

OTIMIZAÇÃO DE PORTFÓLIO DE ATIVOS REAIS UTILIZANDO  
UMA MEDIDA DE RISCO COERENTE

por

**Sergio Vitor de Barros Bruno**

Dissertação de Mestrado submetida ao  
Instituto Nacional de Matemática Pura e Aplicada  
como requerimento para a obtenção do grau de

MESTRE EM MATEMÁTICA

Métodos Matemáticos em Finanças  
Instituto Nacional de Matemática Pura e Aplicada– IMPA

Rio de Janeiro, Março de 2008

**Sergio Vitor de Barros Bruno**

**Otimização de Portfólio de Ativos Reais Utilizando uma  
Medida de Risco Coerente**

Dissertação apresentada como requisito parcial para obtenção do grau de Mestre em Matemática pelo Programa de Pós-graduação em Métodos Matemáticos em Finanças do IMPA. Aprovada pela Comissão Examinadora abaixo assinada.

**Dr. Jorge Passamani Zubelli**

Orientador

IMPA – Instituto Nacional de Matemática Pura e Aplicada

**Dra. Claudia Alejandra Sagastizábal**

Co-orientadora

CEPEL – Centro de Pesquisas de Energia Elétrica

**Dr. Mikhail Solodov**

IMPA– Instituto Nacional de Matemática Pura e Aplicada

**Dr. Fernando Antonio Lucena Aiube**

PUC-Rio – Pontifícia Universidade Católica do Rio de Janeiro

Rio de Janeiro, 27 de Março de 2008

Todos os direitos reservados. É proibida a reprodução total ou parcial do trabalho sem autorização do instituto, do autor e do orientador. Os *softwares* utilizados na preparação deste trabalho não violam direitos. São eles: editor L<sup>A</sup>T<sub>E</sub>X LEd, disponível em <http://www.latexeditor.org>, compilador MiKTeX, disponível em <http://miktex.org>, e editor de imagens xfig. Os resultados foram gerados utilizando-se o *software* MatLab 2007, fruto do acordo entre o IMPA e a Opencadd.

**Sergio Vitor de Barros Bruno**

Ficha Catalográfica

Bruno, Sergio Vitor de Barros

Otimização de Portfólio de Ativos Reais Utilizando uma Medida de Risco Coerente / Sergio Vitor de Barros Bruno; orientador: Jorge Passamani Zubelli; co-orientador: Claudia Alejandra Sagastizábal. — Rio de Janeiro : IMPA, Matemática, 2008.

v., 61 f: il. ; 29,7 cm

1. Dissertação (mestrado) - Instituto Nacional de Matemática Pura e Aplicada, Matemática.

Inclui referências bibliográficas.

1. Matemática – Tese. 2. Otimização de Portfólio. 3. Ativos reais. 4. Medidas de Risco. 5. CVaR. 6. . I. Zubelli, Jorge Passamani. II. Sagastizábal, Claudia Alejandra. III. Instituto Nacional de Matemática Pura e Aplicada. Matemática. IV. Título.

A Luisa, meu grande amor.

## Agradecimentos

Agradeço primeiramente ao Professor Weiler Finamore, por ter me apresentado a pesquisa acadêmica e a Rui Mano, pelos primeiros ensinamentos como engenheiro. A colaboração de ambos foi fundamental para minha formação profissional e acadêmica.

Agradeço também a minha família e minha namorada Luisa, pois me deram todo o apoio que precisei durante a realização desse mestrado e muita compreensão em meus momentos de ausência causados pelos estudos.

A Petrobras apoiou diretamente a realização desse curso através de seu programa de treinamento. Agradeço muito a essa empresa e a Roberto Iachan por viabilizarem meus estudos. Aos amigos e colegas de trabalho, Luiz Carlos e Pierre, agradeço as inúmeras conversas que por diversas vezes resultaram em excelentes sugestões para os problemas encontrados no decorrer desse trabalho.

Por fim, agradeço ao Professor Jorge Zubelli pelos dois anos de ensinamentos que moldaram meu raciocínio matemático e conhecimentos de finanças e à Professora Claudia Sagastizábal pela excelente orientação, sempre presente, da qual obtive diversos aprendizados teóricos e práticos.

## Resumo

Bruno, Sergio Vitor de Barros; Zubelli, Jorge Passamani; Sagastizábal, Claudia Alejandra. **Otimização de Portfólio de Ativos Reais Utilizando uma Medida de Risco Coerente**. Rio de Janeiro, 2008. 61p. Dissertação de Mestrado — Instituto Nacional de Matemática Pura e Aplicada.

A teoria de portfólio moderna baseia-se no paradigma risco-retorno popularizado por Markowitz [17]. Muitas das críticas existentes sobre o modelo de Markowitz referem-se à penalização da volatilidade positiva dos retornos e ao desempenho fraco do modelo, quando utilizado com dados reais. Em testes empíricos, o modelo média-variância apresenta alta sensibilidade à estimação de seus parâmetros e concentração de alocações. No final da década de noventa a pesquisa sobre o assunto ganhou novas forças, com a introdução de outras medidas de risco, em particular aquelas chamadas *coerentes*, que apresentam boas propriedades do ponto de vista da otimização. Este trabalho estuda o problema de otimizar o retorno esperado de um portfólio de investimentos em ativos reais de uma empresa que atua no setor de Gás e Energia. O retorno é dado pelo *Valor Presente Líquido* do lucro líquido da empresa, resultante de sua operação em longo prazo e considerando tanto seu resultado operacional quanto a aquisição dos ativos do portfólio. O risco é controlado mediante restrições ao risco máximo aceitável, medido com a função *Conditional Value at Risk*. A abordagem leva em consideração diferentes cenários de demanda que possuem probabilidades de ocorrência conhecidas. Contrariamente a versões determinísticas, o modelo considerado, de natureza estocástica, evita o conhecido problema de *flaw of averages*, que é o erro obtido em um modelo ao representar uma quantidade incerta por sua média. O programa linear estocástico resultante de nossa abordagem é de grande porte e não é solúvel em tempo computacional razoável, quando aplicados métodos diretos de resolução. Para viabilizar a otimização do modelo estocástico, são aplicadas técnicas de decomposição, que reduzem o esforço computacional e possibilitam a paralelização do processo de resolução do problema. Os resultados numéricos obtidos confirmam as vantagens de utilizar métodos de decomposição para resolver este tipo de problema.

## Palavras-chave

Otimização de Portfólio. Ativos reais. Medidas de Risco. CVaR. .

## Abstract

Bruno, Sergio Vitor de Barros; Zubelli, Jorge Passamani; Sagastizábal, Claudia Alejandra. **Optimization of Real Assets portfolio using a coherent risk measure**. Rio de Janeiro, 2008. 61p.  
MsC Thesis — Instituto Nacional de Matemática Pura e Aplicada.

Modern Portfolio Theory relies on the risk-return paradigm of Markowitz [17]. The original Markowitz model is often subject to criticisms due to the penalization of both negative and positive volatility and the weak performance of the model, when applied to empirical tests. Indeed, the mean-variance model is highly sensitive to estimation errors and often leads to allocation on few assets. By the end of the 90s different measures of risk were introduced, in particular *Coherent measures*, with interesting properties from an optimization perspective. This work studies the problem of optimizing the expected return of the real assets portfolio of a Gas & Energy company. Return is measured by Net Present Value of the company's Net Profit, considering both the long term operational result and the amount of investment capital. Risk is controlled by constraining the maximum acceptable risk, measured with the *Conditional Value at Risk*. Our approach takes into account different demand scenarios with known associated probabilities. Contrary to deterministic counterparts, our stochastic model avoids the well known *flaw of averages*, the error caused by representing an uncertain quantity by its average. The resulting stochastic linear programming problem is large scale and cannot be solved in reasonable times with commercial solvers. For this reason, to solve the problem, we propose decomposition techniques that reduce the overall computational effort and open the way to parallel implementations. Our numerical results confirm the advantages of employing decomposition methods to solve this type of problems.

## Keywords

Portfolio Optimization. Real Assets. Risk Measures. CVaR. .

# Sumário

1	Introdução	<b>11</b>
1.1	Descrição do Problema	11
1.2	Resolução do Problema	15
1.3	Organização e Contribuições do Trabalho	15
2	Modelagem das Incertezas	<b>17</b>
3	Medidas de Risco	<b>22</b>
3.1	Teoria de Portfólio	22
3.2	Algumas Medidas de Risco	24
3.3	Propriedades de Medidas de Risco	25
4	O Problema de Portfólio Estocástico	<b>31</b>
4.1	Formulação Matemática do Problema	31
4.2	Decomposição do Modelo Proposto	33
4.2.1	Decomposição por Planos Cortantes	34
4.2.2	Decomposição por Relaxação Lagrangeana	37
5	Resultados	<b>39</b>
5.1	Custo computacional do modelo Determinístico Equivalente	39
5.2	Validação dos métodos de decomposição	40
5.3	Planos cortantes: Variantes	44
5.4	Resultados obtidos	44
5.5	Fronteira Eficiente	46
5.6	Sensibilidade do nível de confiança	47
5.7	Condicionamento do problema	48
6	Conclusão	<b>50</b>
6.1	Conclusões	50
6.2	Extensões	51
A	Apêndice	<b>56</b>
A.1	Otimização Não Diferenciável	56
A.2	Aproximação por planos cortantes	56
A.3	Decomposição de Benders	57
A.4	Relaxação Lagrangeana	60

## Lista de figuras

1.1	Malha de Gás Natural existente e em construção no Brasil	13
1.2	Representação esquemática da rede de gás	14
1.3	Representação de possibilidade de investimento	14
2.1	Estrutura de cenários em “pente”	18
3.1	Fronteira eficiente de Markowitz	23
3.2	VaR do retorno de um portfólio	23
3.3	CVaR do retorno de um portfólio	25
5.1	Tempo de otimização para determinístico equivalente, por número de cenários	41
5.2	Comparação do nível de alocação proposto pelos modelos estocástico e determinístico	45
5.3	Fronteira Eficiente	47
A.1	Exemplo de um “kink”	57
A.2	Aproximação da função $f(x)$ por planos cortantes	58

## Lista de tabelas

5.1	Tempo de otimização para determinístico equivalente, por número de cenários	40
5.2	Desempenho do determinístico equivalente com restrição de CVaR desativada	42
5.3	Comparação de desempenho dos métodos de decomposição com restrição de CVaR desativada	42
5.4	Desempenho do determinístico equivalente com restrição de CVaR ativada	43
5.5	Comparação de desempenho dos métodos de decomposição com restrição de CVaR ativada	43
5.6	Comparação do nível de alocação proposto pelos modelos estocástico e determinístico.	46
5.7	Teste de sensibilidade do nível de confiança do CVaR	48

# 1 Introdução

Este capítulo introdutório apresenta o problema abordado, as técnicas atuais de resolução, e as contribuições propostas neste trabalho.

## 1.1 Descrição do Problema

O presente trabalho estuda o caso de uma empresa de energia, que possui uma rede de gasodutos e que é responsável pela aquisição, venda e transporte de gás. O caso de interesse é o brasileiro, cuja estrutura de distribuição de Gás Natural possui características bem específicas.

O maior agente da indústria de energia no Brasil é a empresa Petrobras. Seus ativos compreendem campos de produção de óleo e gás, refinarias, usinas termelétricas, unidades de processamento de gás natural (UPGN), além de diversos oleodutos, gasodutos, terminais, embarcações, e outros <sup>1</sup>.

A oferta de gás no Brasil é composta majoritariamente pelo gás produzido nacionalmente pela Petrobras e seus parceiros (64%), e em menor proporção pelo gás importado pela Petrobras de outros países (36%). Esse gás importado atualmente é todo proveniente da Bolívia, através do Gasoduto Brasil-Bolívia, mas já está planejada a entrada em operação de terminais no Rio de Janeiro e no Nordeste para a regaseificação de gás liquefeito. Estes terminais permitirão a aquisição de gás importado de forma mais flexível, através da aquisição de cargas de *gás natural liquefeito* (GNL) da África, por exemplo.

A demanda de gás é composta por três grandes grupos. A demanda das refinarias da Petrobras por combustível e insumos em suas unidades (30%), a demanda das *companhias distribuidoras locais* (CDL), que atendem aos consumidores industriais, residenciais, *gás natural veicular*

<sup>1</sup>mais informações podem ser obtidas no site da empresa, <http://www.petrobras.com.br>

(GNV) e outros (51%), e a demanda das Usinas Termoelétricas, que utilizam o gás como combustível para gerar energia elétrica (19%)<sup>2</sup>.

A demanda de gás natural foi fomentada por políticas nacionais a partir da década de noventa. A participação do gás na matriz energética brasileira subiu de 5,4% em 2000 para 9,3% em 2005, segundo [20]. Isso implicou em diversos investimentos na ampliação da malha de gasodutos, como pode ser visto na Figura 1.1.

Para operar eficientemente a malha de gasodutos, devem ser levados em consideração diversos fatores, como por exemplo, a qualidade do gás transportado. O gás natural chamado *especificado* possui um poder calorífico de 9.400 kcal/m<sup>3</sup> e é composto basicamente por C1 e C2<sup>3</sup>. Porém, quando o gás não está especificado, pode ser encontrado misturado com hidrocarbonetos mais pesados, como C3 e C4, que formam o GLP (Gás Liquefeito de Petróleo, conhecido como gás de cozinha). A mistura eleva o poder calorífico do gás, podendo ser encontrado a até 10.500 kcal/m<sup>3</sup>. Ainda, é possível que o gás esteja misturado a elementos inertes, o que reduz o poder calorífico abaixo do nível do especificado. Sendo assim, devem ser considerados diferentes fornecedores de gás e diversas qualidades do mesmo. O gás produzido nessas fontes só pode ser entregue a clientes que aceitem o produto não especificado, ou então entregue a uma UPGN. A UPGN faz a especificação do gás, processo pelo qual os hidrocarbonetos são separados, produzindo gás especificado e GLP.

Uma representação esquemática da malha operada pela empresa Petrobras pode ser vista na Figura 1.2. Uma das principais decisões de planejamento de longo prazo da empresa é relativa a determinar seu portfólio de investimentos. Para atender à demanda de longo prazo, é esperado que a empresa realize investimentos de ampliação de sua malha de gasodutos, para poder aumentar a capacidade de transporte do gás. A Figura 1.3 representa essa oportunidade de investimento.

A decisão de novos investimentos é difícil pela complexidade do problema, no qual existem vários fornecedores de gás e clientes disputando os mesmos recursos e também porque a inclusão de um investimento na

<sup>2</sup>Fonte: Balanço Energético Nacional, Ministério de Minas e Energia

<sup>3</sup>Na indústria de petróleo é usual se referir a um hidrocarboneto pelo seu número de átomos de carbono. Assim sendo, o metano ( $C_1H_4$ ) é referido por C1, o etano ( $C_2H_6$ ) e eteno ( $C_2H_4$ ) são referidos por C2, etc.



Figura 1.1: Malha de Gás Natural existente e em construção no Brasil

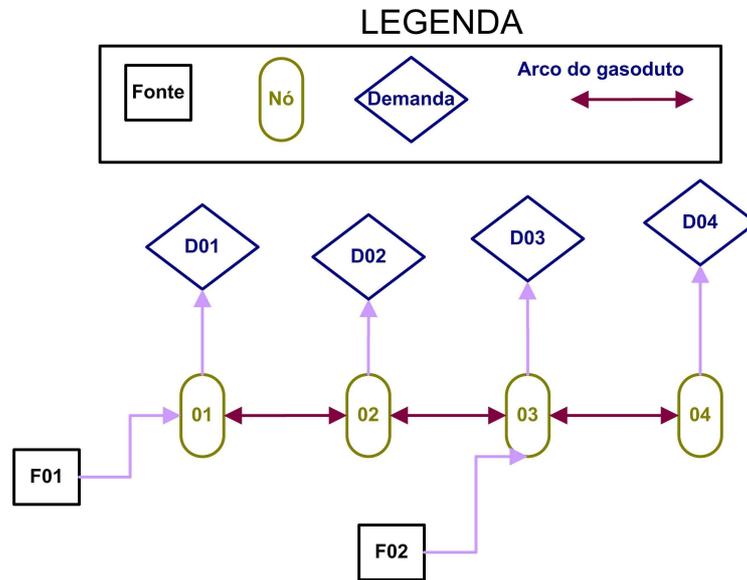


Figura 1.2: Representação esquemática da rede de gás

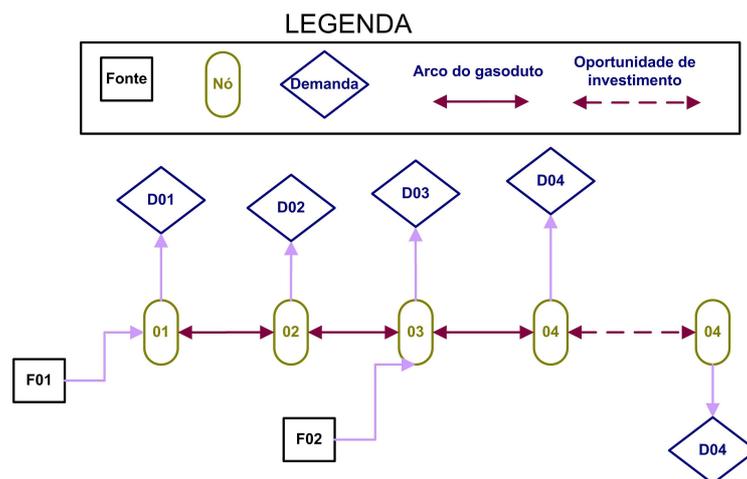


Figura 1.3: Representação de possibilidade de investimento

carteira afeta diretamente a viabilidade econômica de outro projeto. Por outro lado, a demanda de longo prazo é um dado de natureza incerta.

## 1.2 Resolução do Problema

O objetivo deste trabalho é desenvolver um modelo para apoiar a escolha de um portfólio de investimento em ativos reais, isto é, investimento em ativos físicos que apresentam algum grau de irreversibilidade, atendendo a certos critérios de otimalidade e restrições operacionais e financeiras. Os ativos que compõem o portfólio são aqueles tipicamente considerados pela empresa, como gasodutos, UPGNs, terminais de GNL etc. Na Petrobras já existem modelos disponíveis que fazem o tratamento desse problema, porém os mesmos são apenas modelos determinísticos, que não levam em conta as incertezas inerentes a planos de longo prazo. Conforme será mostrado neste trabalho, existem potenciais ganhos de se utilizar modelos estocásticos na resolução desse problema, além da possibilidade de gerenciar melhor diversos riscos financeiros inerentes à decisão de investimento.

## 1.3 Organização e Contribuições do Trabalho

As principais contribuições do presente trabalho são a substituição de um modelo determinístico por um modelo estocástico com restrições de risco, e a proposição de métodos de decomposição para viabilizar a otimização desse modelo.

As vantagens do uso de um modelo estocástico ao invés de um modelo determinístico são descritas no Capítulo 2.

O modelo estocástico pode levar a resultados indesejados em determinados cenários. Para proteger o investidor, o modelo incorpora uma medida de risco coerente, o Conditional Value at Risk (CVaR) para definir níveis máximos aceitáveis de risco do resultado dos portfólios selecionados. As medidas de risco, suas classes e propriedades são descritas em detalhes no Capítulo 3.

O programa linear estocástico proposto como modelo é de grande porte e não é solúvel em tempo computacional aceitável quando são aplicados métodos diretos de resolução. Assim sendo, são aplicadas técnicas de decomposição que reduzem o esforço computacional e possibilitam a

paralelização do processo de resolução do problema. O modelo proposto e as técnicas de decomposição utilizadas são apresentadas no Capítulo 4.

Os resultados obtidos com a aplicação do modelo proposto e o desempenho computacional dos métodos de decomposição utilizados são descritos no Capítulo 5.

Por fim, as conclusões e as possíveis extensões do trabalho são apresentadas no Capítulo 6.

## 2 Modelagem das Incertezas

Por se tratar de um problema de longo prazo (5 anos ou mais), existem diversas fontes de incerteza que devem ser observadas para modelar o problema com boa acurácia.

A natureza operacional de algumas premissas do problema envolve incertezas de cunho técnico, tais como a qualidade do gás recebido de uma fonte, volume disponível de oferta e volume de demanda. Também existem incertezas político-econômicas, como preços de compra e venda, impostos, etc. A demanda termoelétrica é uma das fontes de incerteza mais claras do problema e é a considerada nesse trabalho para se definir uma árvore de cenários para o modelo estocástico, devido a sua relevância e variabilidade.

O planejamento e operação do setor elétrico Brasileiro é fortemente influenciado pelo regime de chuvas. Devido à grande incerteza associada às chuvas, é comum o uso de modelos estocásticos de planejamento de longo prazo. O Operador Nacional do Sistema (ONS), responsável pela operação do sistema elétrico brasileiro, determina o nível de despacho mensal das diversas usinas em operação usando o modelo NEWAVE [14]. Este modelo determina os despachos mensais das usinas brasileiras com base no menor custo esperado total de operação e déficit do sistema, considerando centenas de séries sintéticas de chuvas, produzidas por um simulador.

Para gerar suas previsões de demanda termoelétrica, a Petrobras faz uso do modelo NEWAVE e do SDDP<sup>©</sup> (sigla para *Stochastic Dual Dynamic Programming*), software comercial que também implementa métodos de amostragem de programação estocástica para resolver o problema. Tipicamente, o modelo SDDP considera de 80 a 200 séries de chuva para gerar cenários de demanda termoelétrica.

Já foi mencionado que a decisão de investimento em ativos reais é de longo prazo, havendo casos de ativos que se tornam disponíveis só anos

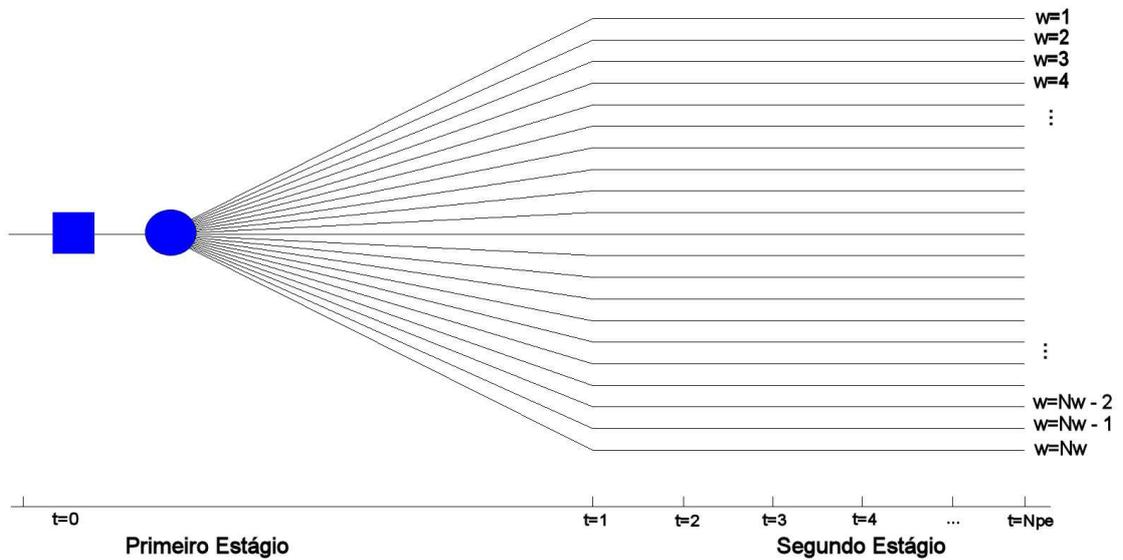


Figura 2.1: Estrutura de cenários em “pente”

depois da decisão de investimento, como no caso de gasodutos, e também casos de outros ativos irreversíveis, como contratos de fornecimento com duração de vários anos.

Por esta razão, o decisor não está em condições de definir uma estratégia do tipo *wait and see*, na qual posterga-se a decisão de investimento até que a informação sobre a incerteza é revelada (a demanda de gás para gerar energia elétrica).

Considerando a demanda termoelétrica como única fonte de incerteza, gerada a partir da informação disponibilizada pelo pacote SDDP, temos então uma estrutura de diferentes cenários futuros conforme a Figura 2.1.

A árvore em “pente” da figura mostra que existe um primeiro estágio de decisão, no qual o portfólio de investimentos é determinado. O retângulo representa a decisão de investimento e o círculo a posterior realização de um cenário de demanda. Há também um segundo estágio, composto por vários ( $N_{pe}$ ) períodos subseqüentes de tempo. No segundo estágio, os investimentos previamente decididos entram em operação e são tomadas as melhores decisões operacionais dado o cenário  $w$  (dentre os  $N_w$  possíveis) de demanda termoelétrica realizado e os investimentos

determinados. O modelo em dois estágios permite decidir o plano de investimentos no presente, levando em conta diferentes alternativas futuras, representadas pela árvore de cenários.

Atualmente na Petrobras não há disponibilidade de modelos que utilizem a informação dos cenários segundo descrito acima. Os modelos em uso são determinísticos, e utilizam a média das demandas dos diversos cenários como demanda futura. Essa simplificação permite a resolução do problema de forma determinística com baixo custo computacional, pois é resolvido apenas um problema linear de porte médio, porém há um impacto negativo sobre a qualidade da decisão ótima fornecida pelo modelo.

Em particular, deve-se ter atenção para não incorrer em erros conhecidos como falhas das médias (*flaw of averages*, em inglês). No modelo determinístico utiliza-se a média como aproximação da distribuição probabilística de demanda, representada de forma discreta pelos cenários. Esses cenários podem possuir grande variabilidade, o que pode degradar fortemente a validade do modelo de médias. Um exemplo ilustrativo desse tipo de problema, destacado por Sam Savage em [24], é o chamado de “bêbado na rodovia”.

No exemplo, o bêbado caminha de forma aleatória, ora estando na pista em um sentido, ora estando na pista oposta. Considerando que a pista possui fluxo intenso de veículos e a posição aleatória do bêbado na pista, espera-se que ele seja atropelado. Porém, é possível fazer outra análise em média, e considerar que a *expectativa* da posição do bêbado na pista é de ficar na faixa central, entre as pistas. Como nesse local não há fluxo de veículos, o bêbado não seria atropelado. Matematicamente, o erro desta análise decorre de possuímos em nosso problema uma variável aleatória  $x$  (posição do bêbado na pista) e calcularmos o resultado do problema,  $f$ , sobre  $E[x]$ , a média de  $x$ , obtendo  $f(E[x])$ . A avaliação correta seria calcular a média do resultado  $E[f(x)]$  ( $E[\ ]$  denota o operador de valor esperado).

Outro exemplo de Savage é o da previsão de demanda de um novo produto. Em uma empresa hipotética, o gerente de produto estima que a demanda para o novo produto da empresa estará entre 50.000 e 150.000 unidades. Pressionado pelo chefe para fornecer um único número, o gerente informa que a demanda média é de 100.000 unidades. Por sua

vez, essa estimativa de demanda é utilizada pelo chefe para determinar a capacidade da fábrica e avaliar a rentabilidade do projeto, que é assim calculada em \$ 10 milhões. Ao implementar o projeto, a equipe terá uma grande surpresa: se a demanda for inferior a 100.000 unidades, o resultado será menor que o esperado. Mas se a demanda superior a 100.000 unidades, a fábrica não poderá atender à demanda excedente, ficando limitada a um retorno de \$ 10 milhões. O lado positivo da incerteza não pode ser aproveitado e portanto o projeto terá certamente um retorno esperado abaixo dos \$ 10 milhões calculados pelo chefe.

Levando esse conceito para o modelo de otimização de portfólio, devemos ter cautela quando, por exemplo, um parâmetro  $b$  do problema for uma variável aleatória. Tipicamente a análise determinística utiliza o modelo com a esperança de  $b$ ,  $E[b]$ , fazendo:

$$\begin{aligned} \min_x \quad & f(x) \\ \text{s.a.} \quad & Ax = E[b]. \end{aligned}$$

Outras abordagens para lidar com a incerteza em modelos podem ser vistas em [21].

A metodologia adotada neste trabalho substitui o problema anterior por um problema de *otimização estocástica*, da forma:

$$\begin{aligned} \min_x \quad & E[f(x)] \\ \text{s.a.} \quad & Ax_w = b_w, \quad \text{para todo } w, \end{aligned}$$

onde  $w$  representa um cenário qualquer (ou amostra) do espaço de probabilidade associado à árvore de  $N_w$  cenários.

Será assumido que os vetores mencionados ao longo do texto são vetores coluna, com dimensão dada pelo contexto, a não ser que seja especificado o contrário. O operador  $T$  indica a transposição do vetor ou matriz.

Vamos ilustrar essa mudança de modelagem com mais um exemplo: suponha que nosso problema consista em gerenciar um portfólio de ativos de mercado. O objetivo é maximizar o retorno no fim do período considerado. Os retornos dos  $N$  ativos são dados pelo vetor  $r$ , onde a componente  $r_i$  indica o retorno do  $i$ -ésimo ativo. Esses retornos são variáveis aleatórias com uma distribuição qualquer, de valor esperado  $E[r]$ . Se temos uma limitação de capital  $C$  para investir, o problema poderia ser formulado como:

$$\begin{aligned} \max_x \quad & E[r]^T x \\ \text{s.a.} \quad & \mathbb{1}^T x = C. \end{aligned} \tag{2-1}$$

Nesse problema,  $x$  é o vetor de variáveis de investimento, que indica quanto devemos investir em cada ativo para maximizar o retorno esperado do portfólio. Já  $\mathbb{1}$  denota o vetor com todas as coordenadas iguais a um e dimensão igual a de  $x$ .

Poderíamos substituir o problema (2-1) por:

$$\begin{aligned} \max_x \quad & \sum_{w=1}^{N_w} p_w r_w^T x \\ \text{s.a.} \quad & \mathbb{1}^T x = C, \end{aligned} \tag{2-2}$$

onde  $p_w$  representa a probabilidade de um cenário (ou amostra)  $w$  do espaço de probabilidade definidos pelos  $N_w$  cenários.

Nota-se que nesse exemplo não ocorre o fenômeno da falha das médias. De fato, a solução de ambos os problemas é a mesma, isto é, o mesmo portfólio. Ainda assim, podemos destacar a qualidade da resposta: no problema (2-1), o retorno é  $E[r]^T x = y$ , um único valor. Se existe uma diretriz corporativa de atingir uma meta  $M < y$ , o resultado desse modelo indica a aprovação do portfólio. Já no modelo (2-2), temos  $r_w^T x = y_w$ , para todo  $w$ , que é uma distribuição de probabilidades de retorno. Podemos utilizar essa informação para avaliar o risco de não atingir a meta direta. Esse risco é dado pela soma da probabilidade daqueles cenários  $w$  para os quais  $r_w^T x < M$ .

## 3 Medidas de Risco

Matematicamente, uma medida de risco é um funcional que quantifica o risco de uma decisão. Para contextualizar melhor a discussão sobre medidas de risco, começaremos esse capítulo com uma breve introdução à teoria de portfólio.

### 3.1 Teoria de Portfólio

A teoria de portfólio moderna é fortemente influenciada por Markowitz, que propôs na década de 50 o paradigma de modelagem para o *problema do investidor*. Este paradigma tornou-se padrão nas seis décadas subseqüentes. O modelo proposto por Markowitz [17] é da forma

$$\begin{aligned} \max_{\lambda} \min_x & V[r(x)] - \lambda \{E[r(x)] - \bar{r}\} \\ \text{s.a.} & \mathbb{1}^T x = 1, \end{aligned} \quad (3-1)$$

onde  $x$  é a fração do capital a ser investido no portfólio e  $V[r(x)]$  e  $E[r(x)]$  são a variância e esperança do retorno  $r(x)$ , respectivamente. O parâmetro  $\bar{r}$  é o retorno esperado, dado fornecido pelo decisor e  $\lambda \geq 0$  é uma penalização associada à aversão ao risco do decisor. Uma representação alternativa do modelo é

$$\begin{aligned} \min_x & V[r(x)] \\ \text{s.a.} & \mathbb{1}^T x = 1, \\ & E[r(x)] = \bar{r}. \end{aligned} \quad (3-2)$$

O coeficiente  $\lambda$  em (3-1) está relacionado com o multiplicador ótimo da última restrição de (3-2). Desse modelo de portfólio deriva-se o conceito de fronteira eficiente, que expressa a relação entre o risco (medido pela variância) e o retorno do portfólio. A fronteira eficiente é obtida resolvendo diversas vezes o problema (3-2), para níveis de retorno esperado distintos, conforme apresentado na Figura 3.1.

Algumas das principais críticas ao modelo de Markowitz são decorrentes do uso da variância como medida de risco. De fato, o modelo de

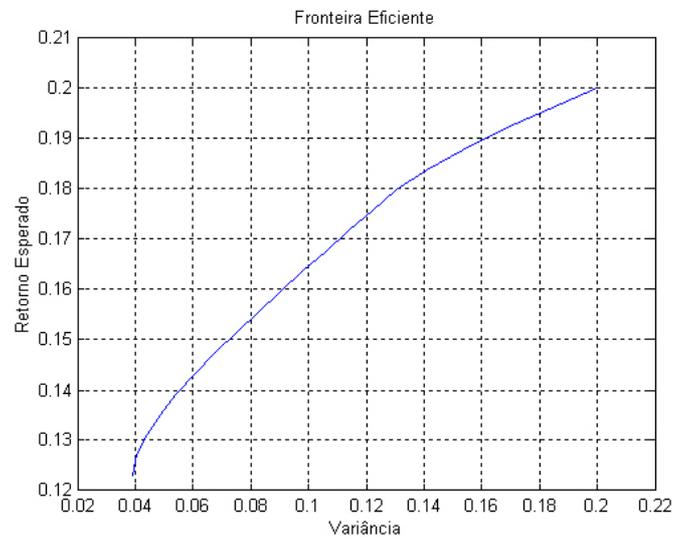


Figura 3.1: Fronteira eficiente de Markowitz

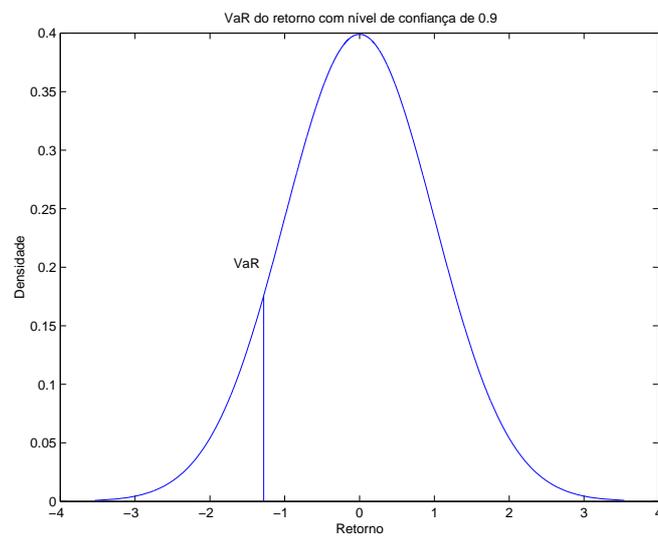


Figura 3.2: VaR do retorno de um portfólio

Markowitz apresenta resultados empíricos de baixa qualidade, ocorrendo pouca diversificação do portfólio e penalização da volatilidade positiva dos retornos dos ativos, assim como alta instabilidade da carteira ao variar os parâmetros do modelo, segundo [12]. Além disso, diversos cuidados devem ser tomados na estimação dos parâmetros e formulação do problema de otimização para não incorrer em erros sérios de análise. Mais informações sobre essas questões de modelagem e formulações por *otimização robusta*, que reduzem os riscos associados ao modelo de Markowitz, podem ser vistos em [19].

O próprio Markowitz reconheceu o problema de se penalizar igualmente a incerteza positiva e negativa das variáveis aleatórias, propondo então um modelo onde a semi-variância é utilizada como medida de risco [16]. Mais recentemente, houveram diversas propostas para o uso do *desvio médio absoluto*, o modelo MAD [11].

Na próxima seção são apresentadas algumas medidas de risco que são de interesse prático e teórico.

### 3.2 Algumas Medidas de Risco

A partir dessa seção será considerado que o retorno estocástico de um portfólio  $\alpha$  é dado pela variável aleatória  $\psi_\alpha(y)$  e que  $\psi_\alpha(y)$  possui função distribuição de probabilidade dada por  $f_{\psi_\alpha}(y)$  e função distribuição acumulada dada por  $F_{\psi_\alpha}(y)$ , onde  $y$  é um nível de retorno. Um quantil do retorno será definido como  $Q_{\psi_\alpha}(p) \equiv F_{\psi_\alpha}^{-1}(p)$ , onde  $p \in [0, 1]$  denota um valor de probabilidade acumulada.

Após o acordo de Basiléia I (1988), popularizou-se o uso de medidas de risco “unilaterais”, baseadas em quantis da distribuição de retornos da carteira, especialmente a medida conhecida como Value at Risk (VaR) [10].

Suponha que o gestor queira monitorar as perdas com um nível de confiança  $\beta$ , onde  $\beta$  é um quantil de  $F_{\psi_\alpha}(y)$ . Podemos definir  $VaR(\beta) \equiv Q_{\psi_\alpha}(\beta)$ . O VaR de um portfólio  $\alpha$  pode ser visto na Figura 3.2. O VaR é freqüentemente definido na literatura como  $E[\psi_\alpha(y)] - Q_{\psi_\alpha}(\beta)$ . Essa translação é uma questão de convenção e não influencia nenhuma das propriedades da medida.

Vários modelos foram propostos, seja para otimizar o VaR da carteira, seja para otimizar o retorno sujeito a uma restrição de máximo

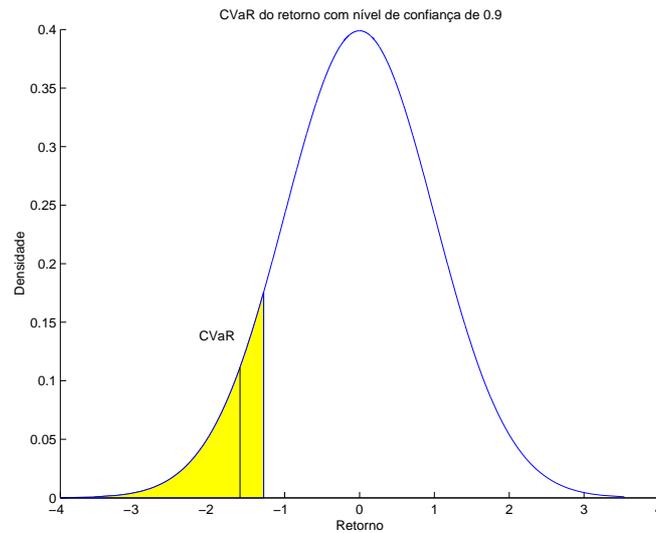


Figura 3.3: CVaR do retorno de um portfólio

VaR aceitável. Esses modelos possuem dificuldades de implementação computacional, pois o emprego do VaR resulta em um problema de otimização não convexa, como veremos a seguir.

Outra medida de grande relevância é o *Conditional Value at Risk* (*CVaR*), ou *Expected Shortfall* (*ES*). O  $CVaR - \beta$  é a média dos quantis dos piores cenários, para um nível de confiança  $\beta$ , com  $\beta \in [0, 1]$ . Matematicamente:

$$CVaR_{\beta}(\alpha) \equiv \frac{1}{1 - \beta} \int_0^{1-\beta} Q_{\psi(\alpha)}(s) ds \quad (3-3)$$

O CVaR de um portfólio  $\alpha$  pode ser visto na Figura 3.3. A relação do CVaR com o VaR é direta: enquanto o VaR delimita a perda dos piores cenários, o CVaR é a média (ou valor esperado) dessa perda. Como consequência direta das definições, para um mesmo nível de confiança, a desigualdade  $CVaR \leq VaR$  é sempre satisfeita.

### 3.3 Propriedades de Medidas de Risco

Como medir o risco pode ser uma tarefa difícil e subjetiva, muitos pesquisadores vêm dedicando esforços para fundamentar uma teoria axiomática de medidas de risco. Também são definidas axiomáticamente diversas propriedades que podem ser desejáveis e classes de medidas que possuem essas propriedades. Alguns exemplos são as medidas de risco coerentes e medidas de risco convexas. Uma boa introdução ao assunto

pode ser vista em [11].

Para fundamentar melhor a escolha de uma medida de risco, primeiro são definidas quais as propriedades desejadas para a medida de risco a ser utilizada no problema e depois são estudadas as diversas medidas ou classes de medidas que atendem às propriedades escolhidas.

O funcional  $R(\alpha)$  expressa, de forma geral, o risco de um portfólio  $\alpha$  determinado por uma medida de risco qualquer.

Vejamos agora uma lista com as principais propriedades reconhecidas como desejáveis para uma medida de risco:

- *Equivalência a Dinheiro*: O risco expressado pela medida é diretamente interpretado em unidade monetária.

Exemplo: o VaR é um exemplo de medida que possui essa propriedade, pois ele informa a quantidade monetária em risco para um determinado nível de confiança.

- *Estimabilidade*: De posse de  $f_{\psi_\alpha}(x)$ ,  $F_{\psi_\alpha}(x)$  ou da função geradora de momentos de  $f_{\psi_\alpha}(x)$ , é possível determinar o risco.

Essa propriedade possui elevada importância prática, pois simplifica o cálculo do risco.

- *Consistência com Dominância Estocástica (fraca)*: O portfólio  $\alpha$  domina estocasticamente (no sentido fraco) o portfólio  $\gamma$  se  $Q_{\psi(\alpha)}(p) \geq Q_{\psi(\gamma)}(p)$ , para todo  $p \in [0, 1]$ . A medida  $R(\alpha)$  é consistente com dominância estocástica se a existência de uma relação de dominância fraca entre dois portfólios  $\alpha$  e  $\gamma$  implica que  $R(\alpha) \leq R(\gamma)$ .

Cabe ressaltar que essa relação é uma dominância em probabilidade, diferente da dominância forte, ou de primeira ordem, que será explanada a seguir.

O VaR naturalmente possui essa propriedade. Se existe dominância estocástica, então  $VaR(\alpha) \leq VaR(\gamma)$ .

- *Sensatez*: A Sensatez (ou Monotonicidade) é consistência com dominância estocástica de primeira ordem (dominância forte). A dominância estocástica de primeira ordem é definida como  $\psi_\alpha(y) \geq \psi_\gamma(y)$ , para todos os cenários. Esta propriedade é mais forte que a dominância fraca. Uma medida é sensata se a existência de uma

relação de dominância forte entre dois portfólios  $\alpha$  e  $\gamma$  implica que  $R(\alpha) \leq R(\gamma)$ .

Essa propriedade também é válida para o VaR. Se ocorre  $\psi_\alpha(y) \geq \psi_\gamma(y)$ , para todos os cenários então  $VaR(\alpha) \leq VaR(\gamma)$ .

- *Constância*: Se o retorno do portfólio  $\alpha$  é determinístico e igual a  $b$ , então  $R(\alpha) = b$ .
- *Homogeneidade Positiva*:  $R(\lambda\alpha) = \lambda R(\alpha)$ , para todo  $\lambda \geq 0$ .

Essa propriedade basicamente implica que a posição tomada em uma carteira tem relação linear com o risco observado.

- *Invariância a Translações*: Somando ao portfólio um retorno certo  $b \equiv 1$ , então  $R(\alpha + \lambda b) = R(\alpha) + \lambda$ , para  $\lambda$  um escalar.
- *Co-monotonicidade Aditiva*: Dado  $(\alpha, \delta)$  co-monotônicos então  $R(\alpha + \delta) = R(\alpha) + R(\delta)$ . Essa propriedade é extremamente importante ao utilizarmos derivativos, pois garante que a medida não é “enganada” por “falsos hedges”.
- *Sub-aditividade*: Uma medida de risco é sub-aditiva se  $R(\alpha) + R(\gamma) \geq R(\alpha + \gamma)$ . Essa propriedade representa o entendimento prático que a diversificação do portfólio leva à redução dos riscos.

Essas propriedades podem ser vistas em mais detalhes, sob a ótica de medidas de satisfação, em [19].

De todas as propriedades acima, a única que o VaR não atende é a de Sub-aditividade. Na prática, isso implica que essa medida de risco não favorece a diversificação do portfólio, como já foi observado por diversos autores [13]. Além disso, em termos de implementação, o VaR não gera uma região viável convexa, conforme exemplificado em [19], o que implica que a otimização de um portfólio utilizando o VaR pode levar a ótimos locais, ao invés de desejáveis ótimos globais.

Em resposta a esses problemas, Artzner et al [2] propuseram uma classe de medidas de risco, baseada em propriedades axiomáticas. Essas medidas são conhecidas como medidas coerentes de risco. As medidas coerentes de risco devem apresentar as propriedades de Sub-aditividade, Homogeneidade Positiva, Sensatez e Invariância a Translações. A medida dessa classe que se popularizou desde então foi o CVaR.

Além de todas as propriedades apresentadas pelo VaR, o CVaR ainda apresenta a Sub-aditividade:  $CVaR(\alpha + \gamma) \leq CVaR(\alpha) + CVaR(\gamma)$ , isto é, a diversificação do portfólio leva à redução do risco.

Outra vantagem do CVaR está em sua implementação: essa medida de risco não só gera uma região viável convexa, permitindo a obtenção de ótimos globais, como também pode ser representada por relações lineares, o que permite a geração de modelos de programação linear, cujo ferramental computacional é muito bem desenvolvido. Essa representação linear é possível ao utilizarmos um número finito de cenários, como no caso de problemas de otimização estocástica.

Iremos agora demonstrar que o CVaR pode ser representado por um conjunto de equações lineares quando o número de cenários é finito. Essa demonstração será feita utilizando uma notação distinta, segundo [13], com o objetivo de maior clareza. Vamos assumir que temos um *vetor de investimentos*  $x$  e um *vetor de incertezas*  $y$  de dimensão  $n$ . A função  $f(x, y)$  é definida como a *perda do portfólio*  $x$ , uma variável aleatória com distribuição em  $\mathbb{R}$  induzida pela distribuição de  $y$ . Para simplificar a demonstração, será assumido sem perda de generalidade que  $y$  possui uma função densidade de probabilidade contínua  $p(y)$ . O  $\beta - VaR$  e  $\beta - CVaR$  serão denotados respectivamente por  $VaR_{\beta(x)}$  e  $\phi_{\beta(x)}$  e serão definidos como

$$VaR_{\beta(x)} = \min\{a \in \mathbb{R} : \Psi(x, a) \geq \beta\}$$

onde

$$\Psi(x, a) = \int_{f(x,y) \leq a} p(y) dy$$

e

$$\phi_{\beta(x)} = \frac{1}{1 - \beta} \int_{f(x,y) \geq VaR_{\beta(x)}} f(x, y) p(y) dy$$

onde  $p(y)$  é a densidade de probabilidade de  $y$ . Dada a função auxiliar  $F_{\beta}(x, a)$ :

$$F_{\beta}(x, a) = a + \frac{1}{1 - \beta} \int_{y \in \mathbb{R}^n} [f(x, y) - a]^+ p(y) dy \quad (3-4)$$

onde  $[f(x)]^+ = f(x)$  quando  $f(x) \geq 0$  e  $[f(x)] = 0$ , caso contrário. Em [22] mostra-se que  $\phi_{\beta(x)} = \min_a F_{\beta}(x, a)$  e  $\min_x \phi_{\beta(x)} = \min_{x,a} F_{\beta}(x, a)$ . Esse resultado permite a formulação de problemas minimizando o CVaR ou minimizando uma função qualquer sujeito a restrições de CVaR, através da representação de  $F_{\beta}(x, a)$ . Essa representação torna-se extremamente

interessante se a distribuição  $f(x, y)$  é discreta ou através de uma amostragem da mesma. Nesses casos, uma aproximação para (3-4) é

$$\tilde{F}_\beta(x, a) = a + \frac{1}{1-\beta} \sum_{j=1}^N [f(x, y) - a]^+ p_j, \quad (3-5)$$

onde  $N$  é o número de cenários (ou amostras) e  $p_j$  a probabilidade associada a cada cenário (ou amostras equiprováveis). Se a função de perdas  $f(x, y)$  é linear com respeito a  $x$ , então a função  $\tilde{F}_\beta(x, a)$  é convexa e linear por partes.

Usando um conjunto de variáveis auxiliares  $z_j$ ,  $j = 1..N$ , então  $\tilde{F}_\beta(x, a)$  pode ser substituída pela função linear  $a + \frac{1}{1-\beta} \sum_{j=1}^N z_j p_j$  e o conjunto de restrições lineares

$$z_j \geq f(x, y_j) - a, \quad z_j \geq 0, \quad j = 1, \dots, N.$$

Assim sendo, um problema de otimização cujo objetivo é maximizar o retorno, que possui uma restrição de CVaR, poderia ser formulado como

$$\begin{aligned} \max_x \quad & E[r(x)] \\ \text{s.a.} \quad & \mathbb{1}^T x = 1, \\ & a + \frac{1}{1-\beta} \sum_{j=1}^N z_j p_j \leq \rho_\beta, \\ & z_j + a + r_w^T x \geq 0 \quad j = 1, \dots, N, \\ & z_j \geq 0, \quad j = 1, \dots, N. \end{aligned} \quad (3-6)$$

Diversos modelos de otimização de portfólio utilizando o CVaR como objetivo ou restrição já foram propostos para aplicações de finanças, como [15] e [25]. Destes, podemos destacar contribuições mais recentes de Fabián em [7] e [8], que propõem algoritmos de decomposição que aproveitam a estrutura do problema. As modelagens propostas são compatíveis com o uso de diversas restrições de CVaR, com limites para diferentes níveis de confiança e intervalos de tempo, permitindo representar de forma detalhada a aversão ao risco do decisor. A indústria de energia elétrica já apresenta várias propostas, como [6] e [18].

O CVaR pertence ainda à classe de *medidas de risco espectrais* [1]. As medidas espectrais estão contidas no conjunto de medidas de risco coerentes. Além das propriedades das medidas coerentes, as medidas espectrais devem possuir Estimabilidade e Co-monotonicidade Aditiva. Esses requisitos implicam que as medidas espectrais possuem todas as propriedades citadas.

A classe de medidas espectrais pode ser representada na forma

$$Esp_{\psi}(\alpha) = \int_0^1 \phi(p) Q_{\psi_{\alpha}}(p) dp$$

onde  $\phi(p)$  é decrescente,  $\phi(1) \equiv 0$ , e  $\int_0^1 \phi(p) dp \equiv 1$ .

Uma medida espectral pode ser vista como uma mudança da medida de probabilidade original para uma medida mais avessa ao risco. A idéia de mudança de medida é um conceito comum em finanças. Um exemplo fundamental surge no estudo de medida martingal equivalente.

## 4 O Problema de Portfólio Estocástico

O modelo apresentado tem como objetivo maximizar o valor esperado do portfólio da empresa relativo às operações de comercialização de gás, no horizonte de cinco ou mais anos, considerando compra e venda de gás, custos de operação da malha, multas contratuais e possibilidades de investimentos. As principais restrições são relativas às ofertas e demandas de gás e a capacidade de fluxo de gás na malha.

Por questão de conveniência, o modelo será representado como um modelo de minimização da perda da empresa. Essa alteração na prática só inverte o sinal dos coeficientes da função objetivo, mas simplifica a representação do problema, pois grande parte da literatura de otimização tem como convenção o estudo de problemas de minimização.

A natureza estocástica do problema encontra-se na demanda de gás, cujas trajetórias de expansão previstas e probabilidades associadas de ocorrência são consideradas conhecidas.

A decisão de investimento tomada no presente independe de qual cenário de demanda e oferta se concretizar no futuro. Esta é uma característica dos modelos estocásticos de dois estágios de decisão: primeiro é decidido o investimento e depois a operação ótima do sistema dimensionado no primeiro estágio. A capacidade do sistema operado no segundo estágio é afetada pela decisão de investimentos através de um conjunto de restrições, como será visto adiante.

Propõe-se o  $\beta - CVaR$  como medida de risco a ser considerada no modelo. Como já foi mencionado, o CVaR é uma medida de risco espectral, possuindo propriedades desejáveis para a otimização de portfólio.

### 4.1 Formulação Matemática do Problema

O problema é caracterizado pelos elementos e notação a seguir. O modelo que representa incerteza na demanda é suposto conhecido. Mais

precisamente, são conhecidos  $N_w$  cenários de demanda  $w$ . Valores típicos de  $N_w$  são oitenta ou duzentos.

O problema a resolver tem a seguinte formulação

$$\min_{z_w, a, x, y_w} \quad c^T x + \sum_{w=1}^{N_w} [p_w q_w^T y_w] \quad (4-1)$$

$$s.a. \quad Ax = b, \quad (4-2)$$

$$a + \frac{1}{1-\beta} \sum_{w=1}^{N_w} p_w z_w \leq \rho \beta, \quad (4-3)$$

$$Tx + Wy_w = h_w, \quad \text{para } w=1, \dots, N_w, \quad (4-4)$$

$$B_w y_w = d_w, \quad \text{para } w=1, \dots, N_w, \quad (4-5)$$

$$z_w + a - c^T x - q_w^T y_w \geq 0, \quad \text{para } w=1, \dots, N_w, \quad (4-6)$$

$$x \geq 0, y_w \geq 0, z_w \geq 0, \quad \text{para } w=1, \dots, N_w, \quad (4-7)$$

onde:

- $x$  representa o conjunto de variáveis de investimento. Existem  $m$  ativos propostos e  $N_{pe}$  períodos no horizonte. Cada variável  $x_{i,j}$  representa a proporção de capital investido no  $i$ -ésimo ativo, no  $j$ -ésimo período, onde  $0 \leq x_{i,j} \leq 1$ , para todo  $i, j$ .
- $c$  representa o vetor de custo de investimento.
- $y_w$  agrupa diversas variáveis, representando, por exemplo, a vazão de gás em um gasoduto, a quantidade de gás adquirida de uma fonte, a quantidade de gás vendido a uma demanda etc, sempre referente ao cenário  $w$ .
- $q^T y_w$  representa as receitas e custos de operação da malha de gás no cenário  $w$ . Os fluxos de caixa são trazidos a valor presente através de uma taxa de desconto, que já está computada em  $q$ .
- A relação (4-2) contém as restrições de primeiro estágio para o investimento. A matriz  $A$  contém restrições para o investimento total, isto é,  $\sum_{pe=1}^{N_{pe}} x_{i,pe} \leq 1$ , para  $i = 1, \dots, m$  onde  $pe \leq N_{pe}$  representa um período.
- $h_w, d_w$  contém os dados das ofertas e demandas nos diversos cenários  $w$ .

- $p_w \in [0, 1]$  é a probabilidade de ocorrência do cenário  $w$  ( $\sum_{w=1}^{N_w} p_w = 1$ ).
- $\rho_\beta$  é o mínimo requerido, em unidade monetária, para a média do retorno dos piores cenários que possuem probabilidade conjunta maior ou igual a  $\beta$ .
- $a$  é uma variável escalar auxiliar para criação das restrições de risco. A abordagem proposta também garante que a variável  $a$  é igual ao VaR do modelo otimizado na solução com restrições de risco ativas. Esta variável equivale ao  $a$  de (3-5) e (3-4).
- a equação (4-1) contém a função objetivo do problema, de natureza estocástica.
- As relações (4-4) e (4-5) contém as restrições de segundo estágio do problema. Nessas restrições encontram-se regras de negócio de natureza física, financeira e econômica.
- As relações (4-3) e (4-6) representam as restrições de CVaR para o primeiro e segundo estágio, respectivamente. Estas relações se equivalem às relações para restringir o risco em (3-6). O CVaR mede o risco da soma do resultado operacional e dos custos de investimento. A inclusão dos custos de investimento pode ser descartada se o montante de investimento for uma constante. Nesse caso, devido à invariância a translações, o montante pode ser considerado modificando diretamente o parâmetro  $\rho_\beta$ .

Ao modelo estocástico gerado e resolvido como o modelo determinístico, sem o uso de técnicas especiais, dá-se o nome de modelo *determinístico equivalente*. O problema descrito por (4-1) - (4-7) será referido dessa forma até o fim do texto.

## 4.2 Decomposição do Modelo Proposto

O problema determinístico equivalente apresenta duas formas de acoplamento:

- Acoplamento por estágios: a equação (4-4) acopla as variáveis de primeiro e segundo estágio do modelo. Esse tipo de acoplamento pode ser contornado aplicando uma técnica de decomposição como a de Benders [5].

- Acoplamento por cenários: na desigualdade (4-3) a variável  $a$  acopla todos os cenários. Esse acoplamento não é solúvel de forma direta com técnicas de decomposição como a de Benders e devemos portanto propor uma reformulação alternativa do modelo.

São apresentadas a seguir duas alternativas para a decomposição do modelo: uma que utiliza decomposição por planos cortantes e outra que utiliza Relaxação Lagrangeana.

#### 4.2.1 Decomposição por Planos Cortantes

Podemos separar o problema descrito anteriormente em um problema de dois níveis. Considerando como variáveis de primeiro estágio aquelas que não dependem do cenário (não são indexadas por  $w$ ) e variáveis de segundo estágio todas as outras, obtemos o problema em dois níveis:

$$\begin{aligned} \min_{a,x} \quad & c^T x + V(a, x) \\ \text{s.a.} \quad & Ax = b, \\ & x \geq 0, \end{aligned} \tag{4-8}$$

onde o modelo de segundo nível (ou estágio) é dado por

$$V(a, x) = \left\{ \begin{array}{l} \min_{z_w, y_w} \quad \sum_{w=1}^{N_w} p_w q_w^T y_w \\ \text{s.a.} \quad \frac{1}{1-\beta} \sum_{w=1}^{N_w} p_w z_w \leq \rho_\beta - a, \\ W y_w = h_w - T x, \quad \text{para } w=1, \dots, N_w, \\ B_w y_w = d_w, \quad \text{para } w=1, \dots, N_w, \\ z_w - q_w^T y_w \geq c^T x - a, \quad \text{para } w=1, \dots, N_w, \\ y_w \geq 0, z_w \geq 0, \quad \text{para } w=1, \dots, N_w. \end{array} \right. \tag{4-9}$$

Em (4-9) a restrição de CVaR dada por  $\frac{1}{1-\beta} \sum_{w=1}^{N_w} p_w z_w \leq \rho_\beta - a$  acopla todos os cenários. Levando então para o primeiro estágio o conjunto de variáveis  $z_w$  e a restrição que causa o acoplamento, obtém-se o seguinte modelo alternativo

$$\begin{aligned}
 \min_{a,x,z_w} \quad & c^T x + \sum_{w=1}^{N_w} p_w V_w(a, x, z_w) \\
 \text{s.a.} \quad & Ax = b, \\
 & a + \frac{1}{1-\beta} \sum_{w=1}^{N_w} p_w z_w \leq \rho\beta, \\
 & x \geq 0, z_w \geq 0, \quad \text{para } w=1, \dots, N_w.
 \end{aligned} \tag{4-10}$$

$$V_w(a, x, z_w) = \begin{cases} \min_{y_w} & q_w^T y_w \\ \text{s.a.} & Wy_w = h_w - Tx, \quad \text{para } w=1, \dots, N_w, \\ & B_w y_w = d_w, \quad \text{para } w=1, \dots, N_w, \\ & -q_w^T y_w \geq c^T x - z_w - a, \quad \text{para } w=1, \dots, N_w, \\ & y_w \geq 0, \quad \text{para } w=1, \dots, N_w. \end{cases}$$

A função  $\sum_{w=1}^{N_w} p_w V_w(a, x, z_w)$  é convexa e poliedral por partes, porque  $V_w(a, x, z_w)$  é definida por uma função objetivo linear e um conjunto de restrições lineares que definem uma região viável poliedral.

A função  $\sum_{w=1}^{N_w} p_w V_w(a, x, z_w)$  pode ser aproximada iterativamente por uma técnica de planos cortantes. Com efeito, para  $\bar{a}, \bar{x}, \bar{z}_w$  fixos, ao se resolver

$$V_w(\bar{a}, \bar{x}, \bar{z}_w) = \begin{cases} \min_{y_w} & q_w^T y_w \\ \text{s.a.} & Wy_w = h_w - T\bar{x}, \quad \text{para } w=1, \dots, N_w, \\ & B_w y_w = d_w, \quad \text{para } w=1, \dots, N_w, \\ & -q_w^T y_w \geq c^T \bar{x} - \bar{z}_w - \bar{a}, \quad \text{para } w=1, \dots, N_w, \\ & y_w \geq 0, \quad \text{para } w=1, \dots, N_w. \end{cases}$$

tem-se o valor da função  $V_w(\bar{a}, \bar{x}, \bar{z}_w)$  e um subgradiente  $g_w(\bar{a}, \bar{x}, \bar{z}_w) = Jc(\bar{a}, \bar{x}, \bar{z}_w)^T \tilde{L}_{\bar{a}, \bar{x}, \bar{z}_w}$ , onde  $\tilde{L}_{\bar{a}, \bar{x}, \bar{z}_w}$  é o multiplicador de Lagrange associado à solução ótima de  $V_w(\bar{a}, \bar{x}, \bar{z}_w)$  e  $Jc(\bar{a}, \bar{x}, \bar{z}_w)$  é o Jacobiano das restrições, de forma similar ao exemplo descrito no Apêndice, na seção A.3 .

Logo, a função  $\sum_{w=1}^{N_w} p_w V_w(a, x, z_w)$  é aproximada por  $\sum_{w=1}^{N_w} p_w \hat{V}_w(a, x, z_w)$ , onde  $\hat{V}_w(a, x, z_w) = \max_k V_w(a^k, x^k, z_w^k) + g_w(a^k, x^k, z_w^k)^T [(a, x, z) - (a^k, x^k, z_w^k)]$

O modelo de planos cortantes obtido para (4-10) é

$$\begin{aligned}
 \min_{a,x,z_w,r} \quad & c^T x + r \\
 \text{s.a.} \quad & Ax = b, \\
 & a + \frac{1}{1-\beta} \sum_{w=1}^{N_w} p_w z_w \leq \rho\beta \\
 r \geq \sum_{w=1}^{N_w} p_w V_w(a^j, x^j, z_w^j) + \sum_{w=1}^{N_w} p_w g_w(a^j, x^j, z_w^j)^T [(a, x, z) - (a^j, x^j, z_w^j)], \\
 & \text{para todo } j \leq k, \\
 0 \geq \|\tilde{z}_{p^l}\|_1 + g'_w(a^l, x^l, z_w^l)^T [(a, x, z) - (a^l, x^l, z_w^l)], \\
 & \text{para todo } l \leq t, \\
 x \geq 0, z_w \geq 0, \quad & \text{para } w=1, \dots, N_w,
 \end{aligned}$$

onde  $k$  é o número de cortes de otimalidade e  $t$  o número de cortes de viabilidade gerados. Neste modelo já são considerados os cortes de viabilidade segundo exemplificado na seção A.3.

Se, ao invés de aproximar  $\sum_{w=1}^{N_w} p_w V_w(a, x, z_w)$  com um corte por iteração, utilizarmos um corte para cada subproblema resolvido na iteração, então a informação do segundo estágio pode ser melhor aproveitada. Esse modelo, do tipo multicortes, seria então da forma

$$\begin{aligned}
 \min_{a,x,z_w,r} \quad & c^T x + \sum_{w=1}^{N_w} p_w r_w \\
 \text{s.a.} \quad & Ax = b, \\
 & a + \frac{1}{1-\beta} \sum_{w=1}^{N_w} p_w z_w \leq \rho\beta, \\
 r_w \geq V_w(a^j, x^j, z_w^j) + g_w(a^j, x^j, z_w^j)^T [(a, x, z) - (a^j, x^j, z_w^j)], \\
 & \text{para todo } j \leq k \text{ e } w=1, \dots, N_w, \\
 0 \geq \|\tilde{z}_{p^l}\|_1 + g'_w(a^l, x^l, z_w^l)^T [(a, x, z) - (a^l, x^l, z_w^l)], \\
 & \text{para todo } l \leq t, \\
 x \geq 0, z_w \geq 0, \quad & \text{para } w=1, \dots, N_w.
 \end{aligned}$$

O modelo multicortes, apesar de aproximar de forma desagregada (mais precisa) cada função  $V_w(a, x, z_w)$ , possui a desvantagem de aumentar o número de variáveis e restrições do modelo de primeiro estágio. A escolha pelo modelo de um corte ou multicortes deve ser feita caso a caso. Como regra geral, caso o custo computacional do problema de primeiro estágio seja muito menor que o dos problemas de segundo estágio, é acon-

selhável utilizar o método multicortes.

#### 4.2.2 Decomposição por Relaxação Lagrangeana

Fazendo uma abordagem clássica, colocamos no primeiro estágio as variáveis que não dependem de cenário  $w$  e no segundo estágio as restantes, obtendo o problema em dois níveis (4-8)-(4-9).

A Relaxação Lagrangeana é eficaz, entre outros casos, quando temos restrições complicadoras que aumentam consideravelmente o custo computacional de resolução. A primeira restrição de (4-9) causa acoplamento dos cenários, sendo uma candidata natural para relaxação. Dualizando a primeira restrição de (4-9) temos

$$\tilde{V}(\lambda, a, x) = \begin{cases} \min_{z_w, y_w} & \sum_{w=1}^{N_w} p_w q_w^T y_w + \lambda \left[ \frac{1}{1-\beta} \sum_{w=1}^{N_w} p_w z_w - \rho_\beta + a \right] \\ \text{s.a.} & W y_w = h_w - T x, \quad \text{para } w=1, \dots, N_w, \\ & B_w y_w = d_w, \quad \text{para } w=1, \dots, N_w, \\ & z_w - q_w^T y_w \geq c^T x - a, \quad \text{para } w=1, \dots, N_w, \\ & y_w \geq 0, z_w \geq 0, \quad \text{para } w=1, \dots, N_w. \end{cases}$$

Assim, é possível reescrever o problema original, agora decomposto por cenários, na forma

$$\begin{aligned} \min_{a, x} & c^T x + M(a, x) \\ \text{s.a.} & Ax = b, \\ & x \geq 0, \end{aligned}$$

onde  $M(a, x) = \max_{\lambda} \theta(\lambda, a, x)$  e

$$\theta(\lambda, a, x) := \lambda [a - \rho_\beta] + \sum_{w=1}^{N_w} \tilde{V}_w(\lambda, a, x) \quad (4-11)$$

para

$$\tilde{V}_w(\lambda, a, x) = \begin{cases} \min_{z_w, y_w} & p_w \left[ q_w^T y_w + \lambda \frac{1}{1-\beta} z_w \right] \\ \text{s.a.} & W y_w = h_w - T x, \quad \text{para } w=1, \dots, N_w, \\ & B_w y_w = d_w, \quad \text{para } w=1, \dots, N_w, \\ & z_w - q_w^T y_w \geq c^T x - a, \quad \text{para } w=1, \dots, N_w, \\ & y_w \geq 0, z_w \geq 0, \quad \text{para } w=1, \dots, N_w. \end{cases} \quad (4-12)$$

Alternativamente, a ordem dos problemas de otimização poderia ser invertida, obtendo

$$\max_{\lambda} M(\lambda)$$

onde

$$M(\lambda) = \begin{cases} \min_{a,x} & c^T x + \tilde{V}(a, x, \lambda) \\ s.a. & Ax = b, \\ & x \geq 0, \end{cases}$$

e

$$\tilde{V}(\lambda, a, x) = \begin{cases} \min_{z_w, y_w} & \sum_{w=1}^{N_w} p_w q_w^T y_w + \lambda \left[ \frac{1}{1-\beta} \sum_{w=1}^{N_w} p_w z_w - \rho\beta + a \right] \\ s.a. & Wy_w = h_w - Tx, \quad \text{para } w=1, \dots, N_w, \\ & B_w y_w = d_w, \quad \text{para } w=1, \dots, N_w, \\ & z_w - q_w^T y_w \geq c^T x - a, \quad \text{para } w=1, \dots, N_w, \\ & y_w \geq 0, z_w \geq 0, \quad \text{para } w=1, \dots, N_w. \end{cases}$$

A resolução computacional desse modelo em três níveis será realizada novamente através de planos cortantes. Os cortes serão gerados de forma similar ao exemplo anterior.

## 5 Resultados

Este capítulo contém os resultados computacionais obtidos com o modelo estocástico proposto. São também apresentados testes de validação, resultados com alguns casos de interesse, e observações de importância prática sobre o condicionamento do modelo.

### 5.1 Custo computacional do modelo Determinístico Equivalente

O modelo determinístico equivalente pode ser resolvido utilizando pacotes comerciais. Utilizando o *solver* CPLEX 11<sup>1</sup>, é possível gerar instâncias do problema com diferentes números de cenários. Para verificar o custo computacional do modelo determinístico equivalente, foi gerado inicialmente um modelo com apenas um cenário. Esse modelo, assim como a versão determinística do problema, é resolvido em cerca de 0,5 segundos. O número de cenários foi então incrementado progressivamente, até que os recursos computacionais do computador utilizado se esgotaram. O resultado deste teste pode ser visto na Tabela 5.1. A medida que o número de cenários aumenta, o tempo de otimização aumenta consideravelmente. Fazendo uma boa escolha de parâmetros do *solver*, é possível obter a solução de modelos determinísticos equivalentes até a ordem de 63 cenários. Além desse número de cenário a máquina esgota sua memória. O modelo de 63 cenários é da ordem de 1.845.000 restrições e 1.860.000 variáveis.

Na Figura 5.1 encontra-se a curva de tempo de otimização dado o número de cenários. O último modelo otimizado leva cerca de 33 minutos para sua geração e otimização, ao passo que o modelo com apenas um cenário toma menos de um segundo para o mesmo. Visualmente nota-se um crescimento exponencial no tempo para otimização. Esse crescimento exponencial desencoraja a busca por um computador com maiores recursos para resolver o problema com 80 cenários, como os casos

<sup>1</sup>página do software em <http://www.ilog.com/products/optimization/>

$N^\circ$ de séries	tempo (s)	$N^\circ$ de séries	tempo (s)	$N^\circ$ de séries	tempo (s)
1	3	22	155	43	403
2	6	23	175	44	449
3	9	24	206	45	456
4	13	25	219	46	469
5	16	26	151	47	505
6	20	27	194	48	591
7	39	28	213	49	703
8	44	29	213	50	785
9	47	30	225	51	883
10	55	31	254	52	1051
11	61	32	298	53	982
12	66	33	308	54	1150
13	70	34	315	55	1522
14	78	35	310	56	1310
15	92	36	322	57	1879
16	92	37	326	58	1442
17	99	38	425	59	1530
18	118	39	370	60	1555
19	113	40	391	61	1707
20	126	41	366	62	2037
21	155	42	382	63	1972

Tabela 5.1: Tempo de otimização para determinístico equivalente, por número de cenários

típicos de interesse. Esses resultados reforçam a necessidade de aplicar métodos de decomposição.

## 5.2 Validação dos métodos de decomposição

Para assegurar que a resolução do modelo, utilizando os métodos de decomposição, resulta em soluções ótimas, foi feita inicialmente uma validação dos métodos, comparando os resultados obtidos pelo modelo determinístico equivalente.

Instâncias do problema com um número reduzido de cenários podem ser resolvidas diretamente, sem a aplicação de decomposição, no modelo determinístico equivalente, conforme visto na seção anterior. O primeiro teste de validação realizado consistiu em gerar algumas instâncias do problema, com 10, 20 30, 40, e 50 cenários.

As instâncias foram resolvidas com o modelo determinístico equivalente, o de Relaxação Lagrangeana e o de planos cortantes. Na Tabela 5.2

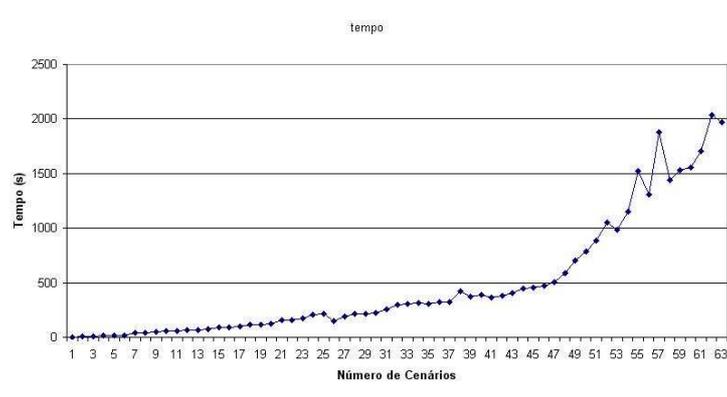


Figura 5.1: Tempo de otimização para determinístico equivalente, por número de cenários

encontra-se o valor da função objetivo (F.O.) e o CVaR obtido para as diversas instâncias resolvidas com o determinístico equivalente. O tempo de execução de cada instância também se encontra na mesma tabela. Esses cenários foram gerados com a restrição de CVaR desativada. Na Tabela 5.3 encontram-se os desvios percentuais dos modelos de Relaxação Lagrangeana e de planos cortantes em relação aos resultados obtidos para o modelo determinístico equivalente.

O segundo teste realizado foi com a restrição de CVaR ativa. Para isso, em cada instância do problema foi estabelecido um limite máximo para a perda equivalente ao CVaR obtido no teste anterior mais 5 ( $\times 10^6$  \$). Dessa forma há a garantia de que a restrição de CVaR está ativa. De forma equivalente ao teste anterior, nas Tabelas 5.4 e 5.5 encontram-se os resultados desse teste.

O *gap*, distância entre o valor ótimo teórico e o valor da melhor solução obtida, foi definido como um valor absoluto de  $10^{-2}(\times 10^6$  \$) para todas as otimizações utilizando métodos de decomposição. Ainda assim, é usual o modelo obter soluções com precisão maior. Soluções ótimas podem ser obtidas com tempo computacional significativamente menor se o *gap* for relaxado. As soluções obtidas com os métodos de decomposição coincidiram com a solução do problema original em todas as instâncias.

O tempo computacional obtido com os métodos de decomposição é substancialmente maior do que o tempo obtido com o modelo determinístico equivalente, nos casos testados até cinquenta cenários. Esse resultado indica que o modelo determinístico equivalente é a melhor opção quando

Modelo	Cenários	Determinístico Equivalente
F.O. (10 <sup>6</sup> \$)		-5513,742980
CVaR (10 <sup>6</sup> \$)	50	-3744,227126
Tempo (s)		383
F.O. (10 <sup>6</sup> \$)		-5522,865797
CVaR (10 <sup>6</sup> \$)	40	-3573,681025
Tempo (s)		197
F.O. (10 <sup>6</sup> \$)		-5633,14629
CVaR (10 <sup>6</sup> \$)	30	-3837,164609
Tempo (s)		118
F.O. (10 <sup>6</sup> \$)		-5494,875716
CVaR (10 <sup>6</sup> \$)	20	-3380,455574
Tempo (s)		65
F.O. (10 <sup>6</sup> \$)		-5629,711133
CVaR (10 <sup>6</sup> \$)	10	-2704,594714
Tempo (s)		30

Tabela 5.2: Desempenho do determinístico equivalente com restrição de CVaR desativada

Modelo	Cenários	Planos Cortantes	Relaxação Lagrangeana
F.O.		0%	0%
CVaR	50	0%	-0,0001%
Tempo		985,1175%	1464,2298%
F.O.		0,0002%	0%
CVaR	40	0%	0,0009%
Tempo		2125,3807%	4336,5482%
F.O.		0,0002%	0%
CVaR	30	0,0002%	0%
Tempo		2955,9322%	2555,0847%
F.O.		0%	0%
CVaR	20	0%	0%
Tempo		5447,6923%	3427,6923%
F.O.		0,0001%	0%
CVaR	10	0,0009%	0,0003%
Tempo		6746,6667%	4493,3333%

Tabela 5.3: Comparação de desempenho dos métodos de decomposição com restrição de CVaR desativada

Modelo	Cenários	Determinístico Equivalente
F.O. (10 <sup>6</sup> \$)		-5511,823838
CVaR (10 <sup>6</sup> \$)	50	-3749,000000
Tempo (s)		282
F.O. (10 <sup>6</sup> \$)		-5522,130062
CVaR (10 <sup>6</sup> \$)	40	-3578,680000
Tempo (s)		205
F.O. (10 <sup>6</sup> \$)		-5630,674312
CVaR (10 <sup>6</sup> \$)	30	-3842,160000
Tempo (s)		133
F.O. (10 <sup>6</sup> \$)		-5492,936618
CVaR (10 <sup>6</sup> \$)	20	-3385,450000
Tempo (s)		74
F.O. (10 <sup>6</sup> \$)		-5628,999271
CVaR (10 <sup>6</sup> \$)	10	-2709,590000
Tempo (s)		32

Tabela 5.4: Desempenho do determinístico equivalente com restrição de CVaR ativada

Modelo	Cenários	Planos Cortantes	Relaxação Lagrangeana
F.O.		0%	0%
CVaR	50	0%	0%
Tempo		6174,4681%	2882,2695%
F.O.		0,0129%	0%
CVaR	40	-0,0033%	0%
Tempo		5861,9512%	2019,0244%
F.O.		0,0053%	0%
CVaR	30	0%	0%
Tempo		5035,3383%	11891,7293%
F.O.		0,0001%	0%
CVaR	20	0%	-0,0001%
Tempo		10121,6216%	3967,5676%
F.O.		0%	0%
CVaR	10	0%	0%
Tempo		12565,6250%	26893,7500%

Tabela 5.5: Comparação de desempenho dos métodos de decomposição com restrição de CVaR ativada

o número de cenários é pequeno.

Ainda assim, os métodos de decomposição utilizados podem ser beneficiados pelo uso de computação paralela. O tempo computacional pode ser severamente reduzido ao otimizarmos os subproblemas de cada cenário nas duas estratégias de decomposição estudadas.

### 5.3 Planos cortantes: Variantes

A decomposição por planos cortantes pode gerar um único corte de otimalidade por iteração, ou um corte para cada cenário, conforme mencionado na seção 4.2.1. Os problemas resolvidos com o modelo de multicortes nesse estudo de caso obtiveram menores tempos de otimização e menores *gaps*, freqüentemente atingindo o *gap* ótimo, igual a zero. Esse resultado indica que o aumento no número de restrições do problema de primeiro estágio é compensado pela convergência acelerada do método.

Como exemplo de desempenho, resolvemos uma mesma instância do problema cinco vezes com cada variante. O tempo médio de execução do modelo multicortes foi de 1500 segundos, enquanto que o tempo médio do modelo de um corte foi de 10530 segundos, cerca de 600% maior que o anterior. Por esses motivos, todos os resultados apresentados a seguir utilizam o modelo multicortes.

### 5.4 Resultados obtidos

As seções anteriores comprovam a superioridade dos métodos de decomposição implementados. Também foi esclarecido que, devido a limitações computacionais, não é possível resolver o problema com oitenta cenários, caso de interesse prático, sem recorrer aos métodos de decomposição. Nessa seção apresentamos os resultados obtidos para o problema estocástico de oitenta cenários.

Primeiramente resolvemos o problema estocástico relaxando a restrição de CVaR. O portfólio obtido possui Valor Presente Líquido (VPL) esperado de 5491 ( $10^6$ \$) e CVaR-90% de 3938 ( $10^6$ \$). O modelo determinístico original apresentava um portfólio cujo VPL esperado era de 5151 ( $10^6$ \$) e CVaR-90% de 3434 ( $10^6$ \$). Esse resultado representa uma melhora de 6,61% no VPL e 14,7% no CVaR. Essa melhora é consequên-

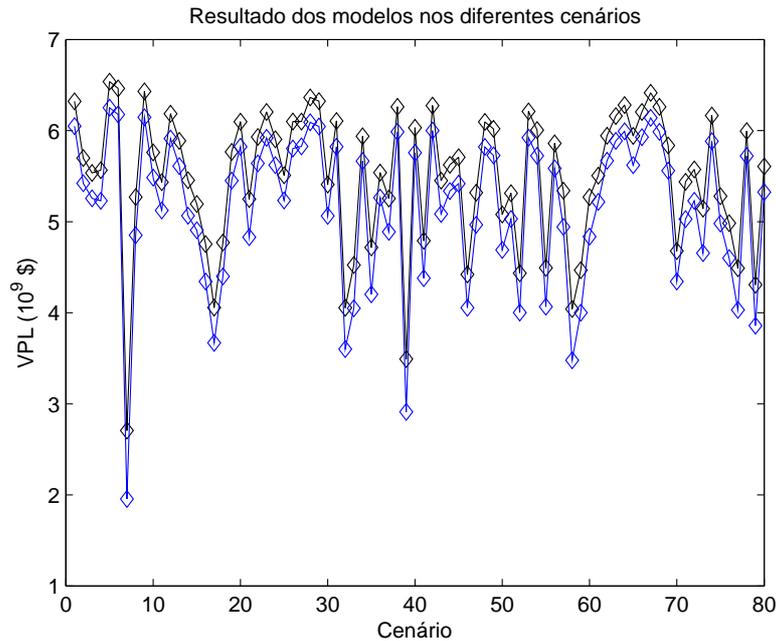


Figura 5.2: Comparação do nível de alocação proposto pelos modelos estocástico e determinístico

cia direta da abordagem estocástica, que evita o fenômeno da falha das médias.

Na Figura 5.2 encontra-se o VPL obtido pelo portfólio determinado no modelo estocástico e o VPL obtido pelo portfólio no modelo determinístico. Verifica-se que o retorno do modelo estocástico domina o retorno do modelo determinístico, isto é, o modelo estocástico apresentou um portfólio com VPL superior ao do modelo determinístico para qualquer cenário realizado.

Claramente o portfólio determinado pelo modelo estocástico é diferente daquele determinado pelo modelo determinístico. Na Tabela 5.6 encontram-se os níveis de alocação propostos por ambos modelos para cada ativo disponível. Verifica-se que, conforme previsto pela teoria, o modelo estocástico diversifica mais o investimento, apresentando menores alocações em alguns ativos e investindo em ativos que não tinham sido escolhidos pelo modelo determinístico.

É ainda possível ativar a restrição de CVaR, obtendo por exemplo uma solução com CVaR de 3941 ( $10^6$ €) e VPL esperado de 5489 ( $10^6$ €). Esse resultado representa uma melhora de 6,57% no VPL e 14,76% no CVaR em relação ao modelo determinístico.

Ativo	Alocação	
	Estocástico	Determinístico
1	100,00%	100,00%
4	62,60%	76,90%
7	100,00%	0,00%
8	100,00%	100,00%
9	25,40%	25,40%
0A	100,00%	100,00%
0B	100,00%	100,00%
0C	17,61%	35,71%
0D	0,00%	0,00%
0E	29,46%	0,00%
0F	0,00%	0,00%
0G	100,00%	100,00%
0H	16,79%	18,55%
0I	35,36%	35,36%
0J	0,00%	0,00%

Tabela 5.6: Comparação do nível de alocação proposto pelos modelos estocástico e determinístico.

### 5.5 Fronteira Eficiente

O conceito de fronteira eficiente, já abordado anteriormente, é usual na teoria de finanças. Podemos também, para esse problema, obter uma fronteira eficiente, variando o parâmetro de CVaR máximo da restrição de CVaR do problema.

A Figura 5.3 representa a fronteira eficiente do problema estudado. Para gerar essa fronteira foi realizado o seguinte procedimento: inicialmente o modelo foi otimizado sem nenhuma restrição de risco, gerando o ponto da curva de menor CVaR. A partir desse ponto foram geradas diversas instâncias do problema, com um requisito de CVaR de 1 ( $10^6$ €) a mais que o problema anterior. O procedimento foi interrompido ao obter a primeira solução inviável, indicando a inexistência de um portfólio que atendesse à restrição de risco informada. Nesse teste foi utilizado nível de confiança para o CVaR de  $\beta=0,9$  (dez por cento piores cenários).

Nota-se que, para o caso estudado, a variabilidade do CVaR é muito menor que a ordem de grandeza dos retornos. Esse fato é consequência da existência de poucos projetos de investimento em ativos que potencialmente mitiguem os riscos dos piores cenários. Esse fato é um indicador de que o decisor que queira uma maior proteção contra o risco deve propor

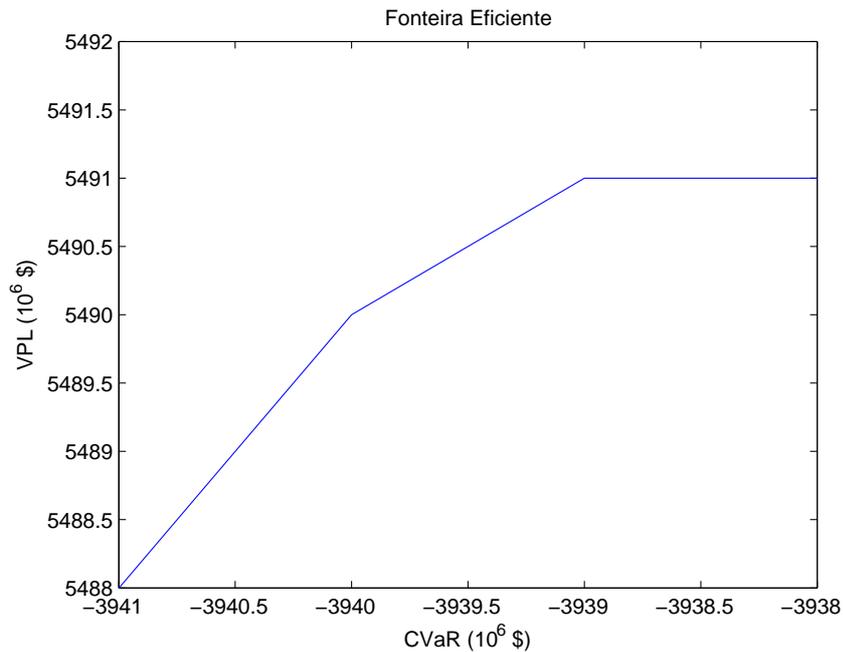


Figura 5.3: Fronteira Eficiente

novos investimentos, a fim de poder melhorar o retorno para os cenários mais desfavoráveis.

### 5.6 Sensibilidade do nível de confiança

Um teste de interesse consiste em analisar a sensibilidade do problema a variações do nível de confiança do CVaR. Todos os testes realizados nas seções anteriores possuem nível de confiança de 0,9 (equivalente aos 10% piores cenários). O portfólio obtido e o valor da função objetivo atingida devem ser diferentes para distintos níveis de confiança.

Neste último teste o nível de confiança do CVaR varia de 0,9 a 0. A Tabela 5.7 resume os resultados obtidos.

Nota-se que o valor da função objetivo aumenta à medida que o nível de confiança requisitado diminui. Esse comportamento é compatível com a teoria, porque a redução do nível de confiança permite que mais cenários com bons retornos compensem os cenários de pior retorno, uma vez que o CVaR é medido através da média dos retornos desses cenários. O CVaR para nível de confiança 0 é, por definição, igual ao valor esperado do VPL.

Nível de confiança	Função objetivo ( $10^6\$$ )	CVaR ( $10^6\$$ )
0,9	-5490,566115	-3940
0,8	-5491,192176	-4274,17641
0,7	-5491,192176	-4556,17697
0,6	-5491,192176	-4754,62429
0,5	-5491,192176	-4907,06641
0,4	-5491,192176	-5044,73603
0,3	-5491,192176	-5173,13853
0,2	-5491,192176	-5287,6091
0,1	-5491,192176	-5391,30776
0	-5491,192176	-5491,19218

Tabela 5.7: Teste de sensibilidade do nível de confiança do CVaR

### 5.7 Condicionamento do problema

As restrições (4-4) e (4-6) acoplam o primeiro e segundo estágios. Segundo a natureza do problema, os portes de  $x$  e  $y$  podem ser tão diferentes que a satisfação de (4-4) e (4-6) pode resultar numericamente impossível. É importante que as matrizes do problema de primeiro estágio e de segundo estágio estejam bem condicionadas para que haja convergência da solução. Este problema de mau condicionamento pode ser resolvido re-escalando as variáveis, segundo mostrado no exemplo a seguir.

Suponha que o problema de segundo estágio resultou em uma inviabilidade da ordem de  $10^{-6}$ . Existem dois investimentos propostos  $x_1, x_2 \in [0, 1]$ . No problema que resultou inviável,  $x_1$  e  $x_2$  possuem valores 1 e 0 fixados pelo primeiro estágio, respectivamente. O corte de viabilidade gerado a partir dos valores marginais é

$$34000(x_1 - 1) + 0,1(x_2 - 0) \geq 10^{-6}.$$

O problema de primeiro estágio nesse caso deveria fazer um pequeno aumento do investimento  $x_2$ , para se evitar a inviabilidade detectada. O aumento seria da ordem de:

$$x_2 = \frac{10^{-6}}{0,1} = 10^{-5}.$$

Na prática, pode acontecer do problema de primeiro estágio não apresentar uma nova solução com aumento do investimento  $x_2$ . De fato, todos os pacotes de otimização numérica estão sujeitos a critérios de

tolerância para determinar a viabilidade de restrições, a otimalidade de soluções, e os limites de variáveis. Suponha que no otimizador utilizado haja, por exemplo, uma tolerância de  $10^{-8}$  para limites de variáveis. As variáveis de investimento do problema em questão serão consideradas dentro dos limites se estiverem entre  $[-10^{-8}, 1+10^{-8}]$ . O valor da variável de investimento  $x_1$  necessário para tornar o problema de primeiro estágio viável com a nova restrição é de

$$x_1 = 1 + \frac{10^{-6}}{34000} < 1 + 10^{-8}.$$

Assim sendo, o otimizador numérico poderá considerar como solução viável e ótima uma solução que incrementa  $x_1$  acima do limite real, ao invés de incrementar  $x_2$ . A solução para esse problema é um re-escalamento adequado das variáveis.

Suponha agora que  $x_1, x_2 \in [0, 10^4]$ . A mesma restrição de viabilidade anterior, com os coeficientes reajustados, seria da forma

$$3,4(x_1 - 10^4) + 10^{-4}(x_2 - 0) \geq 10^{-6}.$$

O valor de  $x_1$  necessário para que o problema de primeiro estágio resulte em soluções viáveis nesse caso é

$$x_1 \geq 10^4 + \frac{10^{-6}}{3,4} > 10^4 + 10^{-8}.$$

Logo, o problema de primeiro estágio não considera o incremento de  $x_1$  uma solução viável. Já o incremento de  $x_2$  é da ordem de

$$x_2 \geq \frac{10^{-6}}{10^{-4}} = 10^{-2}.$$

Assim, o investimento proposto para  $x_2$  será de  $10^{-2}$ .

O problema abordado neste trabalho apresentou um fenômeno de mau condicionamento similar. O mau condicionamento observado foi consequência da existência da restrição de CVaR. As variáveis de investimento foram portanto re-escaladas. O re-escalamento foi feito uma única vez, baseado em valores típicos observados para os cortes gerados. Essa solução foi suficiente para os casos abordados. Como prática geral, é recomendável utilizar um procedimento automático de re-escalamento das variáveis, ajustando seu problema de acordo com a precisão numérica necessária.

## 6 Conclusão

### 6.1 Conclusões

Os últimos avanços na teoria de portfólio foram, principalmente, na melhoria da gestão de riscos. O objetivo deste trabalho foi de desenvolver um modelo para apoiar a escolha de um portfólio de investimento em ativos reais, atendendo a certos critérios de otimalidade e restrições operacionais e financeiras. O modelo estocástico proposto incorpora restrições de CVaR, permitindo realizar gestão de riscos do portfólio.

No estudo de caso realizado verificou-se que o modelo estocástico proposto foi capaz de utilizar a informação de demanda disponível de forma mais eficiente, obtendo alocações com maior lucro esperado e menor risco, quando comparado com o modelo determinístico. Esse comportamento é consequência direta do uso do modelo estocástico, que não incorre na falha das médias.

A inclusão de uma restrição de CVaR no modelo é uma abordagem interessante para a gestão de riscos, pois essa medida possui boas propriedades, tanto do ponto de vista da teoria de finanças quanto do ponto de vista de otimização. Contudo, é imprescindível que o decisor proponha investimentos em ativos que potencialmente mitiguem riscos do portfólio. Um bom exemplo no caso abordado seriam alternativas de flexibilização da oferta e demanda de gás, como investimentos em terminais de Gás Natural Liquefeito (GNL) e contratos flexíveis com consumidores. Sem a presença de alternativas de investimento com esse perfil à disposição, o modelo fica impossibilitado de alocar recursos de forma a atender à restrição de risco imposta.

Um outro uso desse modelo de interesse prático para a indústria é como ferramenta de avaliação (*valuation*, no inglês) de projetos. Investimentos com o objetivo de mitigar riscos, como as alternativas de flexibilização supra-citadas, normalmente possuem retorno esperado baixo ou

até negativo. Por esse motivo, esses projetos podem ser descartados no processo tradicional de avaliação conduzido nas empresas. Com o uso do nosso modelo, a comparação de portfólios com e sem esse ativo permite uma análise incremental, na qual o decisor avalia o impacto da inclusão desse ativo sobre o resultado da companhia, e determina o valor da alternativa como potencial mitigador de riscos.

O uso do modelo estocástico para casos de interesse, que normalmente possuem cerca de oitenta cenários, não é solúvel em tempo computacional razoável. As técnicas de decomposição propostas permitem que o modelo seja resolvido adequadamente. Ainda é possível investir em processamento paralelo para reduzir o tempo computacional, de acordo com a necessidade do decisor.

A incorporação desses avanços no complexo processo de planejamento da indústria do petróleo está apenas começando. Os ganhos associados a esses esforços têm se provado favoráveis na indústria de energia elétrica. As empresas de petróleo que ampliem suas atividades para tornarem-se empresas de energia integradas podem obter potenciais ganhos utilizando essas técnicas em seu planejamento corporativo.

## 6.2 Extensões

- *Redução de Cenários*: Em determinados casos a representação da variabilidade das fontes de incerteza pode requisitar um número elevado de cenários, como por exemplo, utilizando simulação de Monte Carlo. O uso de técnicas de redução de cenários pode ser uma ferramenta útil para compatibilizar o uso desses modelos com as referidas técnicas.
- *Opções Reais*: Considerando que parte dos investimentos previstos na carteira pode ser postergada ou abandonada, existem opções reais sobre a carteira que poderiam ser avaliadas. Essas opções poderiam ser incorporadas ao modelo de análise partindo para uma modelagem conhecida na literatura como *otimização estocástica multi-estágios*. Nesse caso, teríamos uma árvore de cenários, que seria explorada à medida que a informação se revelasse e as decisões de portfólio fossem tomadas. Essa mudança implica em diversas dificuldades adicionais. Primeiro, existe a necessidade de haver dados para a formulação da árvore de cenários. A otimização de

modelos desse tipo é complexa e necessita de técnicas especiais. Já existem aplicações na indústria de energia elétrica, como [9]. Por fim, o CVaR deve ser adaptado nesse ambiente, para que suas propriedades de coerência não sejam perdidas.

- *Decomposição Alternativa:* Outras alternativas de decomposição podem ser estudadas. Uma alternativa seria um modelo diretamente decomposto por Relaxação Lagrangeana, resultando em um modelo em dois níveis. Existem métodos para otimização desse tipo de modelos, como o progressive hedging algorithm [4], e [23], que apresenta métodos mais compatíveis com o acoplamento entre cenários causado pela restrição de CVaR.
- *Métodos de Feixes:* As técnicas de decomposição utilizadas podem ser substituídas por outras mais avançadas, como por exemplo “métodos de feixes” (*bundle methods*, em [5]). É necessário o estudo mais aprofundado dessas técnicas, que envolvem uma complexidade matemática elevada. Entre as vantagens desses métodos, destacam-se um comportamento mais estável do algoritmo entre as iterações, evitando iterações degeneradas, e a possibilidade de eliminar cortes inativos, mantendo o problema de primeiro estágio com um número reduzido de restrições.

## Referências Bibliográficas

- [1] Carlo Acerbi. Spectral measures of risk: A coherent representation of subjective risk aversion. *Journal of Banking & Finance*, 26(7):1505–1518, July 2002. available at <http://ideas.repec.org/a/eee/jbfina/v26y2002i7p1505-1518.html>. 3.3
- [2] P. Artzner, F. Delbaen, J.-M. Eber, and D. Heath. Coherent measures of risk. *Math. Finance*, 9:203–228, 1999. 3.3
- [3] J. F. Benders. Partitioning procedures for solving mixed-variables programming problems. *Numerische Mathematik*, 4:238–252, 1962. A.3
- [4] John R. Birge and François Louveaux. *Introduction to stochastic programming*. Springer-Verlag, New York, 1997. 6.2, A.3
- [5] J. Frédéric Bonnans, J. Ch. Gilbert, C. Lemaréchal, and C. A. Sagastizábal. *Numerical Optimization: Theoretical and Practical Aspects (Universitext)*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2 edition, 2006. 4.2, 6.2, A, A.3
- [6] A. Eichhorn, W. Römisch, and I. Wegner. Polyhedral risk measures in electricity portfolio optimization. *PAMM Proc. Appl. Math. Mech.*, 4(3):7–10, 2004. 3.3
- [7] C. Fábián. Handling cvar objectives and constraints in two-stage stochastic models. *European Journal of Operational Research*, In Press, Corrected Proof, , Available online 29 April 2007(<http://www.sciencedirect.com/science/article/B6VCT-4NKXWJ9-7/2/6db857643b0b75d000764aacd1d1cb5b>), 2004. 3.3

- [8] C. Fábián and Anna Veszpremi. Algorithms for handling cvar-constraints in dynamic stochastic programming models with applications to finance. *The Journal of Risk*, 10(3), 2008. 3.3
- [9] Ronald Hochreiter, Georg Ch. Pflug, and David Wozabal. Multi-stage stochastic electricity portfolio optimization in liberalized energy markets. In F. Ceragioli, Asen L. Dontchev, H. Futura, K. Marti, and Luciano Pandolfi, editors, *System Modelling and Optimization*, volume 199 of *IFIP International Federation for Information Processing*, pages 219–226. Springer, 2005. 6.2
- [10] P. Jorion. *Value at Risk: The New Benchmark for Managing Financial Risk*. McGraw-Hill Professional, first edition, 1996. 3.2
- [11] P. Kall and J. Mayer. *Stochastic Linear Programming: Models, Theory, and Computation*, volume 80 of *International Series in Operations Research and Management Science*. Springer, 2005. 3.1, 3.3
- [12] P. Kaplan. Asset allocation models using the Markowitz approach. Disponível em <http://corporate.morningstar.com/ib/asp/detail.aspx?xmlfile=1409.xml>, 1998. 3.1
- [13] P. Krokmal, J. Palmquist, and S. Uryasev. Portfolio optimization with conditional value-at-risk objective and constraints. *The Journal of Risk*, 4(2):11–27, 2002. 3.3
- [14] M. E. P. Maceira. Programação dinâmica dual estocástica aplicada ao planejamento da operação energética de sistemas hidrotérmicos com representação do processo estocástico de afluências por modelos auto-regressivos periódicos. *Relatório técnico 237/93, Cepel*, 1993. 2
- [15] R. Mansini, W. Ogryczak, and M. G. Speranza. Conditional value at risk and related linear programming models for portfolio optimization. *Annals of Operations Research*, 152(1):227–256, July 2007. 3.3
- [16] H. Markovitz. *Portfolio Selection: Efficient Diversification of Investments*. John Wiley & Sons, first edition, 1959. 3.1

- [17] H. Markowitz. Portfolio selection. *Journal of Finance*, 7(1):77–91, 1952. 3.1
- [18] L. Marzano. *Otimização de Portfólio de Contratos de Energia em Sistemas Hidrotérmicos com Despacho Centralizado*. PhD thesis, Pontifícia Universidade Católica Do Rio De Janeiro - PUC-Rio, 2004. 3.3
- [19] A. Meucci. *Risk and Asset Allocation*. Springer Finance, first edition, 2005. 3.1, 3.3
- [20] C. Prates, E. Pierobon, R. Costa, and V. Figueiredo. Evolução da oferta e da demanda de gás natural no Brasil. *BNDES Setorial*, 2006/09. 1.1
- [21] R. Rockafellar. Coherent approaches to risk in optimization under uncertainty. *INFORMS tutorial*, November 2007. 2
- [22] R. Rockafellar and S. Uryasev. Optimization of conditional value-at-risk. *The Journal of Risk*, 2(3):21–41, 2000. 3.3
- [23] A. Ruszczyński. Decomposition methods in stochastic programming. *Mathematical Programming*, 79(1-3):333–353, 1997. 6.2
- [24] S. Savage. The flaw of averages. *Harvard Business Review*, November:20–21, 2002. 2
- [25] N. Topaloglou, H. Vladimirov, and S.A. Zenios. Cvar models with selective hedging for international asset allocation. *Journal of Banking and Finance*, 26:1535–1561(27), July 2002. 3.3

## A Apêndice

Alguns tópicos de otimização relevantes para o presente trabalho são apresentados nas seguintes seções. Todas as informações aqui contidas podem ser obtidas em mais detalhes em [5].

### A.1 Otimização Não Diferenciável

Nessa seção iremos revisar alguns conceitos de otimização não diferenciável necessários para nossos propósitos.

Seja  $f$  uma função convexa no  $R^n$ . Sabemos que  $f$  é contínua e localmente Lipschitziana. Temos que a derivada direcional  $f'(x, d) := \lim_{t \rightarrow 0} \frac{f(x+td) - f(x)}{t}$  existe para todo  $x$  e  $d$  no  $R^n$ . Conseqüentemente,  $f$  tem um gradiente  $\nabla f(x)$  definido em quase toda parte. Quando  $\nabla f(x)$  não existe, o ponto  $x$  é chamado de *kink*.

Para entender o comportamento de  $f(x)$  perto do *kink*, vamos generalizar o conceito de gradiente. O *subdiferencial* de  $x$ ,  $\partial f(x)$ , é definido como

$$\partial f(x) := \{s \in R^n \mid f(y) \geq f(x) + \langle s, y - x \rangle, \quad \forall y \in R^n\}$$

O subdiferencial  $\partial f(x)$  será então um conjunto não vazio, fechado, convexo e localmente limitado como uma função de  $x$ .  $\partial f(x) = \nabla f(x)$  se, e somente se, a função  $f$  é diferenciável em  $x$ . Cada elemento do conjunto  $\partial f(x)$  é um *subgradiente* de  $f$  no ponto  $x$ .

Na Figura A.1 podemos ver a função  $f(x) = |x|$ , que possui um *kink* na origem. Nesse ponto ocorre  $\partial f(x) = [-1, 1]$ , isto é, todos os reais pertencentes ao intervalo  $[-1, 1]$  são subgradientes da função nesse ponto.

### A.2 Aproximação por planos cortantes

Suponha o problema de minimização irrestrito

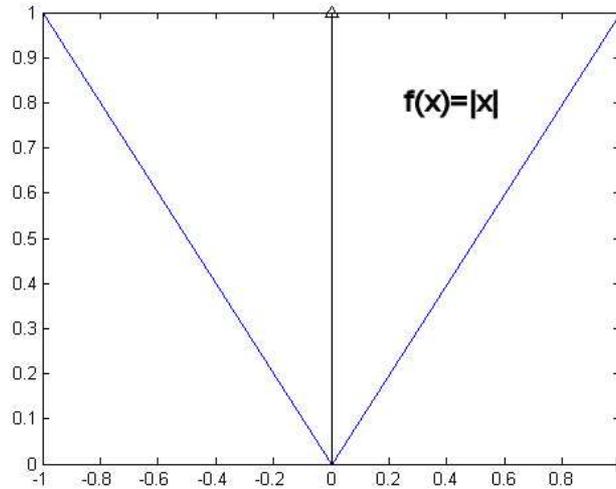


Figura A.1: Exemplo de um “kink”

$$\left\{ \begin{array}{l} \min_{x \in \mathbb{R}} f(x) \end{array} \right. \quad (\text{A-1})$$

onde  $f(x)$  é uma função convexa de difícil representação. Essa função pode ser aproximada por um modelo linear por partes, avaliando a função original em pontos  $x^i$  tal que

$$f_i = f(x^i) \quad e \quad s^i = s(x^i),$$

onde  $s(x^i)$  é o subgradiente da função no ponto  $x^i$ . Construímos então a aproximação

$$\hat{f}_k(x) = \max_{i=1, \dots, k} \{f_i + \langle s^i, x - x^i \rangle\}.$$

O problema (A-1) pode ser aproximado por

$$\left\{ \begin{array}{l} \min_{x, r \in \mathbb{R}} r \\ \text{s.a.} \quad r \geq f_i + \langle s^i, x - x^i \rangle, \quad \text{para todo } i = 1, \dots, k \end{array} \right.$$

A medida que  $k \rightarrow \infty$ , a aproximação  $\hat{f}_k(x)$  converge para  $f(x)$ . A aproximação de  $f(x)$  pelo algoritmo pode ser vista na Figura A.2.

### A.3 Decomposição de Benders

Nesta seção será apresentada a técnica de decomposição de Benders [3], de forma similar a [5]. A versão aqui apresentada é conhecida como método L (*L-shaped method*, em inglês), por apresentar uma extensão

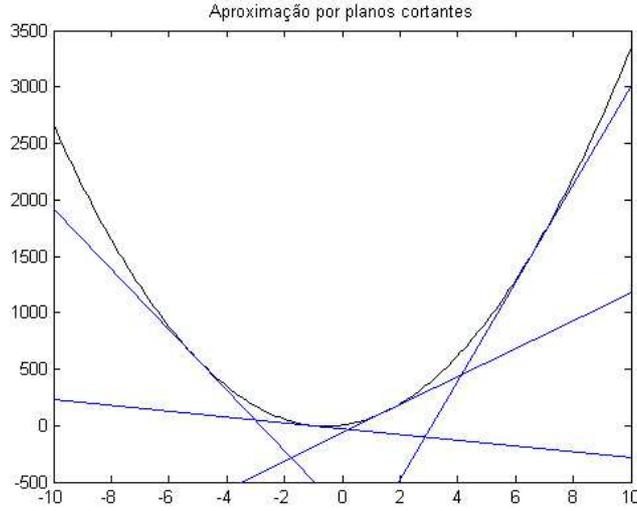


Figura A.2: Aproximação da função  $f(x)$  por planos cortantes

que trata inviabilidades, conforme [4].

Suponha que o problema a ser resolvido é da forma

$$\begin{cases} \min_{p, \tilde{p}} & f_0(p) + \tilde{f}_0(\tilde{p}) \\ s.a. & (p, \tilde{p}) \in P \times \tilde{P}, \\ & c_E(p) + \tilde{c}_E(\tilde{p}) = 0, \\ & c_I(p) + \tilde{c}_I(\tilde{p}) \leq 0, \end{cases}$$

onde as relações que representam o problema de segundo nível (problema de operação do sistema) possuem o superescrito  $\sim$ . Esse mesmo problema pode ser representado da forma

$$\begin{cases} \min_p & f_0(p) + v^{opt}(p) \\ s.a. & p \in P, \end{cases}$$

onde é definida a função de otimalidade

$$v^{opt}(p) = \begin{cases} \min_{\tilde{p}} & \tilde{f}_0(\tilde{p}) \\ s.a. & \tilde{p} \in \tilde{P}, \\ & \tilde{c}_E(\tilde{p}) = -c_E(p), \\ & \tilde{c}_I(\tilde{p}) \leq -c_I(p). \end{cases}$$

Se  $v^{opt}(p)$  é convexa (possivelmente não diferenciável), então é possível calcular um subgradiente usando os multiplicadores de Lagrange associados às restrições de  $v^{opt}(p)$ . Para  $p = p^k$  fixo, obtém-se uma solução  $\tilde{p}_{p^k}$  com multiplicador de Lagrange associado  $\tilde{x}_{p^k}$ , onde

$$v^{opt}(p^k) = \tilde{f}_0(\tilde{p}_{p^k}) \quad e \quad s(p^k) = Jc(p^k)^T \tilde{x}_{p^k} \in \partial v^{opt}(p^k),$$

sendo  $Jc(p^k) = (Jc_E(p^k), Jc_I(p^k))$  o Jacobiano de  $p^k$  na matriz.

Usando um método de planos cortantes, obtemos um problema de primeiro estágio da forma

$$\begin{cases} \min_{p,r} & f_0(p) + r \\ s.a. & p \in P, r \in \mathbb{R}, \\ & r \geq \tilde{f}_0(\tilde{p}_{p^j}) + \langle s(p^j), p - p^j \rangle \quad \text{para todo } j \leq k, \end{cases} \quad (\text{A-2})$$

onde  $k$  é o número de iterações realizadas do algoritmo. A variável  $r$  aproximará o valor de  $v^{opt}(p)$ . Cada restrição para  $r$  em (A-2) é chamada de *corte de otimalidade*.

Pode ocorrer, em alguma iteração, de que o problema de primeiro estágio tenha uma solução  $p_k$  que resultará em um problema de segundo estágio inviável. Essa solução é indesejada, pois não contribui para o processo de otimização. Neste caso, é necessário incluir *cortes de viabilidade*, que restringem o conjunto viável do problema de primeiro estágio, permitindo a geração de uma nova solução  $p_{k+1}$  viável.

Ao obter uma solução  $p_k$  indesejada, recorre-se ao seguinte problema auxiliar:

$$v^{via}(p^k) = \begin{cases} \min_{\tilde{p}} & \|z\|_1 \\ s.a. & \tilde{p} \in \tilde{P}, z_E^+, z_E^-, z_I^+ \geq 0, \\ & \tilde{c}_E(\tilde{p}) + z_E^+ - z_E^- = -c_E(p), \\ & \tilde{c}_I(\tilde{p}) - z_I^+ \leq -c_I(p), \\ & z = (z_E^+ - z_E^-, z_I^+) \in \mathbb{R}^{n_E+n_I}, \end{cases} \quad (\text{A-3})$$

onde  $n_E$  representa o número de restrições de igualdade,  $n_I$  o número de restrições de desigualdade do modelo original e  $\|z\|_1 = \sum_{i=1}^n |z_i|$  é a norma-1 no  $\mathbb{R}^n$ .

A idéia do problema auxiliar é simples.  $v^{via}(p) \geq 0$  para todo  $p$ , mas  $v^{via}(p) = 0$  se, e somente se,  $p$  não resulte em uma inviabilidade para o problema de segundo estágio. Logo, se o problema de primeiro estágio é da forma

$$\left\{ \begin{array}{l} \min_{p,r} \quad f_0(p) + r \\ s.a. \quad p \in P, r \in \mathbb{R}, \\ r \geq \tilde{f}_0(\tilde{p}_{p^j}) + \langle s(p^j), p - p^j \rangle \quad \text{para todo } j \leq k, \\ v^{via}(p) = 0, \end{array} \right.$$

então nunca são geradas soluções indesejadas. Não é possível incorporar diretamente essa restrição ao problema de primeiro estágio, logo utilizamos novamente uma abordagem de planos cortantes para gerar agora os cortes de viabilidade.

Seja  $(\tilde{p}_{p^k}, \tilde{z}_{p^k})$  uma solução para (A-3), e  $\tilde{x}'_{p^k}$  um multiplicador ótimo associado às restrições envolvendo  $c_E(p)$  e  $c_I(p)$ . Temos

$$v^{via}(p^k) = \|\tilde{z}_{p^k}\|_1 \quad e \quad s'(p^k) = Jc(p^k)^T \tilde{x}'_{p^k} \in \partial v^{via}(p^k)$$

formando então os cortes de viabilidade no problema de primeiro estágio

$$\left\{ \begin{array}{l} \min_{p,r} \quad f_0(p) + r \\ s.a. \quad p \in P, r \in \mathbb{R}, \\ r \geq \tilde{f}_0(\tilde{p}_{p^j}) + \langle s(p^j), p - p^j \rangle \quad \text{para todo } j \leq k, \\ 0 \geq \|\tilde{z}_{p^k}\|_1 + \langle s'(p^l), p - p^l \rangle \quad \text{para todo } l \leq t, \end{array} \right.$$

onde  $k$  é o número de cortes de otimalidade e  $t$  o número de cortes de viabilidade gerados.

#### A.4 Relaxação Lagrangeana

Considere o problema primal

$$\left\{ \begin{array}{l} \min_{p \in P} \quad f_0(p) \\ s.a. \quad c_E(p) = 0, \\ c_I(p) \leq 0, \end{array} \right. \quad (\text{A-4})$$

onde  $f_0, c_j : P \rightarrow \mathbb{R}$  e  $P \subset \mathbb{R}^N$ . Associado ao problema (A-4) temos

- um conjunto convexo fechado  $X$ , o espaço dos multiplicadores.
- a função Lagrangeana  $L : P \times X \rightarrow \mathbb{R} \cup \{+\infty\}$ , que possui a propriedade de, para todo  $p \in P$ ,  $\sup_{x \in X} L(p, x) = f_0(p)$  se  $p$  satisfaz as restrições do problema (A-4), e  $\infty$  caso contrário.

Devido a essa última propriedade, temos então que resolver (A-4) é equivalente a resolver

$$\inf_{p \in P} \sup_{x \in X} L(p, x). \quad (\text{A-5})$$

Pode-se provar que resolver (A-5) é equivalente a resolver o problema

$$\sup_{x \in X} \inf_{p \in P} L(p, x). \quad (\text{A-6})$$

Essa inversão em (A-6) é interessante do ponto de vista prático se o problema *dual*

$$\theta(x) = \inf_{p \in P} L(p, x)$$

pode ser resolvido a baixo custo computacional e não é identicamente igual a  $-\infty$ .