorema A3), necessitamos da consistência e a canonicidade do problema. A proposição anterior prova que o problema é superconsistente e, portanto, consistente. Como todas as colunas da matriz A possuem termos positivos e negativos, basta-nos provar que o sistema linear

$$yA = 0, y \geq 1,$$

possui uma solução y viável (ver observação seguinte à definição A.2) para termos sua canonicidade (esta também pode ser garantida pela boa formulação do problema, ver definição A.2 [70]).

3.6 Metodologia aplicada ao Problema de Otimização de Carteira de Investimentos

3.6.1 Técnica de Condensação

Consideremos o modelo 2.12:

$$\max \left[\rho_1 \sum_{i=1}^N R_i \bar{x}_i - \rho_2 \left(\sum_{i=1}^N \bar{x}_i^2 DU_i^2 + \sum_{i=1}^N \sum_{j \neq i} \bar{x}_i DU_i \bar{x}_j DU_j COR_U(i, j) \right) \right]$$

s.a:

$$\sum_{i=1}^N \bar{x}_i \bar{y}_i \leq C_0$$

$$b_i \bar{y}_i \leq \bar{x}_i \leq C_0 \bar{y}_i \quad \forall i$$

$$\bar{x}_i \in \mathbb{R}_+; \bar{y}_i \in \{0, 1\} \quad \forall i$$

onde R_i , DU_i e $COR_U(i, j)$ são dados previamente calculados, obtidos de amostras de retornos passados dos ativos.

Considerando a variável \bar{y}_i no intervalo $[0, 1]$, podemos relaxar o problema como:

$$\min \left[\rho_2 \left(\sum_{i=1}^N \bar{x}_i^2 DU_i^2 + \sum_{i=1}^N \sum_{j \neq i} \bar{x}_i DU_i \bar{x}_j DU_j COR_U(i, j) \right) - \rho_1 \sum_{i=1}^N R_i \bar{x}_i \right]$$

s.a.:

$$\sum_{i=1}^N \bar{x}_i \bar{y}_i - C_0 \leq 0 \tag{3.83}$$

$$-\bar{x}_i + b_i \bar{y}_i \leq 0 \quad \forall i$$

$$\bar{x}_i - C_0 \bar{y}_i \leq 0 \quad \forall i$$

$$-\bar{x}_i \leq 0 \quad \forall i$$

$$\bar{y}_i \leq 1 \quad \forall i$$

$$-\bar{y}_i \leq 0 \quad \forall i$$

Definindo, $\forall i$,

$$x_i = e^{\bar{x}_i},$$

$$y_i = e^{\bar{y}_i},$$

e usando propriedades elementares do logaritmo, o problema 3.83 pode ser escrito como

$$\min \left[\rho_2 \left(\sum_{i=1}^N DU_i^2 (\ln x_i)^2 + \sum_{i=1}^N \sum_{j \neq i}^N DU_i DU_j COR_U(i, j) (\ln x_i) (\ln x_j) \right) - \rho_1 \sum_{i=1}^N R_i \ln x_i \right]$$

s.a:

$$\sum_{i=1}^N (\ln x_i) (\ln y_i) - C_0 \leq 0$$

$$x_i^{-1} y_i^{b_i} \leq 1 \quad \forall i$$

$$x_i y_i^{C_0} \leq 1 \quad \forall i$$

$$x_i^{-1} \leq 1 \quad \forall i$$

$$\frac{1}{e} y_i \leq 1 \quad \forall i$$

$$y_i^{-1} \leq 1 \quad \forall i$$

\Updownarrow

$$\min t_0$$

s.a:

$$\left[\rho_2 \left(\sum_{i=1}^N DU_i^2 (\ln x_i)^2 + \sum_{i=1}^N \sum_{j \neq i}^N DU_i DU_j COR_U(i, j) (\ln x_i) (\ln x_j) \right) \right. \quad (3.84)$$

$$\left. - \rho_1 \sum_{i=1}^N R_i \ln x_i \right] - \ln t_0 \leq 0$$

$$\sum_{i=1}^N (\ln x_i) (\ln y_i) - C_0 \leq 0 \quad (3.85)$$

$$\begin{aligned}
x_i^{-1}y_i^{b_i} &\leq 1 \quad \forall i \\
x_i y_i^{C_0} &\leq 1 \quad \forall i \\
x_i^{-1} &\leq 1 \quad \forall i \\
\frac{1}{\varepsilon}y_i &\leq 1 \quad \forall i \\
y_i^{-1} &\leq 1 \quad \forall i \\
t_0^{-1} &\leq 1 \quad \forall i
\end{aligned} \tag{3.86}$$

Neste ponto, resta-nos transformar 3.84 e 3.85 em restrições posinomiais. Para tanto, usaremos a técnica de condensação, apresentada na seção 3.2:

Para o conjunto de restrições 3.85, usando a aproximação

$$\ln x = \varepsilon^{-1}(x^\varepsilon - 1)$$

temos:

$$\begin{aligned}
&\sum_{i=1}^N \varepsilon^{-2} (x_i^\varepsilon y_i^\varepsilon - x_i^\varepsilon - y_i^\varepsilon + 1) \leq C_0 \\
\implies &\sum_{i=1}^N \varepsilon^{-2} x_i^\varepsilon y_i^\varepsilon - \sum_{i=1}^N \varepsilon^{-2} x_i^\varepsilon - \sum_{i=1}^N \varepsilon^{-2} y_i^\varepsilon + \varepsilon^{-2} N \leq C_0 \\
\implies &\sum_{i=1}^N \varepsilon^{-2} x_i^\varepsilon y_i^\varepsilon + \varepsilon^{-2} N \leq \sum_{i=1}^N \varepsilon^{-2} x_i^\varepsilon + \sum_{i=1}^N \varepsilon^{-2} y_i^\varepsilon + C_0 \\
\implies &\frac{\sum_{i=1}^N \varepsilon^{-2} x_i^\varepsilon y_i^\varepsilon + \varepsilon^{-2} N}{\sum_{i=1}^N \varepsilon^{-2} x_i^\varepsilon + \sum_{i=1}^N \varepsilon^{-2} y_i^\varepsilon + C_0} \leq 1
\end{aligned} \tag{3.87}$$

Pela desigualdade entre as médias aritmética e geométrica (eq18), temos que

$$\sum_{i=1}^N \varepsilon^{-2} x_i^\varepsilon + \sum_{i=1}^N \varepsilon^{-2} y_i^\varepsilon + C_0 \geq \prod_{i=1}^N \left(\frac{1}{\alpha_i^1} \varepsilon^{-2} x_i^\varepsilon \right)^{\alpha_i^1} \left(\frac{1}{\alpha_i^2} \varepsilon^{-2} y_i^\varepsilon \right)^{\alpha_i^2} \left(\frac{1}{\alpha_i^3} C_0 \right)^{\alpha_i^3} \quad (3.88)$$

para α_i^1, α_i^2 e α_i^3 não negativos tais que

$$\sum_{i=1}^N (\alpha_i^1 + \alpha_i^2) + \alpha_i^3 = 1.$$

Mais ainda, pela condição 3.45, se

$$\alpha_i^1 = \frac{\varepsilon^{-2} x_i^\varepsilon}{\sum_{i=1}^N \varepsilon^{-2} x_i^\varepsilon + \sum_{i=1}^N \varepsilon^{-2} y_i^\varepsilon + C_0}, \quad (3.89)$$

$$\alpha_i^2 = \frac{\varepsilon^{-2} y_i^\varepsilon}{\sum_{i=1}^N \varepsilon^{-2} x_i^\varepsilon + \sum_{i=1}^N \varepsilon^{-2} y_i^\varepsilon + C_0} \quad (3.90)$$

e

$$\alpha_i^3 = \frac{C_0}{\sum_{i=1}^N \varepsilon^{-2} x_i^\varepsilon + \sum_{i=1}^N \varepsilon^{-2} y_i^\varepsilon + C_0} \quad (3.91)$$

a inequação 3.88 valerá com a igualdade, isto é

$$\sum_{i=1}^N \varepsilon^{-2} x_i^\varepsilon + \sum_{i=1}^N \varepsilon^{-2} y_i^\varepsilon + C_0 = \prod_{i=1}^N \left(\frac{1}{\alpha_i^1} \varepsilon^{-2} x_i^\varepsilon \right)^{\alpha_i^1} \left(\frac{1}{\alpha_i^2} \varepsilon^{-2} y_i^\varepsilon \right)^{\alpha_i^2} \left(\frac{1}{\alpha_i^3} C_0 \right)^{\alpha_i^3} \quad (3.92)$$

Portanto, para α_i^1 , α_i^2 e α^3 satisfazendo 3.89, 3.90 e 3.91, respectivamente, se a seguinte desigualdade for verdadeira

$$\frac{\sum_{i=1}^N \varepsilon^{-2} x_i^\varepsilon y_i^\varepsilon + \varepsilon^{-2} N}{\prod_{i=1}^N \left(\frac{1}{\alpha_i^1} \varepsilon^{-2} x_i^\varepsilon \right)^{\alpha_i^1} \left(\frac{1}{\alpha_i^2} \varepsilon^{-2} y_i^\varepsilon \right)^{\alpha_i^2} \left(\frac{1}{\alpha_i^1} C_0 \right)^{\alpha^3}} \leq 1, \quad (3.93)$$

teremos a restrição 3.87 satisfeita.

A inequação 3.93 é equivalente a

$$\left(\sum_{i=1}^N \varepsilon^{-2} x_i^\varepsilon y_i^\varepsilon + \varepsilon^{-2} N \right) \cdot \prod_{i=1}^N \left(\frac{1}{\alpha_i^1} \varepsilon^{-2} x_i^\varepsilon \right)^{-\alpha_i^1} \left(\frac{1}{\alpha_i^2} \varepsilon^{-2} y_i^\varepsilon \right)^{-\alpha_i^2} \left(\frac{1}{\alpha_i^1} C_0 \right)^{-\alpha^3} \leq 1 \quad (3.94)$$

\Leftrightarrow

$$\begin{aligned} & \sum_{j=1}^N \left[\varepsilon^{-2} x_j^\varepsilon y_j^\varepsilon \prod_{i=1}^N \left(\frac{1}{\alpha_i^1} \varepsilon^{-2} x_i^\varepsilon \right)^{-\alpha_i^1} \left(\frac{1}{\beta_i} \varepsilon^{-2} y_i^\varepsilon \right)^{-\alpha_i^2} \left(\frac{1}{\delta} C_0 \right)^{-\alpha^3} \right] \\ & + \sum_{j=1}^N \left[\varepsilon^{-2} N \cdot \prod_{i=1}^N \left(\frac{1}{\alpha_i^1} \varepsilon^{-2} x_i^\varepsilon \right)^{-\alpha_i^1} \left(\frac{1}{\beta_i} \varepsilon^{-2} y_i^\varepsilon \right)^{-\alpha_i^2} \left(\frac{1}{\delta} C_0 \right)^{-\alpha^3} \right] \leq 1 \end{aligned} \quad (3.95)$$

Para a restrição 3.84, temos:

$$\rho_2 \left(\sum_{i=1}^N \varepsilon^{-2} (x_i^\varepsilon - 1) DU_i^2 + \sum_{i=1}^N \sum_{j \neq i} M_{ij} \varepsilon^{-2} (x_i^\varepsilon - 1) (x_j^\varepsilon - 1) \right) \\ - \rho_1 \sum_{i=1}^N R_i \varepsilon^{-2} (x_i^\varepsilon - 1) - \varepsilon^{-1} (t_0^\varepsilon - 1) \leq 0$$

⇕

$$\rho_2 \left[\sum_{i=1}^N \varepsilon^{-2} DU_i^2 x_i^\varepsilon + \sum_{i=1}^N \sum_{j \neq i} M_{ij} \varepsilon^{-2} (x_i^\varepsilon x_j^\varepsilon + 1) \right] + \rho_1 \sum_{i=1}^N R_i \varepsilon^{-2} + \varepsilon^{-1} \\ \leq \rho_2 \left[\sum_{i=1}^N \varepsilon^{-2} DU_i^2 + \sum_{i=1}^N \sum_{j \neq i} M_{ij} \varepsilon^{-2} (x_i^\varepsilon + x_j^\varepsilon) \right] + \rho_1 \sum_{i=1}^N R_i \varepsilon^{-2} x_i^\varepsilon + \varepsilon^{-1} t_0^\varepsilon$$

⇕

$$\frac{\rho_2 \left[\sum_{i=1}^N \varepsilon^{-2} DU_i^2 x_i^\varepsilon + \sum_{i=1}^N \sum_{j \neq i} M_{ij} \varepsilon^{-2} (x_i^\varepsilon x_j^\varepsilon + 1) \right] + \rho_1 \sum_{i=1}^N R_i \varepsilon^{-2} + \varepsilon^{-1}}{\rho_2 \left[\sum_{i=1}^N \varepsilon^{-2} DU_i^2 + \sum_{i=1}^N \sum_{j \neq i} M_{ij} \varepsilon^{-2} (x_i^\varepsilon + x_j^\varepsilon) \right] + \rho_1 \sum_{i=1}^N R_i \varepsilon^{-2} x_i^\varepsilon + \varepsilon^{-1} t_0^\varepsilon} \leq 1 \quad (3.96)$$

onde $M_{ij} = DU_i DU_j COR_U(i, j)$. Pela desigualdade entre as médias aritmética e geométrica 3.44, temos que

$$\rho_2 \left[\sum_{i=1}^N \varepsilon^{-2} DU_i^2 + \sum_{i=1}^N \sum_{j \neq i} M_{ij} \varepsilon^{-2} (x_i^\varepsilon + x_j^\varepsilon) \right] + \rho_1 \sum_{i=1}^N R_i \varepsilon^{-2} x_i^\varepsilon + \varepsilon^{-1} t_0^\varepsilon \geq$$

$$\left[\prod_{i=1}^N \left(\frac{1}{\beta_i^1} \rho_2 \varepsilon^{-2} DU_i^2 \right)^{\beta_i^1} \left(\frac{1}{\beta_i^2} \rho_1 (1 + r_i) \varepsilon^{-2} x_i^\varepsilon \right)^{\beta_i^2} \right] \left[\prod_{i=1}^N \prod_{j \neq i} \left(\frac{1}{\beta_i^3} \rho_2 M_{ij} \varepsilon^{-2} x_i^\varepsilon \right)^{\beta_i^3} \left(\frac{1}{\beta_i^4} \rho_2 M_{ij} \varepsilon^{-2} x_j^\varepsilon \right)^{\beta_i^4} \right] \left(\frac{1}{\beta^\varepsilon} \varepsilon^{-1} t_0^\varepsilon \right)^{\beta^\varepsilon} \quad (3.97)$$

para $\beta_i^1, \beta_i^2, \beta_i^3, \beta_i^4$ e β_i^5 não negativos tais que

$$\sum_{i=1}^N (\beta_i^1 + \beta_i^2 + \beta_i^3 + \beta_i^4 + \beta_i^5) = 1.$$

Mais ainda, pela condição 3.45, se

$$\beta_i^1 = \frac{\rho_2 \varepsilon^{-2} DU_i^2}{\rho_2 \left[\sum_{i=1}^N \varepsilon^{-2} DU_i^2 + \sum_{i=1}^N \sum_{j \neq i} M_{ij} \varepsilon^{-2} (x_i^\varepsilon + x_j^\varepsilon) \right] + \rho_1 \sum_{i=1}^N R_i \varepsilon^{-2} x_i^\varepsilon + \varepsilon^{-1} t_0^\varepsilon}, \quad (3.98)$$

$$\beta_i^2 = \frac{\rho_1 (1 + r_i) \varepsilon^{-2} x_i^\varepsilon}{\rho_2 \left[\sum_{i=1}^N \varepsilon^{-2} DU_i^2 + \sum_{i=1}^N \sum_{j \neq i} M_{ij} \varepsilon^{-2} (x_i^\varepsilon + x_j^\varepsilon) \right] + \rho_1 \sum_{i=1}^N R_i \varepsilon^{-2} x_i^\varepsilon + \varepsilon^{-1} t_0^\varepsilon}, \quad (3.99)$$

$$\beta_i^3 = \frac{\rho_2 M_{ij} \varepsilon^{-2} x_i^\varepsilon}{\rho_2 \left[\sum_{i=1}^N \varepsilon^{-2} DU_i^2 + \sum_{i=1}^N \sum_{j \neq i} M_{ij} \varepsilon^{-2} (x_i^\varepsilon + x_j^\varepsilon) \right] + \rho_1 \sum_{i=1}^N R_i \varepsilon^{-2} x_i^\varepsilon + \varepsilon^{-1} t_0^\varepsilon}, \quad (3.100)$$

$$\beta_i^4 = \frac{\rho_2 M_{ij} \varepsilon^{-2} x_j^\varepsilon}{\rho_2 \left[\sum_{i=1}^N \varepsilon^{-2} DU_i^2 + \sum_{i=1}^N \sum_{j \neq i} M_{ij} \varepsilon^{-2} (x_i^\varepsilon + x_j^\varepsilon) \right] + \rho_1 \sum_{i=1}^N R_i \varepsilon^{-2} x_i^\varepsilon + \varepsilon^{-1} t_0^\varepsilon}, \quad (3.101)$$

$$\beta_i^5 = \frac{\varepsilon^{-1} t_0^\varepsilon}{\rho_2 \left[\sum_{i=1}^N \varepsilon^{-2} DU_i^2 + \sum_{i=1}^N \sum_{j \neq i} M_{ij} \varepsilon^{-2} (x_i^\varepsilon + x_j^\varepsilon) \right] + \rho_1 \sum_{i=1}^N R_i \varepsilon^{-2} x_i^\varepsilon + \varepsilon^{-1} t_0^\varepsilon} \quad (3.102)$$

a inequação 3.97 valerá com a igualdade, isto é,

$$\rho_2 \left[\sum_{i=1}^N \varepsilon^{-2} DU_i^2 + \sum_{i=1}^N \sum_{j \neq i} M_{ij} \varepsilon^{-2} (x_i^\varepsilon + x_j^\varepsilon) \right] + \rho_1 \sum_{i=1}^N R_i \varepsilon^{-2} x_i^\varepsilon + \varepsilon^{-1} t_0^\varepsilon =$$

$$\left[\prod_{i=1}^N \left(\frac{1}{\beta_i^1} \rho_2 \varepsilon^{-2} DU_i^2 \right)^{\beta_i^1} \left(\frac{1}{\beta_i^2} \rho_1 (1+r_i) \varepsilon^{-2} x_i^\varepsilon \right)^{\beta_i^2} \right] \left[\prod_{i=1}^N \prod_{j \neq i} \left(\frac{1}{\beta_i^3} \rho_2 M_{ij} \varepsilon^{-2} x_i^\varepsilon \right)^{\beta_i^3} \left(\frac{1}{\beta_i^4} \rho_2 M_{ij} \varepsilon^{-2} x_j^\varepsilon \right)^{\beta_i^4} \right] \left(\frac{1}{\beta^5} \varepsilon^{-1} t_0^\varepsilon \right)^{\beta^5}$$

Portanto, para β_i^1 , β_i^2 , β_i^3 , β_i^4 e β_i^5 satisfazendo 3.98, 3.99, 3.100, 3.101 e 3.102, respectivamente, se a desigualdade

$$\frac{\rho_2 \left[\sum_{i=1}^N \varepsilon^{-2} DU_i^2 x_i^\varepsilon + \sum_{i=1}^N \sum_{j \neq i} M_{ij} \varepsilon^{-2} (x_i^\varepsilon x_j^\varepsilon + 1) \right] + \rho_1 \sum_{i=1}^N R_i \varepsilon^{-2} + \varepsilon^{-1}}{\left[\prod_{i=1}^N \left(\frac{1}{\beta_i^1} \rho_2 \varepsilon^{-2} DU_i^2 \right)^{\beta_i^1} \left(\frac{1}{\beta_i^2} \rho_1 R_i \varepsilon^{-2} x_i^\varepsilon \right)^{\beta_i^2} \right] \left[\prod_{i=1}^N \prod_{j \neq i} \left(\frac{1}{\beta_i^3} \rho_2 M_{ij} \varepsilon^{-2} x_i^\varepsilon \right)^{\beta_i^3} \left(\frac{1}{\beta_i^4} \rho_2 M_{ij} \varepsilon^{-2} x_j^\varepsilon \right)^{\beta_i^4} \right] \left(\frac{1}{\beta^5} \varepsilon^{-1} t_0^\varepsilon \right)^{\beta^5}} \leq 1,$$

(3.103)

for verdadeira, teremos a restrição 3.96 satisfeita.

Assim, o modelo 2.12 na forma condensada é dado por:

$$\min t_0$$

s.a.

$$\frac{\rho_2 \left[\sum_{i=1}^N \epsilon^{-2} DU_i^2 x_i^\epsilon + \sum_{i=1}^N \sum_{j \neq i} M_{ij} \epsilon^{-2} (x_i^\epsilon x_j^\epsilon + 1) \right] + \rho_1 \sum_{i=1}^N R_i \epsilon^{-2} + \epsilon^{-1}}{\left[\prod_{i=1}^N \left(\frac{1}{\beta_i^1} \rho_2 \epsilon^{-2} DU_i^2 \right)^{\beta_i^1} \left(\frac{1}{\beta_i^2} \rho_1 R_i \epsilon^{-2} x_i^\epsilon \right)^{\beta_i^2} \right] \left[\prod_{i=1}^N \prod_{j \neq i} \left(\frac{1}{\beta_i^3} \rho_2 M_{ij} \epsilon^{-2} x_i^\epsilon \right)^{\beta_i^3} \left(\frac{1}{\beta_i^4} \rho_2 M_{ij} \epsilon^{-2} x_j^\epsilon \right)^{\beta_i^4} \right] \left(\frac{1}{\beta^5} \epsilon^{-1} t_0 \right)^{\beta^5}} \leq 1$$

$$\sum_{j=1}^N \left[\epsilon^{-2} x_j^\epsilon y_j^\epsilon \prod_{i=1}^N \left(\frac{1}{\alpha_i^1} \epsilon^{-2} x_i^\epsilon \right)^{-\alpha_i^1} \left(\frac{1}{\beta_i} \epsilon^{-2} y_i^\epsilon \right)^{-\alpha_i^2} \left(\frac{1}{\delta} C_0 \right)^{-\alpha^3} \right] + \sum_{j=1}^N \left[\epsilon^{-2} N \cdot \prod_{i=1}^N \left(\frac{1}{\alpha_i^1} \epsilon^{-2} x_i^\epsilon \right)^{-\alpha_i^1} \left(\frac{1}{\beta_i} \epsilon^{-2} y_i^\epsilon \right)^{-\alpha_i^2} \left(\frac{1}{\delta} C_0 \right)^{-\alpha^3} \right] \leq 1$$

$$x_i^{-1} y_i^{b_i} \leq 1 \quad \forall i$$

$$x_i y_i^{C_0} \leq 1 \quad \forall i$$

$$x_i^{-1} \leq 1 \quad \forall i \tag{3.104}$$

$$\frac{1}{e} y_i \leq 1 \quad \forall i$$

$$y_i^{-1} \leq 1 \quad \forall i$$

3.6.2 Heurística 0-1

Objetivando a integralidade da variável y , aplicamos a heurística desenvolvida na seção 3.4 ao problema 3.104. Obtemos, assim, o seguinte problema de programação geométrica signomial, de acordo com as restrições 3.103 e 3.95:

min t_0

s.a.

$$\begin{aligned}
& \frac{\rho_2 \left[\sum_{i=1}^N \varepsilon^{-2} DU_i^2 x_i^\varepsilon + \sum_{i=1}^N \sum_{j \neq i} M_{ij} \varepsilon^{-2} (x_i^\varepsilon x_j^\varepsilon + 1) \right] + \rho_1 \sum_{i=1}^N R_i \varepsilon^{-2} + \varepsilon^{-1}}{\left[\prod_{i=1}^N \left(\frac{1}{\beta_i^1} \rho_2 \varepsilon^{-2} DU_i^2 \right)^{\beta_i^1} \left(\frac{1}{\beta_i^2} \rho_1 R_i \varepsilon^{-2} x_i^\varepsilon \right)^{\beta_i^2} \right] \left[\prod_{i=1}^N \prod_{j \neq i} \left(\frac{1}{\beta_i^3} \rho_2 M_{ij} \varepsilon^{-2} x_i^\varepsilon \right)^{\beta_i^3} \left(\frac{1}{\beta_i^4} \rho_2 M_{ij} \varepsilon^{-2} x_j^\varepsilon \right)^{\beta_i^4} \right] \left(\frac{1}{\beta_0^5} \varepsilon^{-1} t_0^\varepsilon \right)^{\beta_0^5}} \leq 1 \\
& \sum_{j=1}^N \left[\varepsilon^{-2} x_j^\varepsilon y_j^\varepsilon \prod_{i=1}^N \left(\frac{1}{\alpha_i^1} \varepsilon^{-2} x_i^\varepsilon \right)^{-\alpha_i^1} \left(\frac{1}{\beta_i} \varepsilon^{-2} y_i^\varepsilon \right)^{-\alpha_i^2} \left(\frac{1}{\delta} C_0 \right)^{-\alpha_i^3} \right] + \sum_{j=1}^N \left[\varepsilon^{-2N} \prod_{i=1}^N \left(\frac{1}{\alpha_i^1} \varepsilon^{-2} x_i^\varepsilon \right)^{-\alpha_i^1} \left(\frac{1}{\beta_i} \varepsilon^{-2} y_i^\varepsilon \right)^{-\alpha_i^2} \left(\frac{1}{\delta} C_0 \right)^{-\alpha_i^3} \right] \leq 1 \\
& x_i^{-1} y_i^{b_i} \leq 1 \quad \forall i \\
& x_i y_i^{C_0} \leq 1 \quad \forall i \\
& x_i^{-1} \leq 1 \quad \forall i \\
& \frac{1}{\varepsilon} y_i \leq 1 \quad \forall i \\
& y_i^{-1} \leq 1 \quad \forall i \\
& -\frac{y_i^m}{\rho_0^m M} + \bar{y}_i \leq 1 \quad \forall i
\end{aligned} \tag{3.105}$$

Como para o problema de telecomunicação sem fio, precisamos condensar novamente o problema 3.105. De 3.64, temos, para cada iteração signomial, o seguinte problema primal de programação geométrica posinomial:

min t_0

s.a.

$$\frac{\rho_2 \left[\sum_{i=1}^N \varepsilon^{-2} DU_i^2 x_i^\varepsilon + \sum_{i=1}^N \sum_{j \neq i} M_{ij} \varepsilon^{-2} (x_i^\varepsilon x_j^\varepsilon + 1) \right] + \rho_1 \sum_{i=1}^N R_i \varepsilon^{-2} + \varepsilon^{-1}}{\left[\prod_{i=1}^N \left(\frac{1}{\beta_i^1} \rho_2 \varepsilon^{-2} DU_i^2 \right)^{\beta_i^1} \left(\frac{1}{\beta_i^2} \rho_1 R_i \varepsilon^{-2} x_i^\varepsilon \right)^{\beta_i^2} \right] \left[\prod_{i=1}^N \prod_{j \neq i} \left(\frac{1}{\beta_i^3} \rho_2 M_{ij} \varepsilon^{-2} x_i^\varepsilon \right)^{\beta_i^3} \left(\frac{1}{\beta_i^4} \rho_2 M_{ij} \varepsilon^{-2} x_j^\varepsilon \right)^{\beta_i^4} \right] \left(\frac{1}{\beta_0^5} \varepsilon^{-1} t_0^\varepsilon \right)^{\beta_0^5}} \leq 1$$

$$\sum_{j=1}^N \left[\varepsilon^{-2} x_j^\varepsilon y_j^\varepsilon \prod_{i=1}^N \left(\frac{1}{\alpha_i} \varepsilon^{-2} x_i^\varepsilon \right)^{-\alpha_i^1} \left(\frac{1}{\beta_i} \varepsilon^{-2} y_i^\varepsilon \right)^{-\alpha_i^2} \left(\frac{1}{\delta} C_0 \right)^{-\alpha^3} \right] + \sum_{j=1}^N \left[\varepsilon^{-2N} \prod_{i=1}^N \left(\frac{1}{\alpha_i} \varepsilon^{-2} x_i^\varepsilon \right)^{-\alpha_i^1} \left(\frac{1}{\beta_i} \varepsilon^{-2} y_i^\varepsilon \right)^{-\alpha_i^2} \left(\frac{1}{\delta} C_0 \right)^{-\alpha^3} \right] \leq 1$$

$$x_i^{-1} y_i^{b_i} \leq 1 \quad \forall i$$

$$x_i y_i^{C_0} \leq 1 \quad \forall i$$

$$x_i^{-1} \leq 1 \quad \forall i$$

$$\frac{1}{e} y_i \leq 1 \quad \forall i$$

$$y_i^{-1} \leq 1 \quad \forall i$$

$$(\bar{\delta})^{\bar{\delta}} (\delta_i)^{\delta_i} \left(\frac{y_i^m}{\rho_0^m M} \right)^{-\delta_i} \frac{1}{y_i} \leq 1 \quad \forall i$$

(3.106)

3.6.3 Existência de Solução

Nesta subseção, apresentaremos as proposições relativas à existência de uma solução ótima para o problema 3.106. Para tanto, precisamos das seguintes hipóteses:

- O número de ativos candidatos à composição da carteira é no mínimo igual a dois, isto é, $N \geq 2$;
- A ponderação mínima de cada ativo na carteira varia entre 0 e $\frac{C_0}{N}$, isto é, $0 < b_i < \frac{C_0}{N}$.

Proposição 3.3. *O problema 3.106 é superconsistente.*

Prova:

Consideremos o ponto

$$t = (x_i^0, y_i^0, \bar{y}_i^0) \in \mathbb{R}^{3N} \quad (3.107)$$

onde

$$x_i^0 = b_i, \forall i$$

$$y_i^0 = 0,6, \forall i$$

$$\bar{y}_i^0 = 1, \forall i$$

Vamos verificar que esse ponto é interior viável:

1. Disponibilidade de Capital:

$$\begin{aligned} \sum_{i=1}^N x_i y_i &= \sum_{i=1}^N b_i \cdot 0,6 \\ &< N \cdot \frac{C_0}{N} \cdot 0,6 \\ &< C_0 \end{aligned}$$

2. Ponderação Mínima:

$$\begin{aligned} x_i &= b_i \\ &> b_i \cdot 0,6 \\ &= b_i y_i \end{aligned}$$

3. Ponderação Máxima:

$$\begin{aligned} x_i &= b_i \\ &\leq \frac{C_0}{N} \\ &\leq \frac{C_0}{2} \\ &< C_0 \cdot 0,6 \end{aligned}$$

4. A positividade de x_i é, assim como a limitação de y_i ao intervalo $(0;1)$ são imediatas;

5. Como $y_i^0, m_z, \rho_0, M > 0$ e $\bar{y}_i^0 = 1$, temos imediatamente que

$$-\frac{(y_i^0)^{m_z}}{\rho_0^{m_z} M} + \bar{y}_i^0 \leq 1$$

Portanto, por definição, o problema primal 3.106 é superconsistente.

Observação: Como relatado anteriormente, são necessárias a consistência e a canonicidade do problema para a provar a existência de uma solução ótima (Teorema A3). A proposição anterior garante a consistência. A canonicidade será suposta baseada na boa formulação do problema (Ver definição A.2 [70]).

Capítulo 4

Metodologia de Decomposição Lagrangeana

4.1 Preliminares

Neste capítulo, apresentamos uma metodologia de relaxação e decomposição, conhecida por Decomposição Lagrangeana, para o modelo não-linear 1.3. O objetivo é particioná-lo em subproblemas de menor porte e de resolução mais simples. O caminho escolhido foi a introdução de uma variável auxiliar y_{ij} , que permite a decomposição do problema em três subproblemas: dois deles envolvendo as restrições de localização de ERB e cobertura da região e o outro envolvendo as restrições de controle de potência e taxa de transmissão. Com isso, obtemos dois subproblemas lineares inteiros e um subproblema não-linear contínuo. Na seção 4.2, apresentamos uma revisão sucinta, com trabalhos que consideramos representativos das metodologias utilizadas nesse capítulo. A seção 4.3 mostra a decomposição lagrangeana.

4.2 Referências bibliográficas

Como uma parte considerável dos problemas reais são de grande porte, as técnicas de decomposição têm despertado muito interesse entre os pesquisadores,

originando uma área de alta produção, diferentes abordagens e com constantes inovações. Nosso objetivo nesta seção é apenas apresentar uma coletânea de trabalhos nas linhas de decomposição, especialmente a decomposição de Dantzig-Wolfe [17], e relaxação lagrangeana, para que tenhamos a idéia de onde foram tiradas as técnicas utilizadas na resolução do modelo.

A utilização de algoritmos de pontos interiores em técnicas de decomposição tem mostrado uma multiplicidade de trabalhos teóricos e de aplicações (ver, por exemplo, [1], [12], [26], [27], [65], [82], [87]). Uma diferença essencial em relação à maioria das demais metodologias é que o problema mestre, seja em formulação primal, seja em dual, não é resolvido exatamente, sendo a iteração, usualmente, um ponto próximo ao centro analítico, definido de modo preciso, através de alguma função potencial ou barreira.

MARTINSON e TIND [51] propõem uma estratégia para encontrar pontos interiores no dual do problema mestre restrito, ao invés de calcular o centro analítico da região viável. Esses pontos estão na vizinhança da trajetória central do problema mestre dual, que é necessariamente suposto limitado. O ponto escolhido fica entre a solução ótima e o centro analítico, fazendo desses os casos extremos.

Com essa proposta, os autores obtêm um método que, para a maioria dos problemas testados, demanda um número menor de iterações do problema mestre em relação ao Dantzig-Wolfe clássico e ao método de centros analíticos. Procedimentos similares também foram implementados e testados em [29].

OLIVEIRA e SANTOS [66] em sua proposta de algoritmo primal para programação convexa não suave, determinam para o politopo de cada iteração um ponto próximo à trajetória central, intermediário entre o centro analítico e a solução

ótima. Ele é obtido através de algumas iterações do algoritmo de centros aplicado ao polítopo atual, o suficiente para garantir o decréscimo de uma cota sobre o custo primal ótimo. Posteriormente SANTOS e OLIVEIRA [81], aplicando a metodologia desenvolvida em [66], propõem duas alternativas para o problema originário da decomposição de Dantzig-Wolfe: a primeira requerendo soluções ótimas dos subproblemas escravos, e a segunda, usando apenas soluções sub-ótimas. No segundo caso, aplica-se o método de centros ao escravo até que a solução sub-ótima produza um corte que separe a iteração corrente, isto é, até que a iteração corrente deixe de ser viável. Verifica-se um potencial de redução da complexidade global dos cálculos¹.

Segundo WENTGES [94], uma das desvantagens do procedimento original de decomposição de Dantzig-Wolfe é o fato de seu problema mestre relaxado gerar multiplicadores de Lagrange muito pobres, com limites inferiores relativamente pequenos. Para suprir essa má utilização e melhor usar a informação gerada pelos subproblemas, Wentges propõe uma modificação, a Decomposição de Dantzig-Wolfe Ponderada, para o procedimento usual. Com isso, obtém uma aceleração na convergência, testada computacionalmente em problemas de localização de facilidades capacitadas (*capacity facility location problem*). Nessa versão, ao invés da solução ótima do problema mestre, o autor usa, como multiplicadores de Lagrange para o próximo subproblema, uma ponderação entre os multiplicadores de Lagrange obtidos previamente e a solução ótima do problema mestre.

KLOSE [40] propõe uma heurística lagrangeana, baseada em uma técnica conhecida como *relax and cut*, para resolver o problema de localização de facilidades capacitadas de dois estágios (TSCFLP). O TSCFLP é uma extensão natural do problema de localização de facilidades capacitadas (CFLP), pois considera o custo

¹Alguns testes com problemas lineares podem ser vistos em MONTEIRO [57].

do transporte entre os pontos.

A técnica mais utilizada para resolver o TSCFLP tem sido a decomposição de Benders [6], por exemplo os trabalhos [24], [35], enquanto que métodos de relaxação lagrangeana são usados com visível sucesso para o CFLP [15]. No trabalho [40] essa última aproximação é utilizada. Relaxando as restrições de capacidade para as fábricas e depósitos, obtém-se um problema que pode ser resolvido por métodos de *branch-and-bound*, cuja melhor cota inferior é computada pela decomposição de Dantig-Wolfe Ponderada [94]. Essa cota e a cota superior são melhoradas através do incorporação de restrições válidas para o TSCFLP (que foi relaxado). Finalmente, essas restrições são dualizadas e relaxadas novamente².

Em relação à metodologia de decomposição lagrangeana, metodologia que remonta a RIBEIRO [72], [73], NAGIH e PLATEAU [60] estendem o resultado clássico de dominação da decomposição lagrangeana sobre a relaxação lagrangeana para a programação fracionária côncava-convexa com restrições lineares, no sentido de termos, para um problema primal de maximização, uma cota superior obtida pelo dual da decomposição lagrangeana menor que a cota superior via dual de relaxação lagrangeana. Essa dominação foi provada para programação linear por GUIGNARD e KIM [33], para programação quadrática convexa por MICHELON e MACULAN [54] e para programação hiperbólica por NAGIH e PLATEAU [61]

4.3 Decomposição Lagrangeana

Nesta seção, abordaremos a resolução do problema 1.3, através do uso de algumas técnicas apresentadas na seção 4.2 e no apêndice A. Como veremos no decorrer da seção, introduziremos uma variável auxiliar de cópia y_{ij} , para podermos separar

²Essa técnica foi primeiramente sugerida para o problema de Steiner, por LUCENA [49].

as variáveis inteiras das variáveis contínuas, em seguida relaxamos algumas restrições e decomposmos o problema em três subproblemas de menor porte. Finalmente, indicamos o caminho utilizado para resolver o problema dual obtido e cada um dos subproblemas. O terceiro deles, por se tratar de um problema não-linear, merecerá uma seção específica.

Passemos ao detalhamento do processo de decomposição lagrangeana:

Nosso primeiro passo foi introduzir a variável de cópia y_{ij} :

$$\min \sum_{i=1}^N c_i z_i + C \cdot \sum_{i=1}^N \sum_{j=1}^M p_{ij}$$

s.a:

$$\begin{aligned} y_{ij} &= x_{ij} && \forall i, \forall j \\ \sum_{i=1}^N a_{ij} x_{ij} &\geq 1, && \forall j \\ \sum_{j=1}^M a_{ij} x_{ij} &\leq \Lambda_i z_i, && \forall i \\ \frac{\sum_{k \neq j} a_{ik} h_k p_{ik} + \sum_{s \neq i} \sum_{t=1}^M a_{sj} h_t p_{st} + \eta_0 W}{W h_j p_{ij}} &\geq \gamma_j r_{ij}, && \forall i, \forall j \\ 0 &\leq p_{ij} \leq P_j y_{ij}, && \forall i, \forall j \\ R_j y_{ij} &\leq r_{ij} && \forall i, \forall j \\ x_{ij}, z_i &\in \{0, 1\} && \forall i, \forall j \\ y_{ij} &\in [0, 1] && \forall i, \forall j \end{aligned} \tag{4.1}$$

Em seguida, encontramos o lagrangeano em função da relaxação dos primeiro e terceiro conjuntos de restrições, aos quais associamos, respectivamente, os multiplicadores $\lambda = \{\lambda_{ij}\}$ e $\mu = \{\mu_i\} \geq 0$:

$$L(z, x, p, r, y, \lambda, \mu) = \sum_{i=1}^N c_i z_i + C \cdot \sum_{i=1}^N \sum_{j=1}^M p_{ij} + \sum_{i=1}^N \sum_{j=1}^M \lambda_{ij} (x_{ij} - y_{ij}) + \sum_{i=1}^N \mu_i \left(\sum_{j=1}^M a_{ij} x_{ij} - \Lambda_i z_i \right) \quad (4.2)$$

A partir dessa relaxação, definamos os conjuntos viáveis para cada conjunto de variáveis:

$$Z = \{z_i; \quad z_i \in \{0, 1\}, \forall i\} \quad (4.3)$$

$$X = \{x_{ij}; \quad x_{ij} \in \{0, 1\}, \quad \forall i, \forall j \quad \text{e} \quad \sum_{i=1}^N a_{ij} x_{ij} \geq 1, \quad \forall j\} \quad (4.4)$$

$$Y = \{(p_{ij}, r_{ij}, y_{ij}) ; \quad 0 \leq p_{ij} \leq P_j y_{ij}, \quad R_j y_{ij} \leq r_{ij}, \quad r_{ij} \geq 0, \quad y_{ij} \in [0, 1]$$

$$\text{e} \quad \frac{Wh_j p_{ij}}{M} \geq \gamma_j r_{ij} \} \quad (4.5)$$

$$\sum_{k \neq j} a_{ik} h_k p_{ik} + \sum_{s \neq i} \sum_{t=1}^M a_{sj} h_t p_{st} + \eta_0 W$$

Observemos que foi acrescentado o conjunto de restrições de não-negatividade para a variável r . Devido ao segundo conjunto de restrições de Y (restrições de taxa mínima), aquele conjunto é redundante e não altera a solução do problema. Inserimo-lo apenas para que possamos considerar como não-negativo todo o conjunto de variáveis do subproblema em (p, r, y) .

A função dual associada é:

$$\theta(\lambda, \mu) = \min_{\substack{z \in Z \\ x \in X \\ (p, r, y) \in Y}} L(z, x, p, r, y, \lambda, \mu) \quad (4.6)$$

O problema dual é dado por:

$$\max_{\substack{\mu \geq 0 \\ \lambda}} \theta(\lambda, \mu) \quad (4.7)$$

Pelas características da função objetivo do modelo original e graças à relaxação utilizada, podemos decompor a função dual da seguinte forma:

$$\begin{aligned} \theta(\lambda, \mu) = & \min_{z \in Z} \left(\sum_{i=1}^N (c_i - \mu_i \Lambda_i) z_i \right) + \min_{x \in X} \left(\sum_{i=1}^N \sum_{j=1}^M (\lambda_{ij} - \mu_i a_{ij}) x_{ij} \right) \\ & + \min_{(p,r,y) \in Y} \left(C \cdot \sum_{i=1}^N \sum_{j=1}^M p_{ij} - \sum_{i=1}^N \sum_{j=1}^M \lambda_{ij} y_{ij} \right) \end{aligned}$$

Assim, a função dual pode ser decomposta na resolução de três subproblemas de menor porte, apresentados nas subseções seguintes.

4.3.1 Subproblema em z

$$\min \left(\sum_{i=1}^N (c_i - \mu_i \Lambda_i) z_i \right)$$

s.a.

$$z_i \in \{0, 1\}$$

Temos um subproblema inteiro que pode ser resolvido por inspeção.

4.3.2 Subproblema em x

$$\min \left(\sum_{i=1}^N \sum_{j=1}^M (\lambda_{ij} - \mu_i a_{ij}) x_{ij} \right)$$

s.a.

$$x \in X$$

onde

$$X = \{x_{ij}; \quad x_{ij} \in \{0, 1\}, \quad \forall i, \forall j \quad \text{e} \quad \sum_{i=1}^N a_{ij} x_{ij} \geq 1, \quad \forall j\}$$

A seguir, mostraremos a importante propriedade de integralidade de X . Relaxando as variáveis do subconjunto X , temos o seguinte conjunto:

$$\bar{X} = \{x_{ij}; 0 \leq x_{ij} \leq 1, \quad \forall i, \forall j \quad \text{e} \quad \sum_{i=1}^N a_{ij}x_{ij} \geq 1, \quad \forall j\}$$

Propriedade 4.1.³ *O subconjunto \bar{X} tem a propriedade da integralidade, portanto a solução ótima em \bar{X} é idêntica à solução ótima em X .*

Demonstração:

Seja a região \bar{X} dada pela união disjunta dos conjuntos RV_j , onde

$$RV_j = \begin{cases} a_{1j}x_{1j} + \dots + a_{Nj}x_{Nj} \geq 1, & a_{ij} \in \{0, 1\}, \quad \forall i \\ x_{ij} \in [0, 1], & \forall i \end{cases}$$

para cada $j \in \{1, \dots, M\}$.

Podemos reescrever RV_j da seguinte forma:

$$RV_j = \{x_j \in \mathbb{R}^N; \quad Ax_j \leq b\}$$

onde $x_j = \{x_{1j}, x_{2j}, \dots, x_{Nj}\}$ e A é uma matriz $2N + 1 \times N$, dada por

$$A = \begin{bmatrix} I_N \\ -a_j^T \\ -I_N \end{bmatrix}$$

onde I_N é uma matriz identidade de ordem N e

$$a_j^T = (a_{1j} \quad a_{2j} \quad \dots \quad a_{Nj})$$

³Agradecemos a colaboração do Prof. Alfredo W. Martins Pinto, da UNAM, na demonstração dessa proposição.

e $b \in \mathbb{R}^{2N+1}$ é um vetor coluna dado por

$$b = \begin{pmatrix} (e)_N \\ -1 \\ (0)_N \end{pmatrix}$$

onde $(e) \in \mathbb{R}^N$ é o vetor de uns, e $(0) \in \mathbb{R}^N$ é o vetor nulo.

Para simplificar a notação, fazemos

$$RV_j = RV = \{x_j \in \mathbb{R}^N; \quad Ax_j \leq b\}$$

$$x_j = x = \{x_1, x_2, \dots, x_N\}$$

$$a_j^T = a^T = \begin{pmatrix} a_1 & a_2 & \dots & a_N \end{pmatrix}$$

Por uma proposição clássica de caracterização dos pontos extremos de um conjunto poliedral (ver, por exemplo, [11] pp. 589 e 590), $v \in RV$ é um ponto extremo, se, e somente se, o conjunto

$$A_v = \{a_k; \quad a_k^T \in A \quad \text{e} \quad a_k^T v = b_k, \quad k = 1, \dots, 2N + 1\} \quad (4.8)$$

contém N vetores linearmente independentes.

As seguintes e exclusivas hipóteses sobre o conjunto dos a_k são possíveis:

a. Consideremos os vetores linearmente independentes dados pelas N linhas de I_N .

O único vetor que satisfaz 4.8 é

$$v = (1, 1, \dots, 1) \in \mathbb{R}^N$$

Como $v \in RV$, v é um vértice de RV .

b. Sejam agora os vetores linearmente independentes dados pelas N linhas de $-I_N$.

O único vetor que satisfaz à condição 4.8 é

$$v = (0, 0, \dots, 0) \in \mathbb{R}^N$$

Porém, como $v \notin RV$, v não é ponto extremo de RV .

c. Consideremos, sem perda de generalidade, os vetores linearmente independentes dados pelas k primeiras linhas de I_N e pelas últimas $N - k$ linhas de $-I_N$, com $1 \leq k < N$. Assim, para satisfazer à condição 4.8, devemos ter

$$\begin{cases} v_i = 1, & \forall i = 1, \dots, k \\ v_i = 0, & \forall i = k + 1, \dots, N \end{cases}$$

Nesse caso, v será vértice se existir pelo menos um índice $i \leq k$ tal que $a_i = 1$. De outro modo, v não será vértice, pois v será inviável. Como a posição das k linhas, e conseqüentemente das $N - k$, é arbitrária, e como a quantidade k , e conseqüentemente a quantidade $N - k$, é variável entre 1 e $N - 1$, temos todos os pontos binários de \mathbb{R}^N com pelo menos uma coordenada igual a 1 e uma coordenada igual a 0.

d. Consideremos a linha $-a^T$ de A . Sem perda de generalidade, consideremos que $-a^T$ é linearmente independente com as $k - 1$ primeiras linhas de I_N e com as $N - k$ últimas linhas de $-I_N$, com $1 \leq k < N$. Por construção, devemos ter a coordenada $a_k = 1$. Assim, o vetor $v \in \mathbb{R}^N$ dado por

$$\begin{cases} v_i = 1, & \forall i = 1, \dots, k-1 \\ v_i = 0, & \forall i = k+1, \dots, N \\ v_k \in [0, 1] \end{cases}$$

satisfaz 4.8. Daí, a restrição

$$a_1 v_1 + \dots + a_{k-1} v_{k-1} + a_k v_k + a_{k+1} v_{k+1} + \dots + a_N v_N = 1$$

dada pela linha $-a^T$ se torna

$$\sum_{i=1}^{k-1} a_i + v_k = 1 \iff v_k = 1 - \sum_{i=1}^{k-1} a_i$$

Como $a_i \in \{0, 1\}$, v_k é inteiro. Como $v_k \in [0, 1] \implies v_k \in \{0, 1\}$. Mais ainda, se existe $i < k$ tal que $a_i = 1$, devemos ter $v_k = 0$. Caso contrário, devemos ter $v_k = 1$.

Portanto, para que v seja ponto extremo de RV , devemos ter

$$v_i \in \{0, 1\}, \quad \forall i = 1, \dots, N$$

Portanto, podemos abordar esse subproblema como subproblema linear que pode ser resolvido por qualquer método de pontos interiores com complexidade polinomial.

4.3.3 Subproblema em (p, r, y)

$$\min(C \cdot \sum_{i=1}^N \sum_{j=1}^M p_{ij} - \sum_{i=1}^N \sum_{j=1}^M \lambda_{ij} y_{ij})$$

s.a.

(4.9)

$$(p_{ij}, r_{ij}, y_{ij}) \in Y$$

onde

$$\begin{aligned}
 Y = \{ (p, r, y) ; & \frac{Wh_j p_{ij}}{\sum_{k \neq j} a_{ik} h_k p_{ik} + \sum_{s \neq i} \sum_{t=1}^M a_{sj} h_t p_{st} + \eta_0 W} \geq \gamma_j r_{ij} , \\
 & p_{ij} - P_j y_{ij} \leq 0 , \\
 & R_j y_{ij} \leq r_{ij} , \\
 & y_{ij} \leq 1 , \\
 & (p, r, y) \geq 0 \}
 \end{aligned} \tag{4.10}$$

Antes de prosseguir com a metodologia de resolução, vamos provar a seguinte propriedade.

Propriedade 4.2. *As restrições de QoS e de taxa de transmissão são ativas na solução ótima.*

Demonstração: Consideremos as condições de otimalidade de Karush-Kuhn-Tucker, observando que as restrições de positividade relativas às variáveis p e r são redundantes e, por isso, podem ser, pelo momento, suprimidas. Além disso, adotaremos a forma quadrática, ao invés da fracionária, para o conjunto de restrições de qualidade de serviço. Seja então $u = (u^1, u^2, u^3, u^4, u^5) \in \mathbb{R}^{5NM}$ o vetor não negativo dos multiplicadores de Lagrange relativos a Y , na respectiva ordem em que as restrições se apresentam, e

$$\begin{aligned}
L(p, r, y) = & C \cdot \sum_{i=1}^N \sum_{j=1}^M p_{ij} - \sum_{i=1}^N \sum_{j=1}^M \lambda_{ij} y_{ij} + \sum_{i=1}^N \sum_{j=1}^M u_{ij}^1 [-Wh_j p_{ij} + d_{ij}(p) \gamma_j r_{ij}] \\
& + \sum_{i=1}^N \sum_{j=1}^M u_{ij}^2 [p_{ij} - P_j y_{ij}] + \sum_{i=1}^N \sum_{j=1}^M u_{ij}^3 [R_j y_{ij} - r_{ij}] + \sum_{i=1}^N \sum_{j=1}^M u_{ij}^4 [y_{ij} - 1] \\
& + \sum_{i=1}^N \sum_{j=1}^M u_{ij}^5 [-y_{ij}]
\end{aligned}$$

a função de Lagrange associada, onde

$$d_{ij}(p) = \sum_{k \neq j} a_{ik} h_k p_{ik} + \sum_{s \neq i} \sum_{t=1}^M a_{sj} h_t p_{st} + \eta_0 W \quad (4.11)$$

As condições de otimalidade para esse subproblema são então dadas por:

$$\begin{aligned}
C + u_{ij}^1 (-Wh_j) + h_j a_{ij} \sum_{k \neq j} \gamma_k u_{ik}^1 r_{ik} + h_j \sum_{s \neq i} \sum_{k=1}^M \gamma_k u_{sk}^1 r_{sk} a_{ik} + u_{ij}^2 &= 0 & \forall i, \forall j \\
u_{ij}^1 \gamma_j d_{ij}(p) - u_{ij}^3 &= 0 & \forall i, \forall j \\
-\lambda_{ij} - u_{ij}^2 P_j + u_{ij}^3 R_j + u_{ij}^4 - u_{ij}^5 &= 0 & \forall i, \forall j \\
-Wh_j p_{ij} + d_{ij}(p) \gamma_j r_{ij} &\leq 0 & \forall i, \forall j \\
p_{ij} - P_j y_{ij} &\leq 0 & \forall i, \forall j \\
R_j y_{ij} - r_{ij} &\leq 0 & \forall i, \forall j
\end{aligned}$$

$$\begin{aligned}
y_{ij} - 1 &\leq 0 && \forall i, \forall j \\
u_{ij}^1 (-Wh_j p_{ij} + d_{ij}(p)\gamma_j r_{ij}) &= 0 && \forall i, \forall j \\
u_{ij}^2 (p_{ij} - P_j) &= 0 && \forall i, \forall j \\
u_{ij}^3 (R_j y_{ij} - r_{ij}) &= 0 && \forall i, \forall j \\
u_{ij}^4 (y_{ij} - 1) &= 0 && \forall i, \forall j \\
u_{ij}^5 (-y_{ij}) &= 0 && \forall i, \forall j \\
(p, r, y) &\geq 0
\end{aligned} \tag{4.12}$$

Do primeiro conjunto, obtemos que $u_{ij}^1 > 0$, pois $C > 0$ e $u_{ij}^2 \geq 0$, $\forall i, j$. Pelo segundo conjunto, obtemos $u_{ij}^3 > 0$, pois possui o mesmo sinal de u_{ij}^1 .

Daí, para satisfazermos a complementaridade com u_{ij}^1 , devemos ter

$$-Wh_j p_{ij} + d_{ij}(p)\gamma_j r_{ij} = 0 \quad \forall i, \forall j$$

$$\text{e, com } u_{ij}^3, \quad R_j y_{ij} - r_{ij} = 0 \quad \forall i, \forall j$$

como queríamos demonstrar.

Agora, levando em conta esta propriedade, substituindo na restrição de QoS a expressão de $r_{ij} = R_i y_{ij}$, dada pela última igualdade, e inserindo as variáveis de folga t_{ij} e v_{ij} , o subproblema 4.9 pode ser reescrito da seguinte forma:

$$\min(C. \sum_{i=1}^N \sum_{j=1}^M p_{ij} - \sum_{i=1}^N \sum_{j=1}^M \lambda_{ij} y_{ij})$$

s.a.

$$Wh_j p_{ij} - d_{ij}(p) \gamma_j R_j y_{ij} = 0 \quad (4.13)$$

$$p_{ij} - P_j y_{ij} + t_{ij} = 0$$

$$y_{ij} + v_{ij} - 1 = 0$$

$$(p, y, t, v) \geq 0$$

onde $d_{ij}(p)$ dada por 4.11, é a notação para o denominador da restrição de QoS.

4.3.4 O Problema Mestre

Apresentaremos, a seguir, o problema mestre associado à decomposição segundo a metodologia de linearização.

Por propriedades conhecidas da teoria de programação matemática, sabemos que, independente da decomposição e da função primal, a função θ (4.6) é côncava não-diferenciável e que existe um oráculo capaz de nos fornecer um subgradiente de θ no ponto (λ, μ) em qualquer iteração. Assim, podemos aplicar qualquer algoritmo baseado nesse oráculo na busca da solução ótima do problema.

Consideremos o problema dual 4.7 dado por

$$\max_{\mu \geq 0} \theta(\lambda, \mu)$$

onde

$$\begin{aligned} \theta(\lambda, \mu) = & \min_{z \in Z} \left(\sum_{i=1}^N (c_i - \mu_i \Lambda_i) z_i \right) + \min_{x \in X} \left(\sum_{i=1}^N \sum_{j=1}^M (\lambda_{ij} - \mu_i a_{ij}) x_{ij} \right) \\ & + \min_{(p, r, y) \in Y} \left(C. \sum_{i=1}^N \sum_{j=1}^M p_{ij} - \sum_{i=1}^N \sum_{j=1}^M \lambda_{ij} y_{ij} \right) \end{aligned}$$

e Z , X e Y são dados por 4.3, 4.4 e 4.5, respectivamente.

Podemos reescrevê-lo como

$$\begin{aligned} & \max r \\ & \text{s. a.} \\ & r \leq L(z, x, p, r, y, \lambda, \mu) \\ & B_{lower} \leq r \\ & \mu \geq 0 \end{aligned}$$

onde o lagrangeano $L(z, x, p, r, y, \lambda, \mu)$ é dado por 4.2 e B_{lower} é dado como no algoritmo (ACA), apresentado mais adiante.

Considerando, em cada iteração,

$$\begin{aligned} s_{\lambda}^k &= (x_{ij}^k - y_{ij}^k) \\ s_{\mu}^k &= \Lambda_i z_i^k - \sum_{j=1}^M a_{ij} x_{ij}^k \\ D(\lambda^k, \mu^k) &= L(z^k, x^k, p^k, r^k, y^k, \lambda^k, \mu^k) \end{aligned}$$

o problema mestre pode ser reescrito como

$$\begin{aligned} & \max r \\ & \text{s. a.} \\ & r \leq D(\lambda^k, \mu^k) + (s_{\lambda}^k, s_{\mu}^k)^T (\lambda - \lambda^k, \mu - \mu^k), \quad \forall k = 1, \dots, K \\ & B_{lower} \leq r \\ & \mu \geq 0 \end{aligned}$$

Finalmente, com a substituição

$$d^k = (s_\lambda^k, s_\mu^k)^T (\lambda^k, \mu^k) - D(\lambda^k, \mu^k), \quad \forall k = 1, \dots, K$$

o problema mestre se torna

$$\max r$$

s.a.

$$(s_\lambda, s_\mu)^T (\lambda, \mu) - r \geq d \tag{4.14}$$

$$r \geq B_{lower}$$

$$\mu \geq 0$$

Observação: O problema mestre 4.14 é um problema linear nas variáveis (u, r) .

Apresentaremos a seguir o algoritmo (ACA), baseado nos trabalhos [66] e [23], utilizado para os testes com a metodologia de decomposição (MI1 no capítulo 6). Mais detalhes e testes podem ser encontrados em [57].

Algoritmo de Centros Analíticos (ACA)

Dados λ^0 e $\mu^0 > 0$,

Tome $k = 0$, $\varepsilon > 0$.

Enquanto $B_{upper} - B_{lower} > \varepsilon$

Passo 1: Resolva o problema escravo (refinamento);

Passo 2: Compute $D(\lambda^k, \mu^k)$;

Passo 3: $B_{lower} = \max\{B_{lower}, D(\lambda^k, \mu^k)\}$;

Passo 4: Incorpore o novo corte e atualize o conjunto de localização.

Faça $k = k + 1$.

Fim do Enquanto

Refinamento

Passo 1: Faça $B_{lower} = D(\lambda^k, \mu^k)$;

Passo 2: Obtenha um ponto inicial para o problema mestre (BIG M);

Passo 3: Obtenha um centro analítico resolvendo o problema 4.14;

Passo 4: Faça $B_{upper} = r$.

Capítulo 5

Um Algoritmo de Pontos Interiores Inviável

Neste capítulo, apresentamos um método primal-dual de pontos interiores, em que introduzimos uma nova função mérito, para a qual a direção de Newton continua sendo de descida. Esse algoritmo será utilizado na resolução dos problemas duais de programação geométrica e dos subproblemas em x e em (p, y, t, v) .

O capítulo está dividido da seguinte maneira: na primeira seção, apresentamos o método propriamente dito: introduzimos a nova função mérito, provamos que a direção de Newton é direção de descida e apresentamos o algoritmo (ABP) e na seção 5.2, aplicamos o método ao subproblema em (p, y, t, v) . Tanto para os duais de programação geométrica, quanto para subproblema em x , utilizamos o algoritmo simplificado para problemas lineares.

5.1 Algoritmo Barreira-Penalidade

Estamos interessados em resolver o seguinte problema de programação matemática:

$$\min f(x)$$

s.a.

$$Q(x) = 0 \tag{5.1}$$

$$x \geq 0$$

onde $f : \mathbb{R}^n \rightarrow \mathbb{R}$ é uma função não-linear, de classe $C^1(\mathbb{R}^n)$, $x \in \mathbb{R}^n$, $b \in \mathbb{R}^{m+1}$ e Q é uma função quadrática.

Sabemos que uma condição necessária a ser verificada por alguma solução de 5.1 é que existam $x \in \mathbb{R}^n$, $y \in \mathbb{R}^{m+1}$ e $s \in \mathbb{R}^n$ tais que

$$Q(x) = 0$$

$$\nabla f(x) - \nabla Q(x)T y - s = 0$$

$$x^t s = 0$$

$$x \geq 0$$

$$s \geq 0.$$

Uma abordagem clássica para a resolução de 5.1 é o método da barreira logarítmica, em que se resolve uma sucessão de problemas perturbados da forma:

$$\min f(x) - \mu \sum_i \ln x$$

s.a.

$$Q(x) = 0 \tag{5.2}$$

$$x > 0$$

com $\mu > 0$. Para o problema 5.2, as condições de KKT são

$$\begin{aligned}
Q(x) &= 0 \\
\nabla f(x) - \nabla Q(x)^T y - s &= 0 \\
x^T s - \mu e &= 0 \\
x &> 0 \\
s &> 0
\end{aligned} \tag{5.3}$$

Da mesma forma, resolver o problema perturbado 5.2 consiste em encontrar $x \in \mathbb{R}^n$, $y \in \mathbb{R}^{m+1}$ e $s \in \mathbb{R}^n$, satisfazendo às restrições 5.3.

Sejam $\Omega = \mathbb{R}_+^n \times \mathbb{R}^{m+1} \times \mathbb{R}_+^n$, $H : \Omega_{+++} \rightarrow \mathbb{R}^{2n+m+1}$ e consideremos o seguinte problema:

$$H(w) = \begin{bmatrix} \nabla f(x) - \nabla Q(x)^T y - s \\ Q(x) \\ x^T s - \mu e \end{bmatrix} = 0 \tag{5.4}$$

onde $w = (x, y, s) \in \Omega = \mathbb{R}_+^n \times \mathbb{R}^{m+1} \times \mathbb{R}_+^n$, $\mu > 0$, A é uma matriz $(m+1) \times n$ de posto completo e $f : \mathbb{R}^n \rightarrow \mathbb{R}$ é uma função não-linear continuamente diferenciável.

Particionando $H(w)$, obtemos

$$F(w) = \begin{bmatrix} \nabla f(x) - \nabla Q(x)^T y - s \\ Q(x) \end{bmatrix}$$

$$G(w) = [x^T s - \mu e]$$

e

$$\Omega_{+++} = \{w \in \Omega; \quad G(w) > 0\}$$

Consideremos a função Barreira-Penalidade $\Psi_{\mu\lambda} : \Omega_{++} \rightarrow \mathbb{R}$ contínua dada por

$$\begin{aligned} \Psi_{\mu\lambda}(w) = & x^t s - \ln \frac{\lambda}{\lambda + \|Q(x)\|^2} - \ln \frac{\lambda}{\lambda + \|\nabla f(x) - \nabla Q(x)^T y - s\|^2} \\ & - \ln \frac{\lambda}{\lambda + \|Xs - \mu e\|^2} - \mu \sum_{i=1}^n \ln x_i s_i \end{aligned} \quad (5.5)$$

onde $\lambda > 0$ é o parâmetro de penalidade, $\mu > 0$ é o parâmetro de barreira e X é uma matriz diagonal cujos elementos não nulos são as componentes do vetor x .

Nosso objetivo é propor um algoritmo para resolver 5.4, conseqüentemente, o problema 5.3, via minimização de $\Psi_{\mu\lambda}$. O algoritmo gerará uma seqüência de vetores $\{w^k\} \subset \Omega_{++}$, cujos pontos de acumulação, se existirem, serão soluções de 5.1.

Comentários sobre a escolha da função Barreira-Penalidade $\Psi_{\mu\lambda}$:

1. As parcelas $-\ln \frac{\lambda}{\lambda + \|Q(x)\|^2}$, $-\ln \frac{\lambda}{\lambda + \|\nabla f(x) - \nabla Q(x)^T y - s\|^2}$ e $-\ln \frac{\lambda}{\lambda + \|Xs - \mu e\|^2}$ forçam a viabilidade da solução e sua proximidade da trajetória central, para $\lambda \rightarrow 0$. Mais ainda, as normas no denominador devem ir para zero mais rápido do que o parâmetro λ . De fato, notemos que

$$\frac{\lambda}{\lambda + \|Q(x)\|^2} = \frac{1}{1 + \frac{\|Q(x)\|^2}{\lambda}}$$

Assim, se $\|Q(x)\|^2 \rightarrow 0$ mais rápido que λ , o quociente tende para 1, e o logaritmo, para zero. Já se $\lambda \rightarrow 0$ mais rápido que o quadrado do resíduo, o denominador tende a ∞ , o quociente tende para zero, o logaritmo para $-\infty$, e a função, conseqüentemente, para ∞ . Fenômeno similar é aplicável aos outros dois termos;

2. A parcela $-\mu \sum_{i=1}^n \ln x_i s_i$, de grande uso em algoritmos primais-duais, impede que x e s se aproximem da fronteira da região de positividade;

3. Como conseqüência do item 1 acima, se $\Psi_{\mu\lambda} \rightarrow 0$, quando $(\mu\lambda) \rightarrow 0$, teremos as condições de KKT (5.3) satisfeitas e, portanto, uma solução para o problema inicial. Como o problema é superconsistente e canônico, pelo Teorema A.4, temos que, para cada $\mu > 0$, o sistema 5.3 possui uma única solução em $\mathbb{R}_{++}^{2n} \times \mathbb{R}^{m+1}$.

5.1.1 A Direção de Newton perturbada

Para um vetor $w \in \Omega_{++}$, a direção de Newton perturbada $\Delta w = (\Delta x, \Delta y, \Delta s)$ em w , aplicada a 5.3, é a solução do seguinte sistema de equações:

$$\begin{pmatrix} \nabla Q(x) & 0 & 0 \\ \nabla^2 f(x) - \nabla^2 Q(x)y & -\text{nabla}Q(x)^T & -I \\ S & 0 & X \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta y \\ \Delta s \end{pmatrix} = - \begin{pmatrix} Q(x) \\ \nabla f(x) - \nabla Q(x)^T y - s \\ Xs - \sigma\mu e \end{pmatrix} \quad (5.6)$$

onde X e S são matrizes diagonais cujos elementos não nulos são, respectivamente, as componentes dos vetores x e s , e $\sigma \in (0, 1)$.

Algoritmo Barreira-Penalidade (ABP)

Dados $x^0 > 0$, $s^0 > 0$, y^0 -livre, $\mu^0 > 0$, $0 < \lambda^0 < 1$, $\delta, \beta, \sigma \in (0, 1)$, $\varepsilon > 0$

Enquanto $\max\{x^k s^k, \|Q(x^k)\|, \|\nabla f(x^k) - \nabla Q(x)^T y^k - s^k\|\} \geq \varepsilon$

Determine, pelo Método de Newton, (x, y, s) tal que

$$\Psi_{\mu\lambda}(x, y, s) \leq \beta \Psi_{\mu\lambda}(x^k, y^k, s^k)$$

Faça

$$\lambda^{k+1} = \delta \lambda^k$$

$$(x^{k+1}, y^{k+1}, s^{k+1}) = (x, y, s)$$

$$\mu^{k+1} = \sigma \frac{(x^{k+1})^T s^{k+1}}{n}$$

$$k \leftarrow k + 1$$

Fim do Enquanto

Lema 5.1. *As seguintes propriedades são válidas para um vetor $w \in \Omega_{++}$ e qualquer $\sigma \in (0, 1)$ e $\mu = \sigma \frac{x^T s}{n}$:*

1. $\Psi_{\mu\lambda}(w) \geq x^t s + n\mu \ln \frac{n}{x^t s}$;
2. $\Psi_{\mu\lambda}$ é continuamente diferenciável em w e seu gradiente é dado por

$$\begin{aligned} \nabla \Psi_{\mu\lambda}^T = & \begin{bmatrix} s \\ 0 \\ x \end{bmatrix}^T + \frac{2}{\lambda + \|Q(x)\|^2} [(Q(x)), (Q(x)), (Q(x))]^t \begin{bmatrix} \nabla Q(x) \\ 0 \\ 0 \end{bmatrix} \\ & + \frac{2}{\lambda + \|\nabla f(x) - \nabla Q(x)^T y - s\|^2} [(\nabla f(x) - \nabla Q(x)^T y - s), (\nabla f(x) - \nabla Q(x)^T y - s), \end{aligned}$$

$$\begin{aligned}
& (\nabla f(x) - \nabla Q(x)^T y - s)]^t \begin{bmatrix} \nabla^2 f(x) - \nabla^2 Q(x)y \\ -\nabla Q(x)^T \\ -I \end{bmatrix} \\
& + \frac{2}{\lambda + \|Xs - \mu e\|^2} [(Xs - \mu e), (Xs - \mu e), (Xs - \mu e)]^t \begin{bmatrix} S \\ 0 \\ X \end{bmatrix} - \mu \begin{bmatrix} X^{-1}e \\ 0 \\ S^{-1}e \end{bmatrix};
\end{aligned}$$

onde onde X^{-1} e S^{-1} são matrizes diagonais cujos elementos não nulos são, respectivamente, os inversos das componentes dos vetores x e s ;

3. Se Δw é uma solução do sistema 5.6, então

$$(\nabla \Psi_{\mu\lambda})^t \Delta w \leq -(1 - \sigma) \left(1 - \frac{\sigma}{n}\right) x^t s.$$

isto e, Δw é uma direção de descida para Ψ

Demonstração:

1. Como para todo $\lambda \geq 0$, x , y e s , vale

$$\left\{ \begin{array}{l} \frac{\lambda}{\lambda + \|Q(x)\|^2} \leq 1, \\ \frac{\lambda}{\lambda + \|\nabla f(x) - \nabla Q(x)^T y - s\|^2} \leq 1, \\ \frac{\lambda}{\lambda + \|Xs - \mu e\|^2} \leq 1, \end{array} \right.$$

temos

$$\left\{ \begin{array}{l} \ln \frac{\lambda}{\lambda + \|Q(x)\|^2} \leq 0, \\ \ln \frac{\lambda}{\lambda + \|\nabla f(x) - \nabla Q(x)^T y - s\|^2} \leq 0, \\ \ln \frac{\lambda}{\lambda + \|Xs - \mu e\|^2} \leq 0. \end{array} \right.$$

Daí,

$$\Psi_{\mu\lambda} \geq x^t s - \mu \sum_{i=1}^n \ln x_i s_i = x^t s - \mu \ln \prod_{i=1}^n x_i s_i$$

Pela desigualdade entre as médias aritmética e geométrica, (Lema 3.1), temos

$$\prod x_i s_i \leq \left(\frac{\sum x_i s_i}{n} \right)^n = \left(\frac{x^t s}{n} \right)^n$$

Daí, como a função logarítmica é crescente,

$$\ln \prod x_i s_i \leq \ln \left(\frac{x^t s}{n} \right)^n = n \ln \frac{x^t s}{n} = -n \ln \frac{n}{x^t s}$$

Portanto,

$$\begin{aligned} \Psi_{\mu\lambda} &\geq x^t s + n\mu \ln \frac{n}{x^t s} \\ &= n\mu + n\mu \ln \frac{1}{\mu} \\ &= n\mu(1 - \ln \mu) \\ &\geq n\mu(1 - \mu) \end{aligned}$$

que é positivo para $0 < \mu < 1$. Para essa conclusão, consideramos $\mu = \frac{x^t s}{n}$, pois $\Psi_{\mu\lambda}$ atinge seu mínimo na trajetória central.

2. Trivial.

3.

$$\begin{aligned}
(\Psi_\mu)^t \Delta w &= s \Delta x + x \Delta s + \frac{2}{\mu + \|Q(x)\|^2} Q(x)^T (\nabla Q(x) \Delta x) \\
&+ \frac{2}{\mu + \|\nabla f(x) - \nabla Q(x)^T y - s\|^2} (\nabla f(x) - \nabla Q(x)^T y - s)^t (\nabla^2 f(x) \Delta x - \nabla Q(x)^T \Delta y - \Delta s) \\
&+ \frac{2}{\mu + \|Xs - \mu e\|^2} (Xs - \mu e)^t (S \Delta x + X \Delta s) \\
&= s \Delta x + x \Delta s - \mu \left((X^{-1}e)^T \Delta x + (S^{-1}e)^T \Delta s \right) \\
&- \frac{2\|Q(x)\|^2}{\mu + \|Q(x)\|^2} - \frac{2\|\nabla f(x) - \nabla Q(x)^T y - s\|^2}{\mu + \|\nabla f(x) - \nabla Q(x)^T y - s\|^2} - \frac{2\|Xs - \mu e\|^2}{\mu + \|Xs - \mu e\|^2} \\
&\leq s \Delta x + x \Delta s - \mu (X^{-1}e \Delta x + S^{-1}e \Delta s) \\
&\leq -(1 - \sigma^2) \left(1 - \frac{\sigma}{n}\right) x^t s
\end{aligned}$$

onde $\mu = \sigma \frac{x^t s}{n}$ e a segunda desigualdade decorre do fato de Δw ser solução do sistema 5.6 e de

$$\begin{aligned}
(X^{-1}e)^T \Delta x + (S^{-1}e)^T \Delta s &= \sum_i \frac{s_i \Delta x_i + x_i \Delta s_i}{x_i s_i} \\
&\geq \frac{\sum_i s_i \Delta x_i + x_i \Delta s_i}{x^T s} \\
&= \frac{-x^T s + \sigma \mu n}{x^T s}
\end{aligned}$$

Observação: Com o suporte teórico acima, testamos esse algoritmo para uma série de problemas teste. Os resultados estão apresentados no apêndice B.

Algoritmo Signomial (ASIG)

Dado $t^0 > 0$,

Tome $k = 0, \varepsilon > 0$

Enquanto $Paradasig = \|\|t^{k+1} - t^k\|\| > \varepsilon$

Passo 1: Condense as restrições signomiais usando t^k como ponto de referência;

Passo 2: Use (ABP) para resolver o problema de programação geométrica obtido no passo 1, sendo t sua solução;

Passo 3: Faça $t^{k+1} = t$;

Passo 4: Faça $k = k + 1$.

Fim do Enquanto

5.2 Método aplicado ao subproblema em (p, y, t, v)

As condições de otimalidade de KKT para o problema 4.13 irrestrito perturbado são, $\forall i, \forall j$:

$$\begin{aligned}
 C - w_{ij}^1 W h_j + h_j a_{ij} \sum_{k \neq j}^M \gamma_k w_{ik}^1 R_k y_{ik} + h_j \sum_{s \neq i}^N \sum_{k=1}^M \gamma_k w_{sk}^1 R_k y_{sk} a_{ik} + w_{ij}^2 - \theta_{ij}^1 &= 0 \\
 -\lambda_{ij} + w_{ij}^1 d_{ij}(p) \gamma_j R_j - w_{ij}^2 P_j + w_{ij}^3 - \theta_{ij}^2 &= 0 \\
 w_{ij}^2 - \theta_{ij}^3 &= 0 \\
 w_{ij}^3 - \theta_{ij}^4 &= 0 \\
 \theta_{ij}^1 p_{ij} &= \mu \\
 \theta_{ij}^2 y_{ij} &= \mu \\
 \theta_{ij}^3 t_{ij} &= \mu \\
 \theta_{ij}^4 v_{ij} &= \mu \\
 W h_j p_{ij} - d_{ij}(p) \gamma_j R_j y_{ij} &= 0 \\
 p_{ij} - P_j y_{ij} + t_{ij} &= 0 \\
 y_{ij} + v_{ij} - 1 &= 0 \\
 (p, y, t, v, \theta^1, \theta^2, \theta^3, \theta^4) &\geq 0
 \end{aligned} \tag{5.7}$$

onde $w = \{(w_{ij}^1, w_{ij}^2, w_{ij}^3)\}$ e $\theta = \{(\theta_{ij}^1, \theta_{ij}^2, \theta_{ij}^3, \theta_{ij}^4)\} \geq 0$ os vetores dos multiplicadores de Lagrange para as restrições de igualdade e de desigualdade, respectivamente.

Além das condições de complementaridade, apenas o conjunto de restrições de QoS (nono conjunto) é não linear. Entretanto, já sabemos que são restrições positivas para qualquer solução viável. Assim, para utilização do método Barreira-

Penalidade, apresentado na seção anterior, consideraremos uma perturbação para este conjunto de restrições, isto é

$$Wh_j p_{ij} - d_{ij}(p) \gamma_j R_j y_{ij} = \mu$$

e o seguinte particionamento

$$F(q) = \left[\begin{array}{c} C - w_{ij}^1 Wh_j + h_j a_{ij} \sum_{k \neq j}^M \gamma_k w_{ik}^1 R_k y_{ik} + h_j \sum_{s \neq i}^N \sum_{k=1}^M \gamma_k w_{sk}^1 R_k y_{sk} a_{ik} + w_{ij}^2 - \theta_{ij}^1 \\ -\lambda_{ij} - w_{ij}^1 d_{ij}(p) \gamma_j R_j - w_{ij}^2 P_j + w_{ij}^3 - \theta_{ij}^2 \\ w_{ij}^2 - \theta_{ij}^3 \\ w_{ij}^3 - \theta_{ij}^4 \\ p_{ij} - P_j y_{ij} + t_{ij} \\ y_{ij} + v_{ij} - 1 \end{array} \right]_{6NM \times 1}$$

$$G(q) = \left[\begin{array}{c} \theta_{ij}^1 p_{ij} - \mu \\ \theta_{ij}^2 y_{ij} - \mu \\ \theta_{ij}^3 t_{ij} - \mu \\ \theta_{ij}^4 v_{ij} - \mu \\ Wh_j p_{ij} - d_{ij}(p) \gamma_j R_j y_{ij} - \mu \end{array} \right]_{5NM \times 1}$$

onde $q = (p, y, t, v, \theta^1, \theta^2, \theta^3, \theta^4, w^1, w^2, w^3) \in \mathbb{R}^{11NM}$.

Identificando com o problema 5.4, resolver o subproblema em (p, y, t, v) se resume a encontrar $q \in \Omega$ tal que

$$H(q) = \begin{bmatrix} F(q) \\ G(q) \end{bmatrix}_{(11NM) \times 1} = 0$$

onde $\Omega = \{q \in \mathbb{R}^{11NM} ; 0 \leq p_{ij} ; 0 \leq y_{ij} ; t_{ij}, v_{ij}, \theta_{ij}^1, \theta_{ij}^2, \theta_{ij}^3, \theta_{ij}^4 \geq 0\}$. Além disso, a partição $G(q)$ incorpora a viabilidade primal de QoS.

Capítulo 6

Implementações e Resultados Computacionais

Neste capítulo, apresentaremos os resultados computacionais obtidos para as seguintes metodologias de implementação:

- (MI1) Inspeção aplicada ao subproblema em z e método de pontos interiores, apresentado na seção 5.1, aplicado aos subproblemas em x e em (p, r, y) originários da decomposição de 1.3 (abordagem apresentada no capítulo 4); Método de Centros Analíticos (ACA) aplicado ao dual do problema 1.3;
- (MI2) Programação signomial 0-1 para a relaxação do problemas 1.3, usando o ponto inicial dado por 3.82 (abordagem apresentada no capítulo 3).

O modelo 3.65, linearizado pela aproximação de Taylor na restrição de QoS (ver [57]), ainda foi implementado no *software* XPRESS [96] no Sistema Operacional LINUX e executado em um PC Pentium III 600MHz. A metodologia (MI1) foi implementada em ambiente *MATLAB for Windows* [53] e executada em um Pentium III 850MHz. A metodologia (MI2) também foi implementada em ambiente *MATLAB for Windows* [53], porém executada em um Pentium 4 1,5GHz.

Os parâmetros para um sistema CDMA utilizados nas simulações estão dispostos na tabela abaixo. Os intervalos foram baseados nos valores apresentados em [78].

ITEM	SÍMBOLO	VALOR
Custo de instalação	c_i	$[29; 37] \times 10000$ R\$
Custo da potência	C	8 R\$/W
Largura da banda	W	1250 kHz
Taxa mínima de transmissão	R_j	$[7, 8; 8, 1]$ kbps
QoS requerido	γ_j	$[0, 5; 0, 7]$ dB
Potência máxima de transmissão	P_j	$[0, 48; 0, 52]$ W
Ganho por canal	h_j	$[0, 23; 0, 26]$
Ruído gaussiano	η_0	10^{-6}

(6.1)

Problema 1

O primeiro problema simulado foi a cobertura de uma região dividida em 8 quadrículas, a partir de um subconjunto do conjunto de 4 locais candidatos à instalação de uma ERB, gerando um total de 132 variáveis (das quais 36 inteiras) e 212 restrições, para a metodologia de decomposição lagrangeana e um total de 136 variáveis e 284 restrições para a metodologia de programação geométrica signomial. Os resultados obtidos estão dispostos a seguir:

Para a Metodologia MI1:

No. de it. de Newton ⁽¹⁾	No. médio de it. ⁽²⁾	Tempo (s)	Gap
18	57	44,65	$3,67 \cdot 10^{-5}$

No. de it. PM ⁽³⁾	Solução ótima inteira	ERBs escolhidas
14	65,6548	z_1, z_4

⁽¹⁾ No. de iterações de Newton para obtenção de um ponto interior.

⁽²⁾ No. médio de iterações de Newton para obtenção de um centro analítico, por iteração externa.

(3) No. de iterações para o Problema Mestre.

Para a Metodologia MI2:

Os parâmetros usados foram:

$$m_z = 5$$

$$m_x = 2$$

$$M = 2$$

$$\rho_0 = \frac{1+e}{2}$$

Os resultados obtidos foram:

No. de it. signomiais	Solução ótima inteira	<i>Paradasig</i>	ERBs escolhidas
18	67,9738	$O(10^{-5})$	z_1, z_4

Para o *software* XPRESS:

Problema relaxado (contínuo):

Número de iterações	Solução ótima contínua
63	53,6253

Problema inteiro, a partir da relaxação:

Nó / Nós abertos ⁽⁴⁾	Melhor cota obtida
2/3	60,9253

Solução ótima inteira	ERBs escolhidas
68,1253	z_1, z_4

⁽⁴⁾ Para melhor entendimento, sugerimos a consulta de [50] pp. 94 a 99.

Problema 2

O segundo problema simulado foi o seguinte: uma região dividida em 20 quadrículas, onde foram distribuídos 6 pontos, candidatos à instalação de uma ERB, em locais aleatórios, gerando um total de 486 variáveis (das quais 126 inteiras) e 758 restrições, para a metodologia de decomposição lagrangeana e um total de 492 variáveis e 1010 restrições para a metodologia de programação geométrica signomial. Os resultados estão a seguir:

Para a Metodologia MI1:

No. de it. de Newton	No. médio de it.	Tempo (s)	Gap
19	61	193,82	$7,4 \cdot 10^{-6}$

No. de it. PM	Solução ótima inteira	ERBs escolhidas
20	97,5775	z_1, z_2, z_5

Para a Metodologia MI2:

Os parâmetros usados foram:

$$m_z = 7$$

$$m_x = 2$$

$$M = 2$$

$$\rho_0 = \frac{1+e}{2}$$

Os resultados obtidos foram:

No. de it. signomiais	Solução ótima inteira	<i>Paradasig</i>	ERBs escolhidas
24	99,4344	$O(10^{-4})$	z_1, z_2, z_5

Para o *software* XPRESS:

Número de iterações	Solução ótima contínua
225	91,0426

Problema inteiro, a partir da relaxação:

Nó / Nós abertos	Melhor cota obtida
9/17	93,9618

Solução ótima inteira	ERBs escolhidas
116,4164	z_1, z_2, z_5

Observação: Cada um dos dois problemas analisados apresenta diferentes valores para o ótimo. Como as ERBs escolhidas são as mesmas, essa diferença se deve exclusivamente aos valores ótimos da variável p .

Capítulo 7

Proposta de Trabalhos Futuros

7.1 Quanto à modelagem

Acreditamos que os modelos desenvolvidos podem ser melhorados, de modo a abranger e gerar resultados mais próximos da realidade. A seguir, enumeramos alguns aperfeiçoamentos que podem ser testados:

1. Generalizar o modelo de telecomunicação sem fio para qualquer tipo de serviço: voz, dados, internet [30], [31], [98], [99]. No que diz respeito à transmissão de dados, SARAYDAR C. *et al.* [84] apresentam uma solução para o problema de controle de potência de um sistema unicelular segundo a teoria dos jogos. No contexto estudado, a qualidade de serviço que um terminal recebe é interpretada como uma função utilidade e o controle de potência como um jogo não-cooperativo, em que a solução, fornecida pelo equilíbrio de Nash [64], é provada ineficiente. Os autores introduzem uma precificação da transmissão de potência, de modo a obter um Pareto melhoramento (ver definição em [84]) para a solução do jogo. Em outro trabalho [83], os mesmos autores aplicam a abordagem anterior para um sistema multicelular, obtendo resultados semelhantes. Por isso, acreditamos que tal abordagem pode ser utilizada em nosso

modelo;

2. Possibilitar a escolha do tipo de antena utilizada para transmissão, visando melhores resultados;
3. Acrescentar um estudo de propagação de sinal;
4. Simular resultados a partir de problemas envolvendo áreas de sombra;
5. Generalizar o modelo de carteiras de investimentos para vários investidores.

7.2 Quanto à implementação

Estamos em fase final de implementação da metodologia de programação geométrica signomial 0-1 aplicada ao modelo de otimização de carteiras, usando o ponto inicial 3.107.

Temos ainda a intenção de implementar todas as metodologias em linguagem C, para testarmos o modelo para problemas de maior porte. Porém, independente dessa implementação, pretendemos testar os algoritmos referentes às metodologias estudadas nesse trabalho em outros modelos de telecomunicação sem fio, inclusive o modelo MCP (1.2).

7.3 Quanto à teoria

Temos a intenção de fortalecer os resultados computacionais, demonstrando resultados teóricos acerca da metodologia de programação geométrica signomial.

Apêndice A

Elementos de Programação Geométrica

O objetivo desta seção é apenas expor o problema de programação geométrica posinomial e alguns resultados conhecidos, necessários para o entendimento da metodologia utilizada. Para elaboração desse texto em especial, baseamo-nos em [62] e [7].

A Programação Geométrica foi desenvolvida inicialmente na década de 60 por DUFFIN *et al.* [21]. Recentemente, vários autores têm voltado seu interesse para essa técnica na tentativa de construir algoritmos de pontos interiores eficientes para esses problemas, veja [42], [43] e [97], este último tendo tratado de um problema mais geral, conhecido por problema de programação geométrica signomial.

O Problema Primal de Programação Geométrica (PPG) consiste em

$$\min g_0(t)$$

s.a.

$$g_k(t) \leq 1, \quad k = 1, \dots, p$$

$$t \in \mathbb{R}_{++}^m$$

onde

$$g_k(t) = \sum_{i \in J[k]} c_i \prod_{j=1}^m t_j^{a_{ij}}, \quad k = 0, 1, \dots, p$$

$$J[k] = \{m_k, m_{k+1}, \dots, n_k\}, \quad k = 0, 1, \dots, p \quad (A.1)$$

$$m_0 = 1, \quad m_1 = n_0 + 1, \quad m_2 = n_1 + 1, \quad \dots, \quad m_p = n_{p-1} + 1, \quad n_p = n,$$

os expoentes a_{ij} são constantes arbitrárias, os coeficientes c_i são positivos. As funções g_k são chamadas posinômios e os termos $c_i \prod_{j=1}^m t_j^{a_{ij}}$ são chamados termos posinomiais, n_0 é o número de termos posinomiais na função objetivo, n é o número de termos posinomiais existentes no problema e as variáveis t_j são as variáveis primais.

O Problema Dual de Programação Geométrica (DPG) consiste em

$$\max u(x)$$

s.a.

$$\sum_{i \in J[0]} x_i = 1$$

$$\sum_{i=1}^n a_{ij} x_i = 0$$

$$x \in \mathbb{R}_+^n$$

onde

$$u(x) = \prod_{i=1}^n \left(\frac{c_i}{x_i} \right)^{x_i} \prod_{k=1}^p \lambda_k^{\lambda_k}$$

$$\lambda_k = \sum_{i \in J[k]} x_i$$

$J[k]$ é dado por (A.1)

Definição A.1. Dizemos que o Problema (Primal ou Dual) é consistente se existe pelo menos um ponto viável. O Problema Primal é dito superconsistente se satisfaz à condição de Slater, isto é, existe

$$t^* \in \mathbb{R}_{++}^m$$

tal que

$$g_k(t^*) < 1 \quad k = 1, \dots, p$$

A hipótese da superconsistência garante a existência de multiplicadores de Kuhn-Tucker, o que torna simples a prova da existência de uma solução ótima dual com valor objetivo ótimo igual ao ótimo primal. Sem a superconsistência, o valor ótimo dual não existe necessariamente.

Teorema A.1. *Suponha que o problema primal (PPG) é superconsistente e que sua função objetivo $g_0(t)$ atinja seu valor mínimo em um ponto que satisfaz às restrições primais. Então:*

1. *O problema dual correspondente (DPG) é consistente e a função dual $u(x)$ atinge seu valor máximo em um ponto que satisfaz às restrições duais;*
2. *O valor ótimo da função dual é igual ao valor ótimo da função primal;*
3. *Se t' é um ponto de mínimo do problema primal, existem multiplicadores de Lagrange μ'_k , $k = 1, 2, \dots, p$, não-negativos, tais que a função de Lagrange*

$$L(t, \mu) = g_0(t) + \sum_{k=1}^p \mu_k [g_k(t) - 1]$$

possui a propriedade

$$L(t', \mu) \leq g_0(t') = L(t', \mu') \leq L(t, \mu')$$

para $t_j > 0$ e $\mu_k \geq 0$ arbitrários. Mais ainda, existe um ponto de máximo x' para o problema dual (DPG) cujas componentes são

$$x'_i = \begin{cases} \frac{c_i t_1^{a_{i1}} \dots t_m^{a_{im}}}{g_0(t)}, & i \in J[0], \\ \frac{\mu_k c_i t_1^{a_{i1}} \dots t_m^{a_{im}}}{g_0(t)}, & i \in J[k], \quad k = 1, \dots, p, \end{cases}$$

onde $t = t'$ e $\mu = \mu'$. Além disso,

$$\lambda_k(x') = \frac{\mu'_k}{g_0(t')}, \quad k = 1, 2, \dots, p.$$

4. Se x' é um ponto de máximo para o problema dual (DPG), cada ponto de mínimo t' do problema primal (PPG) satisfaz o seguinte sistema de equações:

$$c_i t_1^{a_{i1}} \dots t_m^{a_{im}} = \begin{cases} x'_i u(x'), & i \in J[0], \\ \frac{x'_i}{\lambda_k(x')}, & i \in J[k], \end{cases}$$

para todo k inteiro positivo tal que $\lambda_k(x') > 0$.

Prova: ver [21]

Teorema A.2. Se o Problema Primal (PPG) é consistente e existe um ponto x^* com componentes positivas satisfazendo às restrições do problema dual (DPG), a função primal $g_0(t)$ atinge seu valor mínimo em um ponto t' que satisfaz às restrições do problema primal.

Prova: ver [21]

Definição A.2. Um problema de programação geométrica posinomial e seu dual são ditos degenerados se pelo menos um dos termos

$$c_{ki} \prod_{j=1}^m t_j^{a_{kij}}$$

se aproxima de zero sem forçar que outro termo primal tenda a mais infinito. Problemas não-degenerados são ditos canônicos¹.

¹O estudo de problemas degenerados pode ser reduzido ao estudo de problemas canônicos equivalentes, veja [21].

Uma definição equivalente para degenerescência é que um problema primal posinomial e seu dual são ditos degenerados se existe um vetor não-nulo, não-positivo no espaço gerado pelas colunas da matriz de expoentes. PETERSON [70] destaca que os problemas de programação geométrica posinomial, se bem formulados, parecem ser todos canônicos na prática.

Observação: A prova da canonicidade de um problema posinomial não é trivial. De fato, dado que as linhas de uma matriz de expoentes A correspondem a termos posinimiais (e as colunas correspondem as variáveis de decisão), devemos checar se cada coluna possui elementos positivos e negativos:

- se existe uma coluna em que isso não acontece, A não é canônica (e a variável de decisão correspondente vai se aproximar de zero ou de mais infinito durante a minimização);
- se todas as colunas satisfizerem, precisamos testar se o sistema linear

$$yA = 0, y \geq 1,$$

possui uma solução viável (onde y um vetor linha e a restrição $y \geq 1$ significa que cada componente de y deve ser maior ou igual a 1. A matriz A será canônica se e somente se existir tal solução viável.

De fato, se o segundo item acontecer, basta notar que tal vetor solução y , quando dividido pela soma de suas componentes, gera um outro vetor que claramente possui todas as coordenadas positivas e satisfaz às condições de ortogonalidade e normalidade, definição de canonicidade.

Definição A.3. *O grau de dificuldade de um Problema de Programação Geométrica é dado por*

$$d = n - (m + 1)$$

onde n é o número de termos posinomiais e m é o número de variáveis primais.

Teorema A.3. *Suponha que o problema é canônico e que o Primal (PPG) é consistente. Então existe t^* tal que*

$$g_0^* = g_0(t^*)$$

onde

$$g_0^* = \min_{t \in X} g_0(t) \quad X = \{t : t_j > 0, g_k(t) \leq 1\}$$

Prova: ver [21]

Teorema A.4. *Suponha que o Problema Primal (PPG) seja superconsistente e canônico. Então, dado $\mu > 0$, o sistema de equações 5.3 possui uma única solução em $\mathbb{R}_{++}^{2n} \times \mathbb{R}^{m+1}$.*

Prova: ver [62]

Apêndice B

Resultados Computacionais

Nesse apêndice, apresentamos os resultados computacionais para os algoritmos ABP e ASIG, testados em problemas de pequeno porte. Todos os problemas foram executados em um Pentium 4 1,5GHz.

Algoritmo Posinomial Barreira-Penalidade

A legenda para leitura da tabela é:

V P:	Número de Variáveis Primais
R P:	Número de Restrições Primais
T:	Número de Termos
G D:	Grau de Dificuldade (ver apêndice A)
It:	Número de Iterações
CPU:	Tempo de CPU em segundos
V Ot P:	Valor Ótimo Primal
Gap:	Gap de Dualidade

Problema	V P	R P	T	G D	It	CPU	V Ot P	Gap
Beck751	7	4	18	10	40	0,77	1,8098.10 ³	2,0478.10 ⁻⁷
Beck752	7	4	18	10	40	0,72	911,8806	8,5658.10 ⁻⁸
Beck753	7	4	18	10	14	0,22	543,6680	6,7987.10 ⁻¹⁰
Dembo761	12	3	31	18	40	2,48	4,8907.10 ⁹	2,0712.10 ⁻⁵
Dembo761	12	3	31	18	40	2,48	4,8907.10 ⁹	2,0712.10 ⁻⁵
Dembo762	12	3	31	18	40	2,48	3,1684	4,6434.10 ⁻⁸
Kort8	7	7	48	40	35	2,25	178,4779	5,5537.10 ⁻⁹
Kort10	10	7	20	9	16	0,28	3,2236.10 ⁷	5,5463.10 ⁻¹¹
Kort11	22	36	73	50	40	6,59	1,8318.10 ³	3,5611.10 ⁻⁸
Rijck781	4	3	31	19	8	0,05	0,0121	6,2613.10 ⁻¹²
Rijck782	3	1	9	5	16	0,11	6,2998.10 ³	4,9103.10 ⁻¹¹
Rijck783	4	1	12	7	10	0,11	1,263.10 ⁵	1,7596.10 ⁻¹¹
Rijck785	4	3	8	3	11	0,05	6,2325.10 ⁵	1,7980.10 ⁻¹¹

Algoritmo Signomial

Capital Budgeting Problem [90]

$$\max z = 20x_1 + 40x_2 + 20x_3 + 15x_4 + 30x_5$$

s.a.

$$5x_1 + 4x_2 + 3x_3 + 7x_4 + 8x_5 \leq 25$$

$$x_1 + 7x_2 + 9x_3 + 4x_4 + 6x_5 \leq 25$$

$$8x_1 + 10x_2 + 2x_3 + x_4 + 10x_5 \leq 25$$

$$x_1, x_2, x_3, x_4, x_5 \in \{0, 1\}$$

Os parâmetros usados foram:

$$m = 30$$

$$M = 2$$

Os resultados obtidos foram:

Iterações signomiais: 5

Custo na Solução ótima: 95

Paradasig: $6,42 \cdot 10^{-10}$

Set Covering Problem [90]

$$\min x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 + x_8$$

s.a.

$$x_1 + x_2 \geq 1$$

$$x_2 + x_3 \geq 1$$

$$x_4 + x_5 \geq 1$$

$$x_7 + x_8 \geq 1$$

$$x_6 + x_7 \geq 1$$

$$x_2 + x_6 \geq 1$$

$$x_1 + x_6 \geq 1$$

$$x_4 + x_7 \geq 1$$

$$x_2 + x_4 \geq 1$$

$$x_5 + x_8 \geq 1$$

$$x_3 + x_5 \geq 1$$

$$x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8 \in \{0, 1\}$$

Os parâmetros usados foram:

$$m = 2$$

$$M = 2$$

Os resultados obtidos foram:

Iterações signomiais:	4
Custo na Solução ótima:	4
<i>Paradasig</i> :	$2,41 \cdot 10^{-10}$

Referências Bibliográficas

- [1] ATKINSON, D. S. e VAIDYA, P. M. A cutting plane algorithm for convex programming that uses analytic centers, *Mathematical Programming*, v.69, pp.1-43, 1985.
- [2] AVRIEL, M. e WILLIAMS, A. C. Complementary geometric programming, *SIAM Journal on Applied Mathematics* v.19, pp.125-141, 1970.
- [3] BAZARAA, M. S., SHERATI, H. D., SHETTY, C. M., *Nonlinear Programming: Theory and Algorithms* John Wiley & Sons Inc., 1993.
- [4] BECK, P. A. e ECKER, J. G. A modified concave simplex algorithm for geometric programming, *Journal of Optimization Theory and Applications* v.15, pp.189-202, 1975.
- [5] BECK, P. A. A Modified Convex Simplex Algorithm for geometric Programming with Subsidiary Problems, PhD Thesis, Rensseler Polytechnic Inst., 1972.
- [6] BENDERS, J. F. Partitioning procedures for solving mixed variables programming problems, *Numerische Mathematik* v. 4, pp.238-252, 1962.
- [7] BEIGHTLER, C. S., PHILIPS, D. T., *Applied Geometric Programming* New York, John Wiley & Sons, 1976.

- [8] BLAU, G. E. e WILDE, D. J., Generalized Polynomial Programming, *Canadian Journal of Chemical Engineering* vol. 47,, 1969.
- [9] BREALEY, R. A & MYERS, S. C., *Principles of Corporate Finance*, 6th ed., Mc Graw-Hill, 2000.
- [10] BRIGHAM, E. F, GAPENSKI, L. C., EHRHARDT, M. C., *Administração Financeira - Teoria e Prática*, São Paulo, Editora Atlas S.A., 2001.
- [11] BERTSEKAS, D. P., *Nonlinear Programming* Belmont, Massachusetts, Athena Scientific, 1995.
- [12] CHARALAMBOUS, C. e CONN, A. R. An efficient method to solve the minimax problem directly, *SIAM Journal on Numerical Analysis* v.15, pp.162-187, 1978.
- [13] CHOI, J. C. e BRICKER, D. L. A heuristic procedure for rounding posynomial geometric programming solutions to discrete values, *Computers ind. Engng* v. 30, n.4, pp. 623-629, 1996.
- [14] COELHO, L. R., GARCIA, N. A. P., RODRIGUES, M. E. C., MELLO, L. S., *Sistemas Rádio Celulares e de Rádio Acesso - Parte 1*, CETUC - PUC, Rio de Janeiro, RJ, 2001.
- [15] CORNUÉJOLS, G., SRIDHARAN, R., THIZY, J. M. A comparison of heuristics and relaxation for the capacited plant location problem, *European Journal of Operational Research* v. 50, pp. 280-297, 1991.
- [16] DAMODARAN, A., *Finanças Corporativas Aplicadas: Manual do Usuário*, Porto Alegre, Bookman, 2002.

- [17] DANTZIG, G. B., WOLFE, P. The decomposition algorithm for linear programming, *Econometrica* v. 29, n. 4, pp. 767-778, 1961.
- [18] DEMBO, R. S. A set of geometric programming test problems and their solutions, *Mathematical Programming* v. 10, pp.192-213, 1976.
- [19] DUARTE, A. M. Jr. Risco: Definições, Tipos, Medição e Recomendações para seu Gerenciamento, *Resenha BM&F* v. 114, pp.25-33, 1996.
- [20] DUARTE, A. M. Jr. A framework for the active management of a global currency fund, *Revista de Econometria* v.17, n.2, pp.1-19, 1997.
- [21] DUFFIN, R. J., PETERSON, E. L., ZENER, C., *Geometric Programming - Theory and Applications* New York/London/Sidney, Jonh Wiley & Sons Inc., 1967.
- [22] DUFFIN, R. J. e PETERSON, E. L. Reversed geometric programming treated by harmonics means, *Indiana University Mathematics Journal* v. 22, pp.531-550, 1972.
- [23] ELHEDHLI, S., GOFFIN, J. -L. The integration of an Interior-Point Cutting Plane method within a Branch and Price Algorithm, *GERAD*, Faculty of Management, Mc Gill University, Montreal, Canada, March 2001.
- [24] GEOFFRION, A. M., GRAVES, G. W. Multicommodity distribution system design by Benders decomposition, *Management Science* v. 20, pp. 822-844, 1974.
- [25] GILHOUSEN, K. S., JACOBS, I. M., PADOVANI, R., VITERBI, A. J.,

- WEAVER, L. A., WHEATLEY III, C. E. On the capacity of a cellular CDMA system, *IEEE Trans. Veh. Technol.* v. 40, pp.303-312, 1991.
- [26] GOFFIN, J.-L., HAURIE, A., VIAL, J.-P. Decomposition and nondifferentiable optimization with the projective algorithm, *Management Science* v.38, pp.284-302, 1992.
- [27] GOFFIN, J.-L., VIAL, J.-P. Cutting planes and column generation techniques with projective algorithm, *Journal of Optimization Theory and Applications* v.65, pp.409-429, 1990.
- [28] GOMES-RUGGIERO, M. A., MARTINEZ, J. M., MORETTI, A. C. Comparing algorithms for solving sparse nonlinear systems of equations, *SIAM Journal on Scientific and Statistical Computing* v. 13, pp. 459-483, 1992.
- [29] GONDZIO, J., SARKISSIAN, R. Column generation with a primal-dual method, *Technical Report 96.6*, Logilab, HEC Geneva, University of Geneva, Suíça, 1996.
- [30] GOODMAN D., MANDAYAN, N. Power Control for Wireless Data, *IEEE Personal Communications Magazine* v. 7, n.2, pp. 48-54, 2000.
- [31] GOODMAN D., MANDAYAN, N. Network Assisted Power Control for Wireless Data, *Mobile Networks and Applications* v. 6, n.5, pp. 409-415, 2001.
- [32] GRIFFITH, R. E., STEWART, R. A. A nonlinear Programming Technique for the Optimization of Continuous Processing Systems, *Management Science* v. 2, pp. 379-392, 1961.

- [33] GUIGNARD, M., KIM, S. Lagrangean decomposition for integer programming: a model yielding stronger Lagrangean bounds, *Mathematical Programming* v. 32, pp. 215-228, 1987.
- [34] HEILBRONER, R. L., *Elementos de Macroeconomia*, 4a ed., Rio de Janeiro, Zahar Editores, 1978.
- [35] HINDI, K. S., BASTA, T. Computationally efficient solution of a multiproduct, two-stage distribution-location problem, *The Journal of the Operational Research Society* v. 45, pp. 1316-1323, 1994.
- [36] JEFFERSON, T., *Manual for the geometric programming code GPROG (CDC) version 2*, Repotr No. 1974/OR/2, Mechanical and Industrial Engineering Department, University of New South Wales, Australia, 1974.
- [37] KAMATH, A. P., KARMARKAR, N. K., RAMAKRISHNAN, K. G., RESENDE, M. G. C. Computational experience with an interior point algorithm on the Satisfiability problem, In: *Ann. Oper. Res.*, v. 25, pp. 43-58, 1990.
- [38] KARMARKAR, N. An interior point algorithm to NP-complete problems - Part I, *Contemp. Math.*, v. 114, pp. 297-308, 1990.
- [39] KARMARKAR, N., RESENDE, M. G. C., RAMAKRISHNAN, K. G. An interior point algorithm to solve computationally difficult set covering problems, *textitMath. Programming* v. 52, pp. 597-618, 1991.
- [40] KLOSE, A. A lagrangean relax-and-cut approach for the two-stage capacited facility location problem, *European Journal of Operational Research*, v. 126, pp. 408-421, 2000.

- [41] KOCHENBERGER, G. A. Geometric Programming - Extensions to Real with Degrees of Difficulty and Loose Constraints, PhD Thesis, University of Colorado, 1969.
- [42] KORTANECK, K. O., NO, H. A second order affine scaling algorithm for geometric programming dual with logarithmic barrier, *Optimization* v. 23, pp. 303-322, 1992.
- [43] KORTANECK, K. O., XU, X., YE, Y. An infeasible interior point method for solving primal and dual geometric programs, *Mathematical Programming* v. 76, pp. 155-181, 1995.
- [44] KOSKOSIDIS, Y. A. & DUARTE, A. M. Jr. A Scenario-Based Approach to Active Asset Allocation, *Journal of Portfolio Management* v.23, pp.74-85, 1997.
- [45] LEE, W. C. Y., *Mobile Celular Telecommunications: Analog and Digital Systems*, New York, Mc Graw Hill, 1995.
- [46] LEE, W. C. Y. Overview of cellular CDMA, *IEEE Trans. Veh. Technol.* v.40, pp.291-302, 1991.
- [47] LIMA, E. L., *Espaços Métricos, Rio de Janeiro*, Livros Técnicos e Científicos Editora S.A., 1993.
- [48] LINTNER, J. The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets, *Review of Economics and Statistics*, 1965.
- [49] LUCENA, A. Steiner problem in graphs: lagrangean relaxation and cutting-planes, *COAL Bulletin, Mathematical Programming Society* v. 21, pp. 2-8, 1992.

- [50] MACAMBIRA, E. M., *Modelos e algoritmos de programação inteira em projetos de redes de telecomunicação* Tese de D.Sc., COPPE/UFRJ, Rio de Janeiro, RJ, Brasil, 2003.
- [51] MARTINSON, R. K., TIND, J. An interior point method in Dantzig-Wolfe decomposition, *Computers & Operations Research*, v. 26, pp. 1195-1216, 1999.
- [52] MATEUS, G. R., LOUREIRO, A. A. F. Introdução à Computação Móvel, *11^o Escola de Computação*, Rio de Janeiro, 1998.
- [53] HANSELMAN, D., LITTLEFIELD, B., *MATLAB 5 Guia do Usuário*, Makron Books, 1999.
- [54] MICHELON, P., MACULAN, N. Lagrangean methods for 0-1 quadratic problems, *Discrete Applied Mathematics* v. 42, pp. 257-269, 1993.
- [55] MITCHELL, J., PARDALOS, P., RESENDE, M. Interior Point methods for Combinatorial Optimization, in *Tamas Terlaky, editor, Interior Point Method in Mathematical Programming*, chapter 11, pp. 417-466, Kluwer Academic Publishers, 1996.
- [56] MONTEIRO, R. D. C., PANG, J. -S. A Potencial Reduction Newton Method for Constrained Equations, *SIAM Journal on Optimization* v. 9, n. 3, pp. 729-754, 1999.
- [57] MONTEIRO, S., *Um método de pontos interiores usando centros analíticos para resolver problemas de programação linear de grande porte*. Dissertação de M.Sc., COPPE/UFRJ, Rio de Janeiro, RJ, Brasil, 2002.

- [58] MONTORO, A. F. FILHO et al.; organizadores PINHO, D. B. E VASCONCELOS, M. A. S., *Manual de Economia*, 3a ed., São Paulo, Editora Saraiva, 1998.
- [59] MOTTA, R. R. & CALÔBA, G. M., *Análise de investimentos: tomada de decisão em projetos industriais*, São Paulo, Editora Atlas S.A., 2002.
- [60] NAGIH, A., PLATEAU, G. Dualité lagrangienne em programmation fractionnaire concave-convexe en variables 0-1, *C. R. Acad. Sci. Paris* t. 331, Série I, pp. 491-496, 2000.
- [61] NAGIH, A., PLATEAU, G. A lagrangean decomposition for 0-1 hyperbolic programming problems, *Int. J. Math. Algor.* v. 1, pp. 299-314, 2000.
- [62] NASCIMENTO, R. Q., *Métodos de Pontos Interiores para Programação Geométrica*. Tese de D.Sc., COPPE/UFRJ, Rio de Janeiro, RJ, Brasil, 1996.
- [63] NASCIMENTO, R. Q., ANDRADE FILHO, M. G. Otimização de um sistema hidrotérmico via programação geométrica, In: *Anais do XXXIII SBPO*, Campos do Jordão, Novembro 6-9, 2001.
- [64] NASH, J. F. Non-cooperative games, *Annals of Mathematics*, v. 54, pp. 289-295, 1951.
- [65] NESTEROV, Y. Complexity estimates of some cutting plane methods based on analytic barrier, *Mathematical Programming* v.69, pp.149-176, 1995.
- [66] OLIVEIRA, P. R., SANTOS, M. A. Using Analytic Center and Cutting Planes Methods for Nonsmooth Convex Programming, *Lectures Notes in Economics and Mathematical Sciences*, Alemanha, v. 481, pp. 339-356, 2000.

- [67] PASCUAL, L. e BEN-ISRAEL, A. Constrained maximization of posynomials by geometric programming, *Journal of Optimization Theory and Applications* v. 5, pp.73-86, 1970.
- [68] PASSY, U. Generalized weighted mean programming, *SIAM Journal on Applied Mathematics* v.15, n.5, pp.1344-1356, 1967.
- [69] PASSY, U. e WILDE, D. J. Generalized Polynomial Optimization *SIAM Appl. Math* v.15, n.5, pp.1344-1356, 1967.
- [70] PETERSON, E. L. Geometric Programming and some of its extensions in Optimizing and Design, *M. Avriel, M. J. Rijckaert and D. J. Wilde, eds* Prentice-Hall, Englewoods Cliffs, NJ, 1973.
- [71] REKLATTIS, G. V. e WILDE, D. J. Geometric programming via primal auxiliary programs, *ORSA 38th National Meeting*, Dallas, Texas, 1971.
- [72] RIBEIRO, C. C., *Algorithmes de Recherche de Plus Courts Chemins avec δ Contraintes: tude Thorique, Implmentation et Parallisation*, Thse de Docteur-Ingénieur, cole Nationale Suprieure des Tlcommunications, Paris, Frnça, 1983.
- [73] RIBEIRO, C. C., MINOUX, M. Solving Hard Constrained Shortest Path Problems by Lagrangean Relaxation and Branch-and-Bound Algorithms, *Methods of δ Operations Research* v.53, pp.305-316, 1986.
- [74] RIJCKAERT, M. J. e MARTENS, X. M. Comparison of generalized geometric programming algorithm, *Journal of Optimization Theory and Applications*, v. 26(2), pp. 205-242, 1978.

- [75] ROSS, S. A., The arbitrage theory of capital asset pricing, *Journal of Economic Theory*, 1976.
- [76] ROSS, S. A., WESTERFIELD, R. W, JAFFE, J. F., *Administração Financeira - Corporate Finance*, São Paulo, Editora Atlas S.A., 2002.
- [77] ROUSSAS, G. G., *A course in mathematical statistics*. 2 ed., San Diego, Academic Press, 1997.
- [78] SAMPATH, A., KUMAR, P. S., HOLTZMAN, J. M. Power Control and resource management for a Multimedia CDMA Wireless System.
- [79] SAMPATH, A., *Integrated Voice/Data CDMA Wireless Systems: Capacity, Access Control and Performance Analysis*. Ph.D. Thesis, Electrical and Computer Engineering Department, Rutgers University, EUA, 1997.
- [80] SANDRONI, P., *Novo Dicionário de Economia*, São Paulo, Editora Best Seller, 1994.
- [81] SANTOS, M. A., OLIVEIRA, P. R. Interior point algorithm for Dantzig and Wolfe Decomposition Principle. Nonconvex Optimization and its Applications, *Series Advances in Convex Analysis and Global Optimization*, Dordrecht/Boston/London, v. 54, pp.473-486, 2001.
- [82] SAPATNEKAR, S. S., RAO, V. B., VAIDYA, P. M. A convex optimization approach to transmissor sizing for CMOS circuits, Proceedings of the IEEE International Conference on Computer Aided Design, *IEEE Computer Society Press*, Los Alamitos, CA, pp.482-485, 1991.

- [83] SARAYDAR, C. U., MANDAYAN, N. B., GOODMAN, D. J. Pricing and Power Control in a Multicell Wireless Data network, *IEEE Journal on Select Areas in Communications, Special Issue: Mobility and Resource Management in NG Wireless Systems*, a ser publicada, 2001.
- [84] SARAYDAR, C. U., MANDAYAN, N. B., GOODMAN, D. J. Efficient power control via pricing in wireless data networks, *IEEE Trans. Comm.*, aceito, 2000.
- [85] SHARPE, W. F. Capital asset prices: a theory of market equilibrium under conditions of risk, *Journal of Finance*, 1964.
- [86] SHI, C. J., VANNELI, J., VLACH, J. An improvement on Karmarkar's algorithm for integer programming, *COAL Bull* v. 21, pp. 23-29, 1992.
- [87] SONNEVEND, G. An analytical center for polyhedrons and new class of global algorithms for linear (smooth convex) programming, *Lectures Notes in Control and Information Sciences* v.84, pp.866-876, Springer-Verlag, New York, 1985.
- [88] STAATS, G. E. Computational Aspects of Geometric Programming with Degrees of Difficulty, Unpublished Doctoral Dissertation, The University of Texas at Austin, Austin, Texas, 1970.
- [89] STRUZAK, R. Frequency Reuse and Power Control in Wireless Networks, *Global Communications*.
- [90] TAHA, H. A., *Operations Research: an introduction*. 5 ed., Mac Millan, 1992.
- [91] TREYNOR, J. Toward a theory of the market value of risky assets. Trabalho não publicado, 1961.

- [92] WALDMAN, H., YACOUB, M. D., *Telecomunicações: Princípios e Tendências*, São Paulo, Editora Érica Ltda, 1997.
- [93] WARNERS, J. P., TERLAKY, T., ROOS, C., JANSEN, B. Potencial reduction algorithms for structured combinatorial optimization problems, *Operations Research* v. 21, pp. 55-64, 1997.
- [94] WENTGES, P. Weighted Dantzig-Wolfe Decomposition for Linear Mixed-integer Programming, *International Transactions in Operational Research* v. 4, pp. 151-162, 1997.
- [95] WILDE, D. J. e BEIGHTLER, C. S. *Foundation of Optimization* Prentice-Hall, Inc Englewood Cliffs, NJ, 1967.
- [96] *XPRESS MP User Guide*, Dash Associates, UK, 1991.
- [97] YANG, H. H., BRICKER, D. L. Investigation of path-following algorithms for signomial geometric programming, *European Journal of Operational Research* v. 103, pp.230-241, 1997.
- [98] YUN, L. C., MESSERSCHIMITT, D. G. Power Control for Variable QoS on a CDMA Channel, In: *Proc. IEEE MILCOM*, pp. 178-182, Fort Monmouth, N. J. Oct. 2-5, 1994.
- [99] YUN, L. C., MESSERSCHIMITT, D. G. Variable Quality of Service in CDMA systems by Statistical Power Control, In: *IEEE International Conference on Communications*, Seattle, June 18-21, 1995.
- [100] ZANDER, J. Performance of Optimun Transmitter Power Control in Cellular Radio Systems, *IEEE Trans. on Vehicular Technology* v. 41, n.1, 1992.

[101] ZENTGRAF, R., *Estatística Objetiva*, Rio de Janeiro, ZTG Editora Ltda., 2001.