

UNIVERSIDADE ESTADUAL DE CAMPINAS  
FACULDADE DE ENGENHARIA ELÉTRICA E DE COMPUTAÇÃO

André du Pin Calmon

VARIAÇÃO DO CONTROLE COMO FONTE DE INCERTEZA



Campinas  
2009

André du Pin Calmon

VARIAÇÃO DO CONTROLE COMO FONTE DE INCERTEZA

Dissertação de mestrado apresentada à Faculdade de Engenharia Elétrica e de Computação como parte dos requisitos exigidos para a obtenção do título de Mestre em Engenharia Elétrica.  
Área de concentração: Automação.

Orientador: João Bosco Ribeiro do Val



Campinas  
2009

Este exemplar corresponde à redação final da Dissertação/Tese defendida por André du Pin Calmon e aprovada através da comissão julgada em 22/05/2009

André du Pin Calmon

VARIAÇÃO DO CONTROLE COMO FONTE DE INCERTEZA

Dissertação de mestrado apresentada à Faculdade de Engenharia Elétrica e de Computação como parte dos requisitos exigidos para a obtenção do título de Mestre em Engenharia Elétrica.  
Área de concentração: Automação.

Orientador: João Bosco Ribeiro do Val



Campinas  
2009

André du Pin Calmon

Engenheiro Eletricista – UnB-DF

VARIAÇÃO DO CONTROLE COMO FONTE DE INCERTEZA

Dissertação de mestrado apresentada à Faculdade de Engenharia Elétrica e de Computação como parte dos requisitos exigidos para a obtenção do título de Mestre em Engenharia Elétrica. Área de concentração: Automação.

Banca Examinadora:

Prof. Dr. João Bosco Ribeiro do Val  
FEEC/UNICAMP

Prof. Dr. Paulo Augusto Valente Ferreira  
FEEC/UNICAMP

Prof. Dr. Marcelo Dutra Fragoso  
LNCC/CNPq

Campinas

2009

FICHA CATALOGRÁFICA ELABORADA PELA  
BIBLIOTECA DA ÁREA DE ENGENHARIA E ARQUITETURA - BAE - UNICAMP

C137v Calmon, André du Pin  
Variação do controle como fonte de incerteza / André  
du Pin Calmon. --Campinas, SP: [s.n.], 2009.

Orientador: João Bosco Ribeiro do Val.  
Dissertação de Mestrado - Universidade Estadual de  
Campinas, Faculdade de Engenharia Elétrica e de  
Computação.

1. Teoria do controle estocástico. 2. Sistemas não-  
lineares. 3. Programação dinâmica. 4. Modelos  
econômicos. 5. Programação convexa. I. Val, João  
Bosco Ribeiro do. II. Universidade Estadual de  
Campinas. Faculdade de Engenharia Elétrica e de  
Computação. III. Título.

Título em Inglês: Control variation as a source of uncertainty

Palavras-chave em Inglês: Stochastic control, Non-Linear systems, Dynamic  
programming, Economic models, Convex programming

Área de concentração: Automação

Titulação: Mestre em Engenharia Elétrica

Banca examinadora: Marcelo Dutra Fragoso, Paulo Augusto Valente Ferreira

Data da defesa: 22/05/2009

Programa de Pós Graduação: Engenharia Elétrica

## COMISSÃO JULGADORA - TESE DE MESTRADO

**Candidato:** André Du Pin Calmon

**Data da Defesa:** 22 de maio de 2009

**Título da Tese:** "Variação do Controle como Fonte de Incerteza"

Prof. Dr. João Bosco Ribeiro do Val (Presidente):

Prof. Dr. Marcelo Dutra Fragoso:

Prof. Dr. Paulo Augusto Valente Ferreira:



PARA OS MEUS PAIS, PAULO E KATYA, E PARA O MEU IRMÃO, FLÁVIO.

# Agradecimentos

Agradeço,

a Deus por ter feito o impossível acontecer.

ao Prof. João Bosco pela orientação acadêmica e de vida. Obrigado por me ajudar a desenvolver e conquistar coisas que eu nunca imaginei que seria capaz.

aos professores Paulo Valente e Pedro Peres, por terem expandido os meus horizontes.

aos meus amigos do DT, Renato, Ricardo, Walter, Taís, Ceará, Giuliano, Júlio, Tiago, Jurandyr, Alessandro, Vitor e Éber, pela grande amizade e pelas fantásticas horas do café.

aos meus amigos de Brasília, Nathalie, Renata, Fontes, Maíce, Chico e Maguila.

aos meus pais, pelo amor e por sempre terem acreditado em mim (mesmo quando eu não acreditava).

ao meu irmão, que é um anjo na minha vida. Obrigado por ser quem você é.

ao Pasqual, thanks for keeping the flame burning.

ao Rodrigo Caun, pelo exemplo de coragem.

aos membros da APOGEEU.

à Raquel, pelo amor, carinho, atenção e paciência.

In retrospect it is interesting to note that the original problem that started my research is still outstanding - namely the problem of planning or scheduling dynamically over time, particularly planning dynamically under uncertainty. If such a problem could be successfully solved it could (eventually through better planning) contribute to the well-being and stability of the world.

George B. Dantzig

# Resumo

Este trabalho apresenta a caracterização teórica e a estratégia de controle para sistemas estocásticos em tempo discreto onde a variação da ação de controle aumenta a incerteza sobre o estado (sistemas VCAI). Este tipo de sistema possui várias aplicações práticas, como em problemas de política monetária, medicina e, de forma geral, em problemas onde um modelo dinâmico completo do sistema é complexo demais para ser conhecido. Utilizando ferramentas da análise de funções não suaves, mostra-se para um sistema VCAI multidimensional que a convexidade é uma invariante da função valor da Programação Dinâmica quando o custo por estágio é convexo. Esta estratégia indica a existência de uma região no espaço de estados onde a ação ótima de controle é de não variação (denominada *região de não-variação*), estando de acordo com a natureza cautelosa do controle de sistemas subdeterminados. Adicionalmente, estudou-se algoritmos para a obtenção da política ótima de controle para sistemas VCAI, com ênfase no caso mono-entrada avaliado através de uma função custo quadrática. Finalmente, os resultados obtidos foram aplicados no problema da condução da política monetária pelo Banco Central.

Palavras-chave: Controle Estocástico, Sistemas Não-Lineares, Programação Dinâmica Aproximada, Análise Convexa, Modelos Macroeconômicos.

# Abstract

This dissertation presents a theoretical framework and the control strategy for discrete-time stochastic systems for which the control variations increase state uncertainty (CVIU systems). This type of system model can be useful in many practical situations, such as in monetary policy problems, medicine and biology, and, in general, in problems for which a complete dynamic model is too complex to be feasible. The optimal control strategy for a multidimensional CVIU system associated with a convex cost functional is devised using dynamic programming and tools from nonsmooth analysis. Furthermore, this strategy points to a region in the state space in which the optimal action is of no variation (the *region of no variation*), as expected from the cautionary nature of controlling underdetermined systems. Numerical strategies for obtaining the optimal policy in CVIU systems were developed, with focus on the single-input input case evaluated through a quadratic cost functional. These results are illustrated through a numerical example in economics.

Key-words: Stochastic Control, Non-Linear Systems, Approximate Dynamic Programming, Convex Analysis, Macroeconomic Models.

# Lista de Figuras

1.1	Ilustração de um sistema VCAI . . . . .	4
3.1	Comportamento de $\partial_u V(x, u)$ para diferentes $x$ . . . . .	22
3.2	Entrada ótima $u^*$ como uma função do valor do estado . . . . .	26
4.1	Diagrama de blocos da Programação Dinâmica Aproximada. . . . .	28
4.2	Comportamento de $h(y)$ . . . . .	32
4.3	Custo mínimo aproximado e política ótima aproximada obtidos utilizando Programação Linear Aproximada. . . . .	38
5.1	Comportamento da inflação, do PIB, da taxa de juros ótima e da variação ótima da taxa de juros, $v_k^*$ , para o modelo VCAI com $\sigma_\pi = \sigma_y = 0.2$ e para o RLQ. . . . .	44
5.2	Comportamento da inflação, do PIB, da taxa de juros ótima e da variação ótima da taxa de juros, $v_k^*$ , para o modelo VCAI com $\sigma_\pi = \sigma_y = 0.1$ e para o RLQ. . . . .	45
5.3	Curvas do valor histórico da taxa de juros, da taxa de juros gerada pelo modelo RLQ tradicional e a taxa de juros obtida utilizando um modelo VCAI. . . . .	46

# Lista de Tabelas

5.1 Estimativas para o modelo trimestral da economia americana. . . . . 42

# Lista de Símbolos

$\mathbb{R}$	conjunto dos números reais
$\mathbb{Z}$	conjunto dos números inteiro
$k$	índice de tempo
$\bar{x}$	vetor de estados do sistema VCAI original
$x$	vetor de estados do sistema VCAI após a mudança de variáveis
$\bar{u}$	vetor da entrada de controle do sistema VCAI
$u$ ou $v$	vetor da variação da entrada de controle de um sistema VCAI
$u^*$	ação de controle ótima
$u^*(\cdot)$	política de controle ótima
$\tilde{u}^*(\cdot)$	aproximação da política ótima
$\epsilon_k, \varepsilon_k, \hat{\epsilon}_k$	vetores de variáveis aleatórias
$\bar{C}_k(\cdot)$ e $C_k(\cdot)$	funções custo para um dado instante $k$
$A, B, \bar{A}, \bar{B}$	matrizes da porção linear de um sistema VCAI
$\bar{\Sigma}, \Sigma, \hat{\Sigma}$	matrizes semi-definida positivas da porção estocástica de um sistema VCAI
$\phi_i(\cdot)$	função de magnitude
$J(\cdot, \cdot)$	função custo
$J_k(\cdot, \cdot)$	custo para um dado instante $k$ da programação dinâmica
$J_k^*(\cdot)$	custo acumulado até o instante $k$
$J^*(\cdot)$	custo mínimo
$(\cdot)^T$	indica o operador transposição em matrizes
$diag(\cdot)$	indica uma matriz diagonal
$tr(A)$	notação para representar o traço da matriz $A$
$\nabla$	operador gradiente
$x^i$	$i$ -ésima componente de um vetor $x$
$\nabla_{x^i}$	indica derivada parcial na componente $x^i$
$\partial$	gradiente generalizado
$\partial_{x^i}$	gradiente generalizado na componente $x^i$
$\alpha$	fator de desconto
$T$	operador da programação dinâmica
$e_i$	$i$ -ésimo vetor canônico no $\mathbb{R}^n$ ou $\mathbb{R}^m$
$\mathcal{R}_i$	região no espaço de estados
$e^{\max}$	erro máximo de aproximação
$\kappa(\cdot)$	função de relevância
$\pi_k$	inflação em um instante $k$
$y_k$	hiato do produto em um instante $k$
$i_k$	taxa de juros em um instantes $k$

# Sumário

<b>1</b>	<b>Introdução</b>	<b>1</b>
1.1	Incerteza e Tomada de Decisão . . . . .	1
1.1.1	Banco Central . . . . .	2
1.1.2	Medicina . . . . .	2
1.2	O Problema Geral . . . . .	3
1.3	Contribuições deste Trabalho . . . . .	3
1.4	Estrutura da Dissertação . . . . .	4
<b>2</b>	<b>Construção do Modelo e Resultados Úteis</b>	<b>6</b>
2.1	Construção do Modelo - Caso Mono Entrada . . . . .	6
2.2	Construção do Modelo - Caso Geral . . . . .	7
2.3	Resultados Úteis . . . . .	9
2.3.1	Resultado da Análise de Funções Convexas . . . . .	9
2.3.2	Resultados da Teoria de Controle Ótimo . . . . .	12
<b>3</b>	<b>Resultados Principais</b>	<b>15</b>
3.1	Desafios para Obter a Política Ótima . . . . .	16
3.2	Caracterização da Convexidade . . . . .	17
3.3	Quando a Não-Variação é Ótima . . . . .	21
3.4	A Programação Dinâmica . . . . .	24
3.4.1	Um Exemplo Bidimensional . . . . .	25
<b>4</b>	<b>Desafios Numéricos</b>	<b>27</b>
4.1	Programação Dinâmica Aproximada . . . . .	28
4.2	O Caso Quadrático Mono-Entrada . . . . .	30
4.2.1	Aproximação por Mínimos Quadrados . . . . .	32

4.2.2	Aproximação por Truncamento . . . . .	33
4.3	Programação Dinâmica Linear . . . . .	34
4.3.1	Exemplo escalar . . . . .	36
<b>5</b>	<b>Aplicação em um problema macroeconômico</b>	<b>39</b>
5.1	O Modelo . . . . .	39
5.1.1	O Modelo no Espaço de Estados . . . . .	42
5.2	Experimentos Numéricos . . . . .	43
5.2.1	LQR versus VCAI . . . . .	43
5.2.2	Comparação com Dados Reais . . . . .	45
<b>6</b>	<b>Conclusões e Perspectivas</b>	<b>47</b>
	<b>Bibliografia</b>	<b>50</b>

# Introdução

## 1.1 Incerteza e Tomada de Decisão

O Brasil é um país de heurísticas. Soluções para problemas enfrentados por empresas e pelo governo são definidas de forma “ad-hoc”, ignorando as conseqüências a longo prazo e a existência de ferramentas adequadas para a tomada de decisão em cenários com incerteza. Semelhante à maioria dos países em desenvolvimento, o Brasil é marcado por desigualdade social e pelo uso ineficiente de recursos naturais. Desta forma, mesmo possuindo uma das maiores áreas cultiváveis do mundo, parcela significativa da população passa fome e sofre de má-nutrição por causa de problemas de logística e gerência. Além disso, várias políticas estão sendo continuamente questionadas devido a um sistema de distribuição de recursos arcaico e sub-ótimo.

De modo geral, incerteza é endêmica no Brasil, afetando diversas decisões na economia e, em particular, as decisões referentes aos investimentos. É difícil para um investidor lidar com mudanças no ambiente econômico, assim como é difícil para um empreendedor crescer em um ambiente complexo e volátil gerado por uma economia frágil e serviços de logística caros e ineficientes. Todos estes problemas afligem nossa sociedade, e a maioria dos jovens brasileiros acaba possuindo oportunidades profissionais e sociais limitadas, não por causa da falta de habilidade ou potencial, mas devido a fatores além do seu alcance e controle.

Dentro deste contexto, é fundamental estudar o processo de tomada de decisão em cenários dominados por incerteza e volatilidade. Sendo assim, este trabalho faz parte de uma linha de pesquisa desenvolvida dentro da Faculdade de Engenharia Elétrica e da Computação da Universidade Estadual de Campinas que busca estudar o processo decisório e otimização dinâmica, particularmente sob incerteza.

Na última década, diversos resultados da Teoria de Controle foram utilizados em aplicações não tradicionais, como em Economia [9], Medicina [14] e Biologia [13], [12]. Estas

áreas servem como motivação para este trabalho, fornecendo novos desafios e situações onde a tomada de decisão e o controle de sistemas ocorrem em ambientes não-lineares e incertos.

Com o objetivo de apresentar o problema estudado neste trabalho de forma intuitiva, serão analisadas duas situações práticas em que decisões devem ser tomadas na presença de risco: a condução da política monetária por parte do Banco Central e a definição da dosagem ótima de medicamentos a ser aplicada a um paciente.

### 1.1.1 Banco Central

Considere o problema da condução da política monetária pelo Banco Central (BC) em um regime de metas inflacionárias [15], [19]. Ao definir um aumento ou redução na taxa de juros básica da Economia, o BC se depara com um contexto de incerteza em relação a possível reação dos agentes econômicos. Caso haja uma variação muito grande na taxa de juros básica da economia, o BC corre o risco de provocar um excesso de otimismo ou pessimismo que pode resultar em conseqüências indesejadas e imprevisíveis como o aumento da inflação ou redução do Produto Interno Bruto (PIB). Por outro lado, caso haja uma variação muito pequena da taxa de juros, os agentes econômicos podem começar a duvidar da competência e da credibilidade do compromisso do BC em atingir a meta inflacionária. Este é um exemplo no qual a variação da política pelo tomador de decisões (no caso, o BC) pode levar a um aumento da incerteza sobre o sistema.

Note que, nesse caso, a incerteza decorre da dificuldade de desenvolver um modelo dinâmico adequado para descrever a economia e a reação dos agentes econômicos às políticas a serem implementadas devido à impossibilidade de utilizar técnicas tradicionais de identificação. As únicas informações disponíveis para a criação de um modelo são as políticas anteriores do BC e um número limitado de estados da economia. Logo, a incerteza do modelo não está associada apenas ao estado do sistema, mas também à política adotada pelo BC. Quanto maior a variação da política, menor será a confiança sobre o modelo. Desta forma, dada a natureza desta incerteza, verifica-se que ela não pode ser adequadamente descrita mediante modelos de saltos ou modelos tradicionais de incerteza aditiva ou multiplicativa.

### 1.1.2 Medicina

O mesmo tipo de incerteza pode ser observado no problema da definição da dosagem de medicamentos para um paciente. A resposta de um paciente à variações de um medicamento é não-linear e incerta [18], [11], [10]. Mudanças grandes de dosagens podem levar a conseqüências incertas, enquanto variações pequenas podem não surtir efeito algum. Neste caso, também é impossível utilizar técnicas tradicionais para construir um modelo da resposta de

um indivíduo a variações de um medicamento.

## 1.2 O Problema Geral

Considere um sistema onde a entrada de controle é escolhida com informação completa do estado. Entretanto, suponha que não temos acesso a um modelo dinâmico que descreve adequadamente este sistema. Suponha, também, que como nos exemplos em economia e medicina, não é possível utilizar técnicas tradicionais de identificação para obter modelos precisos deste sistema. Esta dificuldade de identificação pode ocorrer em várias situações, particularmente em sistemas utilizados em situações críticas, onde não é possível fazer experimentos (como no caso do corpo humano ou a economia nacional).

Nestes casos, é apenas possível obter um modelo do sistema através de dados históricos, que normalmente oferecem uma quantidade limitada de controles e estados. Logo, o modelo linear obtido para descrever o sistema será preciso apenas para uma faixa de valores do estado e do controle. Além disso, devido a não-linearidades ou outras dinâmicas indeterminadas do sistema, grandes variações do controle poderão levar o sistema para regiões onde o erro de aproximação torna-se muito grande. Neste caso, o erro de aproximação corresponderá à incerteza gerada por variações da magnitude da política. Chamaremos esta classe de sistemas de **Sistemas Estocásticos nos quais a Variação da Magnitude do Controle Aumenta a Incerteza (VCAI)**

A Figura 1.2 ilustra o possível comportamento de um sistema VCAI bidimensional. Suponha que, para o sistema da figura, temos como objetivo levar o estado para a origem. Observe que a ação de controle que leva, em média, o estado para a origem não será necessariamente a ação ótima, pois pode gerar um aumento muito grande da incerteza sobre a posição do sistema, indicada pela região circular maior. Como mostrado na figura, o controle ótimo pode corresponder a uma ação intermediária que, apesar de não levar o sistema para a origem, não acarretará um aumento significativo da incerteza sobre o comportamento do sistema.

## 1.3 Contribuições deste Trabalho

O estudo deste tipo de sistemas estocásticos VCAI não possui precedentes na literatura. Neste sentido, este trabalho objetiva propiciar diversas contribuições para o estudo desta classe de sistemas. As principais delas são:

- Um modelo geral para descrição de sistemas VCAI;

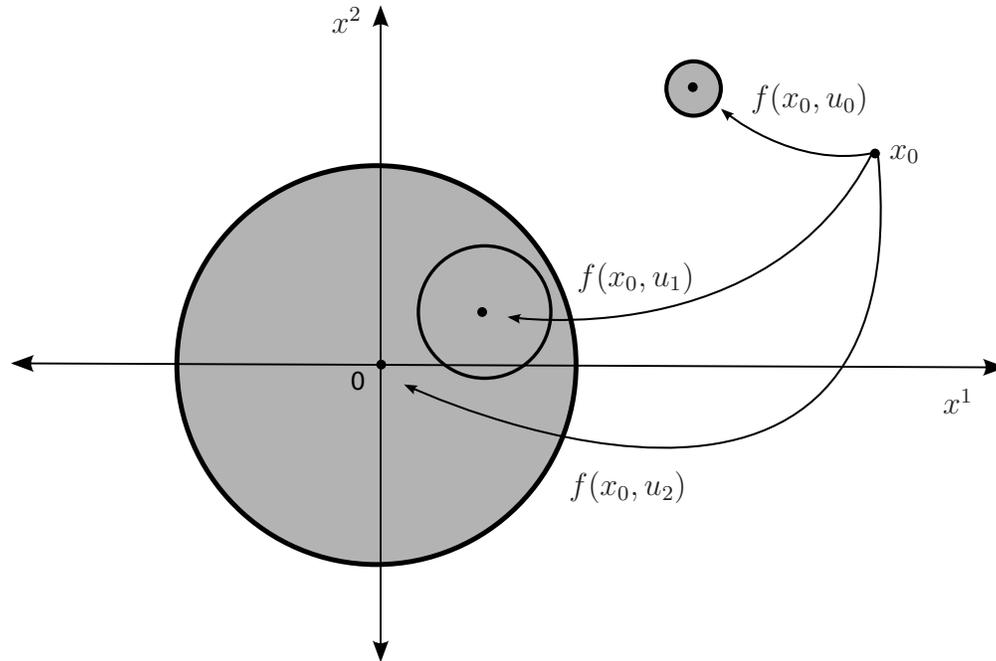


Figura 1.1: Exemplo de um sistema VCAI bidimensional com equação de estado  $x_1 = f(x, u)$ . No instante  $k = 0$  o sistema está em  $x_0$  e a entrada atual é  $u_0$ . Além disso,  $|u_1| < |u_2|$ . Os círculos representam uma região onde o estado estará localizado com probabilidade  $p$ .

- Descrição matemática da política de controle ótima para sistemas VCAI avaliados através de uma função de custo convexa;
- Demonstração da existência de uma *região de não-variação* para esta classe de sistemas, formando um novo paradigma sobre modelos para incerteza.
- Um estudo dos desafios numéricos para a implementação de algoritmos para obtenção da política ótima;
- Proposta de estratégias numéricas para obter a política ótima em sistemas VCAI;
- Aplicação deste tipo de sistema em problemas de condução da política monetária pelo Banco Central. Será observado que sistemas VCAI caracterizam este processo decisório de forma mais adequada do que os modelos existentes na literatura.

## 1.4 Estrutura da Dissertação

No próximo capítulo, serão introduzidas algumas ferramentas matemáticas que serão úteis na resolução deste problema de otimização. Feito isso, no Capítulo 3 a política de controle ótima para sistemas VCAI será caracterizada analiticamente. Além disso, será apresentada

---

a existência de uma *região de não-variação* para esta classe de sistemas, formando um novo paradigma sobre modelos para incerteza. No Capítulo 4, algumas estratégias para obter numericamente a política ótima serão discutidas e o erro destas aproximações serão calculados. No Capítulo 5, o modelo será utilizado para descrever o problema de tomada de decisão enfrentado pelo BC e o seu desempenho será comparado com modelos existentes na literatura. Finalmente, o Capítulo 6 apresentará uma breve conclusão e uma agenda para trabalhos futuros será proposta.

## Construção do Modelo e Resultados Úteis

O primeiro passo no estudo de sistemas VCAI é a criação de um modelo dinâmico que descreve adequadamente o seu comportamento. Este modelo deverá ser capaz de representar o aumento da incerteza gerada pela variação da ação de controle e, ao mesmo tempo, deverá ser simples o suficiente para ser tratado analiticamente. Desta forma, neste capítulo será proposto um modelo geral para sistemas VCAI. A equação de estado irá possuir dois componentes: uma parte determinística linear e uma porção estocástica que descreverá o impacto da variação do controle sobre a incerteza na posição do estado. Será apresentado, primeiramente, um modelo para o caso mono entrada e em seqüência um modelo geral. Além disso, o problema de controle de sistemas VCAI será colocado na forma de um problema de controle estocástico. Feito isso, serão introduzidas algumas ferramentas matemáticas e resultados teóricos que serão úteis para a solução deste problema de otimização.

### 2.1 Construção do Modelo - Caso Mono Entrada

Considere um sistema cuja dinâmica em tempo discreto é dada pela seguinte equação de estado:

$$\bar{x}_{k+1} = \bar{A}_k \bar{x}_k + \bar{b}_k \bar{u}_k + \epsilon_k \quad (2.1)$$

onde  $\bar{A} \in \mathbb{R}^{n \times n}$ ,  $\bar{b} \in \mathbb{R}^{n \times 1}$ ,  $\bar{u}_k \in \mathbb{R}^m$  e  $\epsilon_k$  é um vetor de variáveis aleatórias que representa a incerteza sobre o sistema. Observe que o sistema possui uma componente determinística linear e uma componente aleatória. Suponha que, a cada instante  $k$ , um tomador de decisão tem acesso a um controle,  $u_k$ , que é definido supondo observação completa do estado. Além disso, suponha que a variação do controle entre um instante de tempo e outro é simplesmente

$$v_k = \bar{u}_k - \bar{u}_{k-1}.$$

Conforme discutido na seção anterior, o foco deste trabalho são sistemas onde a variação do controle aumenta a incerteza sobre a sua dinâmica. Desta forma, a seqüência  $\{\epsilon_k\}$  será modelada como um processo estocástico e, neste caso, será uma função do *valor absoluto das variações de controle*  $|v_k|$ , representando a incerteza gerada pela variação da magnitude do controle. De forma a criar essa relação no modelo, será assumido que a seqüência de ruídos é modulada por  $|v_k|$  através da relação:

$$\epsilon_k = (\bar{\sigma}_k + \sigma_k |v_k|) \bar{\epsilon}_k, \quad \bar{\sigma}_k, \sigma \in \mathbb{R}^+, \quad \bar{\epsilon}_k \sim N(0, I_{n \times n})$$

em que  $\{\bar{\epsilon}_k\}$  é uma seqüência i.i.d. e  $I_{n \times n}$  é a matriz identidade de ordem  $n$ . Para uma janela de tempo  $k \in [0, N]$ , o desempenho do sistema será avaliado através da *função custo*:

$$\bar{J}(\bar{x}_0, \bar{\pi}) = E \left[ \sum_{k=0}^{N-1} \bar{C}_k(\bar{x}_k, \bar{u}_k) + \bar{C}_N(\bar{x}_N) \right] \quad (2.2)$$

em que  $\bar{\pi} = \{\bar{u}_0(x), \bar{u}_1(x), \dots, \bar{u}_{N-1}(x)\}$  é uma seqüência de funções,  $\bar{u}_k : \mathbb{R} \rightarrow \mathbb{R}$  que representa a política de controle,  $\bar{C}_k$  é uma função não-negativa e convexa em  $\bar{x}_k$  e  $\bar{u}_k$  para todo  $k \in [0, N-1]$ . Além disso, o custo terminal  $\bar{C}_N(\bar{x}_N)$  é uma função não-negativa e convexa.

## 2.2 Construção do Modelo - Caso Geral

O modelo mono entrada pode ser expandido para um caso mais geral considerando um sistema com múltiplas entradas. O modelo geral para sistemas VCAI será:

$$\bar{x}_{k+1}(\bar{x}_k, \bar{u}_k) = \bar{A}_k \bar{x}_k + \bar{B}_k \bar{u}_k + (\dot{\Sigma}_k + \sum_{i=1}^m \phi_i(v_k^i) \bar{\Sigma}_k^i) \bar{\epsilon}_k$$

em que  $\dot{\Sigma}_k, \bar{\Sigma}_k^i \in \mathbb{R}^{n \times n}$  são matrizes simétricas e semi-definidas positivas,  $\bar{\epsilon}_k$  é um vetor de variáveis aleatórias, e a  $i$ -ésima componente de  $v$  é denotada por  $v^i$ , ou seja,

$$v = [v^1 \quad v^2 \quad \dots \quad v^m]^T$$

e  $\phi_i : \mathbb{R} \rightarrow \mathbb{R}$  é uma função positiva, convexa, Lipschitz e não diferenciável na origem que chamaremos de **função de magnitude**, por ser a função que representa o impacto da variação da magnitude do controle sobre a incerteza do sistema. Exemplos de funções  $\phi_i$  são:

$$\phi_i(v^i) = |v^i|, \quad \phi_i(v^i) = \begin{cases} v^i, & v^i > 0 \\ 0, & v^i < 0 \end{cases}, \quad \phi_i(v^i) = \begin{cases} (v^i + c)^2 - c^2, & v^i > 0 \\ (v^i - c)^2 - c^2, & v^i < 0 \end{cases}$$

Novamente, será suposto que o desempenho do sistema é avaliado através da função custo dada pela Equação 2.2. Logo, deseja-se resolver o seguinte problema de controle estocástico

$$\begin{aligned} \min_{\bar{\pi} \in \bar{\Pi}} \bar{J}(\bar{x}_0, \bar{\pi}) &= E \left[ \sum_{k=0}^{N-1} \bar{C}_k(\bar{x}_k, \bar{u}_k) + \bar{C}_N(\bar{x}_N) \right] \\ \text{s.a. } \bar{x}_{k+1}(\bar{x}_k, \bar{u}_k) &= \bar{A}_k \bar{x}_k + \bar{B}_k \bar{u}_k + (\dot{\bar{\Sigma}}_k + \sum_{i=1}^m \phi_i(v_k^i) \bar{\Sigma}_k^i) \varepsilon_k \end{aligned}$$

onde  $\bar{\Pi}$  é o espaço de funções que contém todas as políticas admissíveis.

Observe que este problema de otimização depende tanto do controle,  $\bar{u}$ , quanto da variação do controle,  $v$ . De forma a simplificá-lo, definimos o sistema aumentado:

$$x_{k+1}(x_k, u_k) = A_k x_k + B_k u_k + (\bar{\Sigma}_k + \sum_{i=1}^m \phi_i(u_k^i) \Sigma_i) \varepsilon_k \quad (2.3)$$

em que  $\varepsilon_k$  é um vetor de variáveis aleatórias,  $u_k = v_k$  e

$$x_k := \begin{bmatrix} \bar{x}_k \\ \bar{u}_{k-1} \end{bmatrix}, \quad A_k := \begin{bmatrix} \bar{A}_k & \bar{B}_k \\ \mathbf{0}_{m \times n} & \mathbf{I}_{m \times m} \end{bmatrix}, \quad B_k := \begin{bmatrix} \bar{B}_k \\ \mathbf{I}_{m \times m} \end{bmatrix}, \quad \bar{\Sigma}_k := \begin{bmatrix} \dot{\bar{\Sigma}}_k \\ \mathbf{0}_{m \times n} \end{bmatrix}, \quad \Sigma_k^i := \begin{bmatrix} \bar{\Sigma}_k^i \\ \mathbf{0}_{m \times n} \end{bmatrix}$$

sendo  $\mathbf{0}$  uma matriz de zeros e  $\mathbf{I}$  a matriz identidade.

Através de uma mudança de variáveis em (2.2) as funções  $C_k(x_k, u_k) = \bar{C}_k(\bar{x}_k, \bar{u}_k)$  e  $C_N(x_N) = \bar{C}_N(\bar{x}_N)$  são obtidas, gerando a função custo convexa equivalente:

$$J(x_0, \pi) = E \left[ \sum_{k=0}^{N-1} C_k(x_k, v_k) + C_N(x_N) \right] \quad (2.4)$$

em que  $\pi = \{u_0(x), u_1(x), \dots, u_{N-1}(x)\}$  é uma seqüência de funções que caracteriza a política de variação do controle.

Com o estado aumentado, a entrada do sistema passa a ser a *variação do controle*. Logo, o problema a ser resolvido é

$$\begin{aligned} \min_{\pi \in \Pi} J(x_0, \pi) &= E \left[ \sum_{k=0}^{N-1} C_k(x_k, u_k) + C_N(x_N) \right] \\ \text{s.a. } x_{k+1}(x_k, u_k) &= A_k x_k + B_k u_k + (\bar{\Sigma}_k + \sum_{i=1}^m \phi_i(u_k^i) \Sigma_k^i) \varepsilon_k \\ x_k \in S \subset \mathbb{R}^n, \quad u_k \in U \subset \mathbb{R}^m, \quad \varepsilon_k \in D \subset \mathbb{R}^n, \quad x_0 = x \in \mathbb{R}^n \end{aligned}$$

em que  $\Pi$  é o espaço de funções que contém todas as políticas admissíveis,  $S$  é o espaço de estados,  $U$  é o espaço de controle,  $D$  é o espaço que contém todos os distúrbios e o estado inicial,  $x_0$ , é conhecido. Como  $\varepsilon$  possui distribuição normal, para uma maior simplicidade, será suposto que  $S = D = \mathbb{R}^n$ .

## 2.3 Resultados Úteis

Nos seguintes resultados, será assumido que as funções custo  $C_k(\cdot, \cdot)$  para cada  $k$  e  $C_N(\cdot)$  são funções convexas e, portanto, serão consideradas algumas propriedades desta classe de funções que serão úteis.

As demonstrações das próximas proposições podem ser encontradas nas referências indicadas.

### 2.3.1 Resultado da Análise de Funções Convexas

**Proposição 2.3.1** ([4]). *Suponha  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  diferenciável, isto é,  $\nabla f$  existe em cada ponto do domínio de  $f$ , que é aberto. Logo,  $f$  será convexa se e somente se  $\text{dom}(f)$  for convexo e*

$$f(x) - f(x_0) \leq \nabla f(x_0)^T(x - x_0) \quad (2.5)$$

**Proposição 2.3.2** ([4]). *Suponha  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $A \in \mathbb{R}^{n \times n}$  e  $b \in \mathbb{R}^n$ . Se  $f$  for convexa, a função  $g : \mathbb{R}^n \rightarrow \mathbb{R}$  definida como*

$$g(x) = f(Ax + b)$$

*também será convexa.*

**Definição 2.3.1** ([7]). *Seja  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  Lipschitz na vizinhança de  $x \in \mathbb{R}^n$ . Além disso, seja  $\Omega$  qualquer conjunto de medida nula em  $\mathbb{R}^n$ , e seja  $\Omega_f$  o conjunto de pontos em  $\mathbb{R}^n$  onde  $f$  não é diferenciável. Logo, o gradiente generalizado em  $x$ , denotado por  $\partial f(x)$ , será o conjunto*

$$\partial f(x) = \text{co} \left\{ \lim_{x_i \rightarrow x} \nabla f(x_i) : x_i \notin \Omega, x_i \notin \Omega_f \right\}$$

onde “co” significa “fecho convexo”.

**Proposição 2.3.3** ([7]). *Seja  $f$  Lipschitz próxima de cada ponto em um subconjunto aberto e convexo  $U \subset \mathbb{R}^n$ . Logo,  $f$  será convexa em  $U$  se e somente se  $\partial f$  for monótona em  $U$ , i.e.,*

$$\langle \zeta - \tilde{\zeta}, x - \tilde{x} \rangle \geq 0, \quad \forall x, \tilde{x} \in U, \zeta \in \partial f(x), \tilde{\zeta} \in \partial f(\tilde{x})$$

**Proposição 2.3.4** ([4]). *Sejam dois conjuntos não vazios e convexas  $U \subset \mathbb{R}^n$  e  $V \subset \mathbb{R}^m$ . Logo, se  $f : U \times V \rightarrow \mathbb{R}$  for convexa em  $V$  e  $U$ , então a função  $g$  dada por*

$$g(x) = \inf_{y \in V} f(x, y)$$

*será convexa em  $x$ , caso  $g(x) > -\infty$  para todo  $x \in U$ .*

**Proposição 2.3.5** (Regra da Cadeia [7]). *Seja  $f = g \circ h$ , onde  $h : \mathbb{R}^m \rightarrow \mathbb{R}^n$  e  $g : \mathbb{R}^n \rightarrow \mathbb{R}$ . Logo,*

$$\partial f(x) \subset \text{co} \{ \partial g(h(x)) \partial h(x) \}.$$

*Se  $f$  for convexo, a inclusão torna-se uma igualdade. Note que  $\partial h(x) \in \mathbb{R}^{n \times m}$  e  $\partial g(h(x)) \in \mathbb{R}^{1 \times n}$ .*

**Proposição 2.3.6** (Extremo Local [7]). *Se  $f : \mathbb{R} \rightarrow \mathbb{R}^n$  atingir um mínimo ou máximo local em  $x_0$ , então  $0 \in \partial f(x_0)$ .*

**Proposição 2.3.7** (Funcionais Integrais [7]). *Seja uma função Lipschitz e mensurável  $g : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}$ , tal que*

$$|g(x_1, y) - g(x_0, y)| \leq K(y) \|x_1 - x_0\|, \forall x_1, x_0 \in \mathbb{R}^n; y \in \mathbb{R}^m$$

*Além disso, seja  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  definida como*

$$f(x) = \int_U g(x, y) dy, U \subset \mathbb{R}^m$$

*Logo*

$$\partial f(x) = \partial \int_U g(x, y) dy \subset \int_U \partial_x g(x, y) dy$$

**Proposição 2.3.8** ([4]). *Para  $x \in \mathbb{R}^n$  e  $u \in \mathbb{R}$  seja  $(x, u) \rightarrow V(x, u)$  uma função estritamente convexa, diferenciável e limitada inferiormente. Além disso, seja  $u^*(x)$  definido como*

$$u^*(x) = \arg \min_{u \in \mathbb{R}} V(x, u) \tag{2.6}$$

*Conseqüentemente, a função  $x \rightarrow V^*(x)$  definida como*

$$V^*(x) = \inf_{u \in \mathbb{R}} V(x, u) \tag{2.7}$$

*será diferenciável para todo  $x$ .*

**Proposição 2.3.9** ([6]). *Para  $x \in \mathbb{R}^n$  e  $u \in \mathbb{R}$ , seja  $(x, u) \rightarrow V(x, u)$  uma função contínua e convexa, e seja  $u^*$  e  $V^*$  dadas por (2.6) e (2.7). Logo,*

$$\partial V^*(x) = \text{co} \{ \partial_x V(x, u) : u \in u^*(x) \}$$

**Lema 2.3.1.** *Considere  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  uma função convexa e diferenciável,  $\varepsilon$  um vetor de variáveis aleatórias independentes com média nula. Além disso, suponha que as componentes de um vetor  $x$  são dadas por  $x^i$ , de modo que*

$$x = [x^1 \quad \dots \quad x^i \quad \dots \quad x^n]^T$$

Logo, sendo  $\varepsilon_0 \in \mathbb{R}^n$ ,  $\bar{\Sigma} \in \mathbb{R}^{n \times n}$  e  $x = \varepsilon_0 + \bar{\Sigma}\varepsilon$ ,

$$E [\nabla_{x^i} f(x) \varepsilon^i] \geq 0,$$

em que  $\nabla_{x^i}$  denota a derivada parcial da  $i$ -ésima componente de  $x$  e  $E[\cdot]$  representa o valor esperado da variável aleatória. Se  $f$  for estritamente convexa, a desigualdade será estrita.

*Demonstração.* Como  $f$  é uma função convexa, vale a desigualdade

$$(\nabla f(x) - \nabla f(\tilde{y}))^T (x - y) \geq 0 \quad (2.8)$$

Fazendo

$$\begin{aligned} x &= \varepsilon_0 + \bar{\Sigma}\varepsilon \\ y &= \varepsilon_0 + \bar{\Sigma}(\varepsilon - \varepsilon e_i) \end{aligned}$$

em que  $e_i$  é o vetor canônico cuja  $i$ -ésima componente vale 1 e todas as outras são nulas, obtém-se

$$(\nabla f(x) - \nabla f(y))^T (x - y) = (\nabla_{x^i} f(x) - \nabla_{x^i} f(y)) \varepsilon^i$$

De (2.8), tem-se a desigualdade

$$\nabla_{x^i} f(x) \varepsilon^i \geq \nabla_{x^i} f(y) \varepsilon^i \quad (2.9)$$

Entretanto, como nenhuma componente de  $y$  depende de  $\varepsilon^i$ , e como as componentes de  $\varepsilon$  são independentes e possuem média nula, note que

$$E [\nabla_{x^i} f(y) \varepsilon^i] = E [\nabla_{x^i} f(y)] E [\varepsilon^i] = 0$$

Logo, calculando o valor esperado nos dois lados da desigualdade (2.9),

$$E [\nabla_{x^i} f(x) \varepsilon^i] \geq 0$$

completando a prova. Observe que caso  $f$  seja estritamente convexa, a desigualdade em (2.8) será estrita e, conseqüentemente, a desigualdade acima também será estrita. □

**Lema 2.3.2.** *Considere  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  uma função convexa e diferenciável,  $\varepsilon$  um vetor de variáveis aleatórias independentes com distribuição normal e média nula, e as matrizes  $\bar{\Sigma}, \Sigma \in \mathbb{R}^{n \times n}$  simétricas semi-definidas positivas. Logo,*

$$E[\nabla^T f(\varepsilon_0 + \bar{\Sigma}\varepsilon) \Sigma \varepsilon] \geq 0, \quad \forall \varepsilon_0 \in \mathbb{R}^n$$

Além disso, se  $f$  for estritamente convexa e  $\Sigma$  possuir ao menos um autovalor positivo, a desigualdade acima será estrita.

*Demonstração.* Como  $\Sigma$  é uma matriz simétrica e semi-definida positiva, existe uma matriz ortogonal  $T$  e uma matriz diagonal semi-definida positiva  $\Lambda$  tal que

$$\Sigma = T\Lambda T^T$$

Desta forma, tem-se

$$E[\nabla^T f(\varepsilon_0 + \bar{\Sigma}\varepsilon)\Sigma\varepsilon] = E[\nabla^T f(\varepsilon_0 + \bar{\Sigma}\varepsilon)T\Lambda T^T\varepsilon] \quad (2.10)$$

Fazendo uma mudança de variável  $y = Tx$ , considere uma função  $h : \mathbb{R}^n \rightarrow \mathbb{R}$  de modo que

$$f(x) = f(T^T Tx) = f(T^T y) = h(y)$$

Logo,

$$\nabla h(y) = T^T \nabla f(T^T y) = T^T \nabla f(x)$$

Pela Proposição 2.3.2,  $h$  também será uma função convexa. Além disso, observe que

$$E[\nabla^T f(\varepsilon_0 + \bar{\Sigma}\varepsilon)\Sigma\varepsilon] = E[\nabla^T h(T^T \varepsilon_0 + T^T \bar{\Sigma}\varepsilon)\Lambda T^T \varepsilon]$$

Fazendo uma mudança de variável  $\varepsilon = T^T w$  e denotando por  $\lambda_i$ , o  $i$ -ésimo autovalor de  $\Lambda$ , (2.10) torna-se

$$\begin{aligned} E[\nabla^T f(\varepsilon_0 + \bar{\Sigma}\varepsilon)\Sigma\varepsilon] &= E[\nabla^T h(T^T \varepsilon_0 + T^T \bar{\Sigma} T w)\Lambda w] \\ &= \sum_{i=1}^n \lambda_i E[\nabla_{x^i} h(T^T \varepsilon_0 + T^T \bar{\Sigma} T w)w^i] \end{aligned} \quad (2.11)$$

Observe que como  $T$  é uma matriz ortogonal, se  $\varepsilon$  for um vetor de variáveis aleatórias gaussianas independentes,  $w$  também será. Utilizando o Lema 2.3.1 e o fato que  $\Lambda$  é semi-definida positiva, note que todos os termos do somatório são não-negativos. Conseqüentemente,

$$E[\nabla^T f(\varepsilon_0 + \bar{\Sigma}\varepsilon)\Sigma\varepsilon] \geq 0$$

Se  $f$  for estritamente convexa e  $\Sigma$  possuir ao menos um autovalor positivo, o somatório em (2.11) será positivo e o a desigualdade acima será estrita. □

### 2.3.2 Resultados da Teoria de Controle Ótimo

Os resultados a seguir apresentam a estratégia clássica para obter a política de controle ótima em sistemas dinâmicos avaliados através de um funcional de custo. Esta estratégia, desenvolvida por Richard Bellman na década de 50, é chamada de *Programação Dinâmica* e é a pedra fundamental da teoria de controle estocástico.

**Proposição 2.3.10** (Programação Dinâmica [2]). *Considere o problema de controle estocástico, no qual  $\Pi$  é o espaço de funções que contém todas políticas admissíveis e o estado inicial  $x_0$  é dado.*

$$\begin{aligned} \min_{\pi \in \Pi} J(x_0, \pi) &= E \left[ \sum_{k=0}^{N-1} C_k(x_k, u_k) + C_N(x_N) \right] \\ \text{s.a. } x_{k+1}(x_k, u_k) &= f_k(x_k, u_k, \varepsilon_k) \\ x_k &\in S \subset \mathbb{R}^n, u_k \in U \subset \mathbb{R}^m, \varepsilon_k \in D \subset \mathbb{R}^n \end{aligned}$$

O controle ótimo e o custo mínimo  $J^*(x_0)$  para um estado inicial  $x_0$  pode ser obtido através do seguinte algoritmo, que retrocede do instante  $k = N - 1$  até  $k = 0$  :

1. Defina  $k = N - 1$  e  $J_N^* = C_N(x_N)$ ;

2. Defina  $J_k(x, u)$  como

$$J_k(x, u) = E[C_k(x, u) + J_{k+1}(f_k(x, u, \varepsilon_k))]$$

3. Calcule, para todo  $x \in \mathbb{R}^n$ ,

$$J_k^*(x) = \min_{u_k} E[C_k(x_k, u_k) + J_{k+1}(f_k(x_k, u_k, \varepsilon_k))]$$

A função  $u_k^*(x)$  que atender a condição

$$J_k^*(x) = J_k(x, v_x) \leq J_k(x, u), \forall x \in \mathbb{R}^n, u \in \mathbb{R}^m$$

tal que  $u_k^*(x) = v_x, \forall x \in \mathbb{R}^n$ , é a **política ótima**. A função  $J_k^*$  a cada instante  $k$  é o **custo acumulado**.

4. Se  $k \neq 0$ , faça  $k = k - 1$  e retorne para o passo 2. Caso contrário, pare.

Desta forma, a política ótima  $u_k^*(x)$  é obtida para cada  $0 \leq k < N$  e, em particular,  $J^*(x) = J_0(x, u_0^*(x))$ .

Em diversas situações práticas, deseja-se otimizar o desempenho de um sistema para um horizonte de tempo infinito com custo descontado. Neste tipo de cenário, temos o seguinte problema de otimização:

$$\begin{aligned} \min_{\pi \in \Pi} J(x_0, \pi) &= \lim_{N \rightarrow \infty} E \left[ \sum_{k=0}^N \alpha^k C(x_k, u_k) \right] \\ \text{s.a. } x_{k+1}(x_k, u_k) &= f(x_k, u_k, \varepsilon_k), \quad 0 < \alpha < 1 \end{aligned} \tag{2.12}$$

em que  $\alpha$  é o fator de desconto. Além disso, será considerado que, para  $M \in \mathbb{R}$ ,  $C$  é limitado para qualquer subconjunto compacto de  $S \times U$ , ou seja,

$$|C(x, u)| < M \quad \forall (x, u) \in S \times U$$

De forma a simplificar a notação, o operador da programação dinâmica  $T$  sobre uma função  $J : S \rightarrow \mathbb{R}$ , sendo  $S \subset \mathbb{R}^n$  um conjunto compacto e convexo, é definido como

$$(TJ)(x) = \min_{u \in U(x)} E\{C(x, u) + \alpha J(f(x, u))\}, \quad x \in S. \quad (2.13)$$

Um resultado fundamental para a solução do problema exposto na Equação 2.12 é a garantia da convergência do algoritmo da Programação Dinâmica, mesmo para um horizonte infinito. Este resultado está exposto na proposição abaixo.

**Proposição 2.3.11** ([1]). *Para qualquer função limitada  $J : S \rightarrow \mathbb{R}$ , a função de custo ótima,  $J^*$ , do problema 2.12 satisfaz:*

$$J^*(x) = \lim_{N \rightarrow \infty} (T^N J)(x), \quad \forall x \in S$$

A partir deste resultado, percebe-se que o custo mínimo é a solução de uma equação do ponto fixo, chamada de *Equação de Bellman*. A proposição a seguir trata deste fato.

**Proposição 2.3.12** ([1]). *A função de custo ótima,  $J^*$ , do problema 2.12 satisfaz:*

$$J^*(x) = TJ^* \quad (2.14)$$

além disso,  $J^*$ , é a única solução desta equação. Esta equação é chamada de **Equação de Bellman**.

Conseqüentemente, para um horizonte infinito de otimização, basta resolver a Equação 2.14 para determinar o custo ótimo do problema 2.12. Inclusive, em algumas situações, particularmente quando deseja-se encontrar a política ótima de um sistema invariante no tempo para um horizonte grande de otimização, resolver numericamente a Equação de Bellman pode exigir um custo computacional menor do que utilizar a programação dinâmica. Sendo assim, dependendo do sistema e do horizonte de otimização, pode ser mais eficiente aproximar o problema por um problema de horizonte infinito e fazer uma escolha adequada da atenuação  $\alpha$ .

## Resultados Principais

Uma vez construído o modelo para sistemas VCAI e definido o problema de controle estocástico a ser resolvido, o próximo passo é obter matematicamente a política ótima. O procedimento tradicional neste caso é utilizar Programação Dinâmica (PD) que, apesar de ser uma ferramenta teórica muito poderosa para resolução de problemas de otimização dinâmica, possui um custo computacional que cresce exponencialmente com o horizonte de otimização e o número de estados analisados. Este fenômeno é conhecido como a “Maldição da Dimensionalidade”[2].

Mais problemas surgem quando o espaço de estados é contínuo, ou seja, consiste em um número infinito não enumerável de elementos. Como é impossível calcular a política ótima para cada estado, é necessário obter funções que mapeiam o valor do estado para o controle ótimo. Entretanto, para qualquer problema que fuja do Regulador Linear Quadrático, dificilmente será possível obter soluções analíticas, sendo necessário utilizar aproximações numéricas. Neste sentido, para determinar a política ótima de controle para um sistema estocástico, é necessário trabalhar em duas frentes distintas. A primeira frente é teórica, consistindo em analisar detalhadamente o problema a ser resolvido de forma a identificar características da função custo que se preservam ao longo do horizonte de otimização. A segunda frente é computacional, onde estas características são utilizadas para obter algoritmos para calcular numericamente a política ótima. É interessante notar que ambas as frentes apresentam desafios muito intrincados e vêm sendo largamente estudadas na literatura. Saber relacionar a frente teórica com a frente computacional é fundamental para obter implementações eficientes de controladores e extrair características particulares de um problema de controle estocástico.

Este capítulo tratará da frente teórica, de forma a obter regularidades do problema de controle ótimo de sistemas VCAI. Primeiramente, será feita uma breve discussão dos desafios da utilização da PD neste problema. Em sequência, serão obtidas algumas características do

custo acumulado que podem ser úteis para o desenvolvimento de estratégias numéricas para calcular a política ótima. Os resultados obtidos serão sintetizados em um teorema para a programação dinâmica em sistemas VCAI. Finalmente, será demonstrada a existência de uma “região de não-variação”, ou seja, uma região no espaço de estados onde a melhor estratégia de controle é de não variação da política.

### 3.1 Desafios para Obter a Política Ótima

Conforme visto na Seção 2.2, temos que o nosso problema de controle ótimo é:

$$\begin{aligned} \min_{\pi \in \Pi} J(x_0, \pi) &= E \left[ \sum_{k=0}^{N-1} C_k(x_k, u_k) + C_N(x_N) \right] \\ \text{s.a. } x_{k+1}(x_k, u_k) &= A_k x_k + B_k u_k + (\bar{\Sigma}_k + \sum_{i=1}^m \phi_i(u_k^i) \Sigma_k^i) \varepsilon_k \\ u_k &\in \mathbb{R}^m, \quad x_k \in \mathbb{R}^n, \quad x_0 = x \in \mathbb{R}^n \end{aligned}$$

em que  $x_k \in \mathbb{R}^n$ ,  $\varepsilon \sim N(0, I_{n \times n})$  e  $u_k^i$  denota a  $i$ -ésima componente de  $u_k$ , ou seja,

$$u_k = [u^1 \quad u^2 \quad \dots \quad u^m]^T,$$

Além disso, para todo  $k, i \in \mathbb{N}$   $A_k, \bar{\Sigma}_k, \Sigma_k^i \in \mathbb{R}^{n \times n}$ ,  $B^k \in \mathbb{R}^{m \times m}$  e  $\phi^i : \mathbb{R}^m \rightarrow \mathbb{R}$  é uma função contínua, convexa e não diferenciável na origem.

Para obter a política ótima, algumas questões devem ser respondidas. As principais delas são:

- Como lidar com o fato que a parte estocástica do nosso sistema é não-linear?
- Se a função custo for convexa, o custo mínimo também será convexo?
- O custo mínimo será diferenciável mesmo  $\phi^i(u_k^i)$  sendo não diferenciável na origem?
- Existirá uma política ótima de controle?
- Que outras características podemos extrair da solução ótima?

No restante deste capítulo, estas perguntas serão respondidas e será apresentado um exemplo simples para ilustrar os resultados obtidos. De modo a simplificar a notação, serão definidas as funções  $V : \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}$  e  $V^* : \mathbb{R}^n \rightarrow \mathbb{R}$  como

$$V(x, u) = C(x, u) + E [F(x_1)] \tag{3.1}$$

e

$$V^*(x) = \inf_{u \in \mathbb{R}^m} V(x, u), \quad (3.2)$$

onde  $C : \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}$  e  $F : \mathbb{R}^n \rightarrow \mathbb{R}$  são funções convexas, não-negativas e Lipschitz. Consequentemente, nas próximas seções estaremos analisando o seguinte problema

$$\begin{aligned} \min_{u \in \mathbb{R}^m} V(x, u) &= C(x, u) + E[F(x_1)] \\ \text{s.a. } x_1(x, u) &= Ax + Bu + (\bar{\Sigma} + \sum_{i=1}^m \phi^i(u^i)\Sigma^i)\varepsilon \end{aligned}$$

Observe que esta expressão corresponde ao problema de PD em um passo para sistemas VCAI. A função  $F$  representa o custo acumulado, a função  $C$  é o custo em um dado instante e  $V^*$  será o custo mínimo para um passo. Além disso, o vetor aleatório  $x_1$  é determinado por (2.3) com  $x_0 = x$  e  $u_0 = u$ . Note que como o sistema é homogêneo no tempo, i.e., se  $x_k = x$  e  $u_k = u$ , estaremos avaliando o valor esperado em (3.1) de  $x_{k+1}$ .

## 3.2 Caracterização da Convexidade

Como será utilizado PD para obter a política ótima em sistemas onde a variação do controle aumenta a incerteza, um passo preliminar e fundamental é caracterizar a convexidade das funções  $V$  e  $V^*$ . O próximo lema mostra que se  $C$  e  $F$  forem convexas,  $V$  e  $V^*$  também serão.

**Lema 3.2.1.** *As funções  $V(x, u)$  e  $V^*(x)$  dadas por (3.1) e (3.2), respectivamente, são convexas.*

*Demonstração.* Para simplificar a notação, o valor esperado em (3.1) será denotado pela função  $\mathcal{E}_F : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}$ , dada por

$$\mathcal{E}_F(x, u) = E[F(x_1(x, u))] \quad (3.3)$$

onde

$$x_1(x, u) = Ax + Bu + (\bar{\Sigma} + \sum_{i=1}^m \Sigma^i \phi_i(u^i))\varepsilon.$$

Esta expressão é equivalente a:

$$x_1(x, u) = Ax + bu + \bar{\Sigma}\varepsilon + W\phi(u)$$

onde

$$W = [\Sigma^1\varepsilon \quad \Sigma^2\varepsilon \quad \dots \quad \Sigma^m\varepsilon]$$

e

$$\phi(u) = [\phi_1(u^1) \quad \phi_2(u^2) \quad \dots \quad \phi_m(u^m)]^T.$$

A partir da Definição 2.3.1, obtém-se

$$\partial x_1(x, u) = [A \quad B + W\mathbb{D}_u], \quad (3.4)$$

onde  $\mathbb{D}_u$  é o gradiente generalizado de  $\phi$  na origem, sendo dado por  $\mathbb{D}_u = \cup_{i=1}^m \text{co}\{D_-^i, D_+^i\}$ , onde

$$D_+^i = \text{diag} \left( 0, \dots, \lim_{u^i \downarrow 0} \partial \phi_i, \dots, 0 \right)$$

$$D_-^i = \text{diag} \left( 0, \dots, \lim_{u^i \uparrow 0} \partial \phi_i, \dots, 0 \right)$$

O próximo passo é mostrar que  $\mathcal{E}_F(x, u)$  é convexa. Seja  $\zeta \in \partial \mathcal{E}_F(x, u)$  e  $\tilde{\zeta} \in \partial \mathcal{E}_F(x, \tilde{u})$ . A partir da Proposição 2.3.5, tem-se

$$\zeta \in E[\beta^T \gamma], \quad \beta \in \partial F(x_1(x, u)), \quad \gamma \in \partial x_1(x, u)$$

$$\tilde{\zeta} \in E[\tilde{\beta}^T \tilde{\gamma}], \quad \tilde{\beta} \in \partial F(x_1(x, \tilde{u})), \quad \tilde{\gamma} \in \partial x_1(x, \tilde{u}).$$

onde

$$\gamma = [A \quad B + W D], \quad D \in \mathbb{D}_u$$

$$\tilde{\gamma} = [A \quad B + W \tilde{D}], \quad \tilde{D} \in \mathbb{D}_{\tilde{u}}.$$

Note que  $D$  e  $\tilde{D}$  são matrizes diagonais e seus elementos serão denotados por  $d^i$  e  $\tilde{d}^i$  respectivamente, logo

$$D = \text{diag}(d_1, d_2, \dots, d_m)$$

$$\tilde{D} = \text{diag}(\tilde{d}_1, \tilde{d}_2, \dots, \tilde{d}_m)$$

Para maior simplicidade, definimos

$$\Delta = \begin{bmatrix} x - \tilde{x} \\ u - \tilde{u} \end{bmatrix}, \quad x, \tilde{x} \in \mathbb{R}^n; \quad u, \tilde{u} \in \mathbb{R}$$

Partindo-se da Proposição 2.3.3, para provar que  $\mathcal{E}_F(x, u)$  é uma função convexa basta mostrar que

$$E[\beta^T \gamma - \tilde{\beta}^T \tilde{\gamma}] \Delta \geq 0, \quad \beta \in \partial F(x_1(x, u)), \quad \tilde{\beta} \in \partial F(x_1(x, \tilde{u})). \quad (3.5)$$

Note que

$$\gamma \Delta = [A \quad B + W D] \Delta \quad (3.6)$$

$$= Ax + Bu + \sum_{i=1}^m d_i u^i \Sigma^i \varepsilon - \left[ A \tilde{x} + B \tilde{u} + \sum_{i=1}^m d_i \tilde{u}^i \Sigma^i \varepsilon \right]$$

e

$$\begin{aligned}
-\tilde{\gamma}\Delta &= - [A \ B + W \tilde{D}] \Delta \\
&= A\tilde{x} + B\tilde{u} + \sum_{i=1}^m \tilde{d}_i \tilde{u}^i \Sigma^i \varepsilon - \left[ Ax + Bu + \sum_{i=1}^m \tilde{d}_i u^i \Sigma^i \varepsilon \right]
\end{aligned} \tag{3.7}$$

Por hipótese, a função de magnitude,  $\phi$ , é convexa, positiva e não diferenciável na origem. Sendo assim, supondo que  $c^i \in \mathbb{R}$  é uma constante não-negativa, pode-se utilizar a Proposição 2.3.3 com  $\zeta = d_i$ ,  $x = u^i$  e  $x_0 = 0$  para obter a igualdade

$$d_i u^i = \phi_i(u^i) + c^i, \quad \forall i \in [1, m]$$

de forma análoga, existe uma constante  $\tilde{c}^i \in \mathbb{R}$  tal que

$$\tilde{d}_i \tilde{u}^i = \phi_i(\tilde{u}^i) + \tilde{c}^i, \quad \forall i \in [1, m]$$

Logo, as Equações (3.6) e (3.7) podem ser escritas como

$$\begin{aligned}
\gamma\Delta &= Ax + Bu + \left( \bar{\Sigma} + \sum_{i=1}^m \phi_i(u^i) \Sigma^i \right) \varepsilon - \left[ A\tilde{x} + B\tilde{u} + \left( \bar{\Sigma} + \sum_{i=1}^m (d_i \tilde{u}^i - c^i) \Sigma^i \right) \varepsilon \right] \\
&= x_1(x, u) - x_1(\tilde{x}, \tilde{u}) + \left( \sum_{i=1}^m (-d_i \tilde{u}^i + c^i + \phi_i(\tilde{u}^i)) \Sigma^i \right) \varepsilon
\end{aligned}$$

e

$$\begin{aligned}
-\tilde{\gamma}\Delta &= A\tilde{x} + B\tilde{u} + \left( \bar{\Sigma} + \sum_{i=1}^m \phi_i(\tilde{u}^i) \Sigma^i \right) \varepsilon - \left[ Ax + Bu + \left( \bar{\Sigma} + \sum_{i=1}^m (\tilde{d}_i u^i - \tilde{c}^i) \Sigma^i \right) \varepsilon \right] \\
&= x_1(\tilde{x}, \tilde{u}) - x_1(x, u) + \left( \sum_{i=1}^m (-\tilde{d}_i u^i + \tilde{c}^i + \phi_i(u^i)) \Sigma^i \right) \varepsilon
\end{aligned}$$

A partir destas equações e notando que  $F$  é uma função convexa, pode-se usar a Proposição 2.3.1 de forma a obter

$$E[\beta^T \gamma] \Delta \geq \mathcal{E}_F(x, u) - \mathcal{E}_F(\tilde{x}, \tilde{u}) + E \left[ \beta^T \left( \sum_{i=1}^m (-d_i \tilde{u}^i + c^i + \phi_i(\tilde{u}^i)) \Sigma^i \right) \varepsilon \right] \tag{3.8}$$

e

$$-E[\tilde{\beta}^T \tilde{\gamma}^T] \Delta \geq \mathcal{E}_F(\tilde{x}, \tilde{u}) - \mathcal{E}_F(x, u) + E \left[ \tilde{\beta}^T \left( \sum_{i=1}^m (-\tilde{d}_i u^i + \tilde{c}^i + \phi_i(u^i)) \Sigma^i \right) \varepsilon \right] \tag{3.9}$$

Além disso, como a função  $\phi_i$  é convexa para todo  $i$ , tem-se

$$-d_i \tilde{u}^i + c^i + \phi_i(\tilde{u}^i) = d_i(u^i - \tilde{u}^i) - \phi_i(u^i) + \phi_i(\tilde{u}^i) \geq 0, \quad \forall i \in [1, m]$$

e

$$-\tilde{d}_i u^i + \tilde{c}^i + \phi_i(u^i) = \tilde{d}_i(\tilde{u}^i - u^i) - \phi_i(\tilde{u}^i) + \phi_i(u^i) \geq 0, \forall i \in [1, m]$$

Para simplificar a notação, sejam  $p \in \mathbb{R}$  e  $\tilde{p} \in \mathbb{R}$  duas constantes positivas definidas como

$$p^i = -d_i \tilde{u}^i + c^i + \phi_i(\tilde{u}^i) \geq 0$$

e

$$\tilde{p}^i = -\tilde{d}_i u^i + \tilde{c}^i + \phi_i(u^i) \geq 0$$

Logo, somando (3.8) e (3.9), e lembrando que  $F$  e  $\phi$  são funções Lipschitz, obtém-se

$$E[\beta^T \gamma] \Delta - E[\tilde{\beta}^T \tilde{\gamma}] \Delta \geq \sum_{i=1}^m p^i E[\beta^T \Sigma^i \varepsilon] + \sum_{i=1}^m \tilde{p}^i E[\tilde{\beta}^T \Sigma^i \varepsilon] \quad (3.10)$$

Como  $\Sigma^i$  é definida positiva para qualquer  $i \in [1, m]$ , segue diretamente do Lema 2.3.2 que

$$E[\beta^T \Sigma^i \varepsilon] \geq 0 \quad \text{e} \quad E[\tilde{\beta}^T \Sigma^i \varepsilon] \geq 0, \quad \forall i \in [1, m]$$

Desta forma, como  $p^i \geq 0$  e  $\tilde{p}^i \geq 0$  para todo  $i$ , tem-se que

$$\sum_{i=1}^m p^i E[\beta^T \Sigma^i \varepsilon] + \sum_{i=1}^m \tilde{p}^i E[\tilde{\beta}^T \Sigma^i \varepsilon] \geq 0$$

e, finalmente,

$$E[\beta^T \gamma - \tilde{\beta}^T \tilde{\gamma}] \Delta \geq 0, \quad \beta \in \partial F(x_1(x, u)), \quad \tilde{\beta} \in \partial F(x_1(\tilde{x}, \tilde{u})).$$

Consequentemente,  $(x, u) \rightarrow \mathcal{E}_F(x, u)$  é uma função convexa. Finalmente, como  $C$  é convexa,  $(x, u) \rightarrow V(x, u)$  e  $x \rightarrow V^*(x)$  (devido à Proposição 2.3.4) também serão convexas, completando a prova. □

É possível mostrar que  $V^*$  será diferenciável caso  $V$  seja uma função estritamente convexa e diferenciável. Além disso, a partir de  $V$ , podemos caracterizar o gradiente generalizado de  $V^*$ . Estes fatos estão enunciados no próximo Lema.

**Lema 3.2.2.** *Seja  $V$  definida como em (3.1) e  $V^*$  como em (3.2). Além disso, seja  $U_x^*$  definida como*

$$U_x^* = \{u \in \mathbb{R}^m : u \in \arg \min_{u \in \mathbb{U}} V(x, u)\}. \quad (3.11)$$

Logo,

1.  $\partial V^*(x) = \text{co}\{\partial_x V(x, u) : u \in U_x^*\}$ ;
2.  $V^*(x)$  será diferenciável se  $V$  for uma função estritamente convexa e diferenciável em  $x$ .

*Demonstração.* A demonstração segue diretamente das Proposições 2.3.8 e 2.3.9. □

### 3.3 Quando a Não-Variação é Ótima

Baseado no fato que a função custo é convexa, é possível proceder com o intuito de determinar o sinal de cada componente de  $u^*$  utilizando somente o valor do estado  $x$ . Suponha uma função  $(x, u) \rightarrow f(x, u)$  diferenciável e convexa em  $u$ . O sinal do mínimo de cada componente de  $u$ , denotado por  $u^i$ , pode ser obtido analisando  $\langle \nabla_u f|_{u=0}, e_i \rangle$  para cada  $x$ , onde  $e_i \in \mathbb{R}^m$  é o vetor canônico que possui apenas a  $i$ -ésima componente não nula. Desta forma, se para um dado elemento  $u^i$  de  $u$ ,  $\langle \nabla_u f|_{u=0}, e_i \rangle > 0$  ( $< 0$ ), então a função é crescente (decrecente) na origem para a componente  $u^i$  e, conseqüentemente, o mínimo  $u^{*i}$  estará no semi-plano negativo (positivo). Obviamente, se  $\nabla_u f|_{u=0} = 0$  a solução ótima será  $u^* = 0$ . Repetindo este procedimento para cada elemento de  $u$ , determina-se em que quadrante estará o controle ótimo.

Observe que esta análise não pode ser aplicada a  $V$  em (3.1) pois, mesmo que  $V$  seja convexa, ela não será necessariamente diferenciável em  $u = 0$  (inclusive, para sistemas VCAI, esse nunca será o caso). O próximo lema apresenta um resultado que trata desta questão e indica que poderá haver uma região no espaço de estados onde  $u^* = 0$  se a função custo for Lipschitz. Além disso, para o caso onde  $F$  for estritamente convexa, diferenciável e coerciva, e  $\nabla_u C(x, u)$  for uma função monótona e contínua em  $x$ ,  $\forall u \in \mathbb{R}^m$ , será mostrado que uma região onde  $u^* = 0$  sempre existirá.

**Lema 3.3.1.** *Para a função  $V$  dada por (3.1),  $V^*$  descrita por (3.11) e  $u^* = [u^{*1} \dots u^{*i} \dots u^{*m}]^T$ , tem-se, para todo  $i \in \{1, 2, \dots, m\}$ ,*

$$\begin{cases} u^{*i} > 0, & \text{se } x \in \mathcal{R}_1^i(V), \\ u^{*i} < 0, & \text{se } x \in \mathcal{R}_2^i(V), \\ u^{*i} = 0, & \text{se } x \in \mathcal{R}_3^i(V). \end{cases} \quad (3.12)$$

onde

$$\mathcal{R}_1^i(V) = \{x : x \in \mathbb{R}^n, \lim_{u^i \downarrow 0} \nabla_{u^i} V(x, u^i e_i) < 0\}, \quad (3.13)$$

$$\mathcal{R}_2^i(V) = \{x : x \in \mathbb{R}^n, \lim_{u^i \uparrow 0} \nabla_{u^i} V(x, u^i e_i) > 0\}, \quad (3.14)$$

$$\mathcal{R}_3^i(V) = \overline{\mathcal{R}_1 \cup \mathcal{R}_2}. \quad (3.15)$$

e o vetor  $u^i e_i$  é um vetor onde apenas a  $i$ -ésima componente é não nula.

*Demonstração.* Como  $\partial_{u^i} V$  é um ponto ou um intervalo fechado da reta, obtemos da convexidade de  $V$  uma noção de que  $\partial_{u^i} V$  é não-decrescente em  $u^i$  no sentido que, para dois vetores  $u_1$  e  $u_2$  onde  $u_1^i \neq u_2^i$  e  $u_1^j = u_2^j, \forall j \neq i$ ,

$$u_1^i \leq u_2^i \iff \exists \gamma_1^i \leq \gamma_2^i, \quad \gamma_1^i \in \partial_{u^i} V(u_1^i, x), \gamma_2^i \in \partial_{u^i} V(u_2^i, x).$$

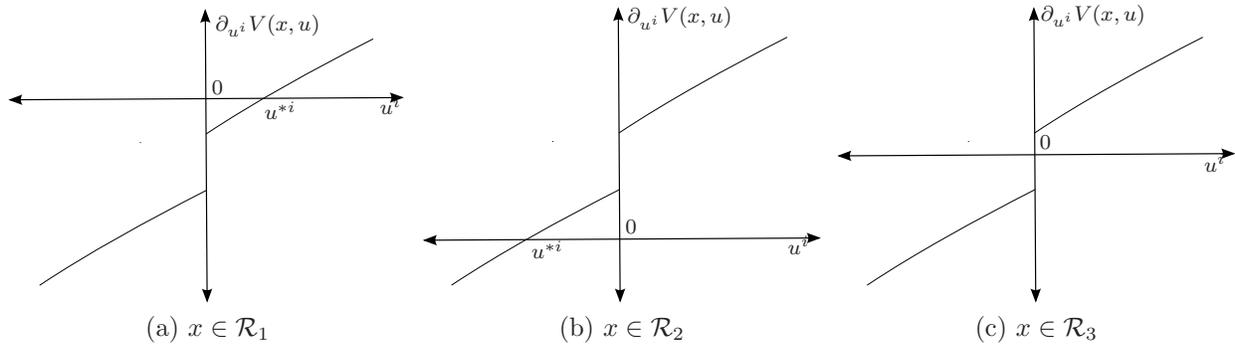


Figura 3.1: Comportamento de  $\partial_{u^i} V(x, u)$  para diferentes  $x$ . Em (c),  $u^{*i} = 0$ .

Conseqüentemente, podemos determinar o sinal de  $u^*(x)$  observando apenas o sinal de  $\partial_{u^i} V|_{u=0}$  para todo  $i$ . Se  $\gamma^i > 0, \forall \gamma^i \in \partial_{u^i} V|_{u=0}$ , então  $V$  é crescente na origem e  $u^{*i} < 0$ . Analogamente, se  $\gamma^i < 0, \forall \gamma^i \in \partial_{u^i} V|_{u=0}$  então  $V$  é decrescente na origem e  $u^{*i} > 0$ . Claramente, se  $0 \in \partial_u V|_{u=0}$  então  $u^* = 0$ , conforme constatado na Proposição 2.3.6.

Note que na análise acima não é necessário verificar todos os elementos de  $\partial_u V|_{u=0}$  para determinar o sinal de  $u^*$ . Baseado no fato que  $\partial_{u^i} V$  é crescente em  $u^i$  para todo  $i \in \{1, 2, \dots, m\}$  e utilizando a definição de gradiente generalizado, observe que

$$\begin{aligned} \lim_{u^i \downarrow 0} \nabla_{u^i} V(x, u^i e_i) &= \max \partial_{u^i} V|_{u=0}, \\ \lim_{u^i \uparrow 0} \nabla_{u^i} V(x, u^i e_i) &= \min \partial_{u^i} V|_{u=0}, \end{aligned}$$

onde os limites são calculados evitando conjuntos de medida nula e pontos onde  $u \rightarrow V(x, u)$  não são diferenciáveis.

A partir destes argumentos, obtemos as seguintes condições

$$\begin{aligned} \lim_{u^i \downarrow 0} \nabla_{u^i} V(x, u^i e_i) < 0 &\implies u^{*i}(x) > 0, \\ \lim_{u^i \uparrow 0} \nabla_{u^i} V(x, u^i e_i) > 0 &\implies u^{*i}(x) < 0, \end{aligned}$$

o que resulta, para cada componente  $u^i$  de  $u$ , em três regiões complementares definidas no enunciado do lema como (3.13), (3.14) and (3.15) e que cobrem o espaço de estados.  $\square$

**Comentário 3.3.1.** O comportamento de  $\partial_{u^i} V$  em função do valor de  $x$  é ilustrado na Figura 3.1.

**Definição 3.3.1.** A região  $\mathcal{R}_3$ , definida como

$$\mathcal{R}_3 = \{x : 0 \in \partial_u V(x, 0)\}, \quad (3.16)$$

será chamada de **região de não-variação**.

**Lema 3.3.2.** *Se  $F$  for uma função estritamente convexa, diferenciável e coerciva, e  $\nabla_u C(x, u)$  for uma função monótona e contínua em  $x$ ,  $\forall u \in \mathbb{R}^m$ , a região  $\mathcal{R}_3$  será um hipervolume.*

*Demonstração.* Lembrando que o vetor  $u_0^i$  possui apenas a  $i$ -ésima componente não nula, a região  $\mathcal{R}_3$  pode ser definida como

$$\mathcal{R}_3 = \left\{ x : \lim_{u^i \downarrow 0} \nabla_{u^i} V(x, u^i e_i) > 0, \lim_{u^i \uparrow 0} \nabla_{u^i} V(x, u^i e_i) < 0, \forall i \in \{1, 2, \dots, m\} \right\} \quad (3.17)$$

Sendo assim, se  $C$  e  $F$  são diferenciáveis e denotando a  $i$ -ésima base canônica do espaço  $\mathbb{R}^m$  por  $e_i$ , temos

$$\lim_{u^i \downarrow 0} \nabla_{u^i} V(x, u^i e_i) = \nabla_{u^i} C(x, 0) + E \left[ (B e_i + \Sigma_i d_+^i \varepsilon)^T \nabla F(Ax + \bar{\Sigma} \varepsilon) \right],$$

$$\lim_{u^i \uparrow 0} \nabla_{u^i} V(x, u^i e_i) = \nabla_{u^i} C(x, 0) + E \left[ (B e_i + \Sigma_i d_-^i \varepsilon)^T \nabla F(Ax + \bar{\Sigma} \varepsilon) \right],$$

em que

$$d_+^i = \lim_{u^i \downarrow 0} \nabla_{u^i} \phi_i(u^i)$$

e

$$d_-^i = \lim_{u^i \uparrow 0} \nabla_{u^i} \phi_i(u^i).$$

Observe que como  $\phi_i(u^i)$  é convexa e atinge o mínimo na origem,  $d_+^i > 0$  e  $d_-^i < 0$ ,  $\forall i$ . Além disso, a região (3.17) pode ser reescrita como

$$\mathcal{R}_3 = \left\{ x : -d_+^i E[\nabla^T F(Ax + \bar{\Sigma} \varepsilon) \Sigma_i \varepsilon] < \nabla_{u^i} C(x, 0) + E[\nabla F(Ax + \bar{\Sigma} \varepsilon)] B e_i < -d_-^i E[\nabla^T F(Ax + \bar{\Sigma} \varepsilon) \Sigma_i \varepsilon], \forall i \in \{1, \dots, m\} \right\}$$

Segue diretamente do Lema 2.3.2 que, se  $F$  for estritamente convexa,

$$E[\nabla^T F(Ax + \bar{\Sigma} \varepsilon) \Sigma_i \varepsilon] > 0$$

Consequentemente,

$$-d_+^i E[\nabla^T F(Ax + \bar{\Sigma} \varepsilon) \Sigma_i \varepsilon] < 0, \quad -d_-^i E[\nabla^T F(Ax + \bar{\Sigma} \varepsilon) \Sigma_i \varepsilon] > 0, \quad \forall i \in \{1, \dots, m\}$$

Entretanto, por hipótese,  $\nabla F$  é uma função estritamente crescente e contínua, pois  $F$  é estritamente convexa, diferenciável e coerciva, e  $\nabla_u C(x, u)$  é monótona em  $x$ . Ao somar as duas funções, existirá  $x$  tal que  $\nabla_{u^i} C(x, 0) + E[\nabla F(Ax + \bar{\Sigma} \varepsilon)] B e_i$  poderá assumir valores arbitrariamente próximos da origem. Sendo assim,

$$\exists x : |\nabla_{u^i} C(x, 0) + E[\nabla F(Ax + \bar{\Sigma} \varepsilon)] B e_i| < \delta, \forall i \in \{1, \dots, m\} \text{ e } \delta > 0$$

Desta forma,  $\mathcal{R}_3$  será um hipervolume no espaço de estados.  $\square$

**Comentário 3.3.2.** Note que para qualquer função quadrática positiva  $C : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}$  da forma

$$C(x, u) = x^T Q x + u^T P x + u^T R u, \quad Q \in \mathbb{R}_+^{n \times n}, \quad P \in \mathbb{R}^{m \times n}, \quad R \in \mathbb{R}_+^{m \times m}$$

a função  $g(x) = \nabla_u C(x, u)$  será monótona em  $x$ , satisfazendo a hipótese do Lema 3.3.2.

## 3.4 A Programação Dinâmica

A partir dos Lemas anteriores, é possível caracterizar o método da Programação Dinâmica aplicado ao sistema (2.3). Vamos denotar por  $J_k^*$  o custo acumulado ótimo em um instante  $k$  para cada  $x$  como

$$J_k^*(x) = \inf_{\pi \in \Pi} E \left[ \sum_{n=k}^{N-1} C_n(x_n, u_n) + C_N(x_N) \right]$$

A próxima proposição fornece uma condição suficiente para a existência das políticas ótimas de realimentação  $\pi_k^*(x)$ .

**Proposição 3.4.1** ([3]). *Suponha que (2.4) é finita para todo  $x_0 \in \mathbb{R}^n$  e  $\pi \in \Pi$ . Logo, se o conjunto*

$$H_k(x, \lambda) = \{u \in \mathbb{R}^m : C_k(x, u) + E \left[ J_{k+1}^*(A_k x + B_k u + [\bar{\Sigma}_k + \sum_{i=1}^m \Sigma^i \phi_i(u^i)] \varepsilon_k) \right] \leq \lambda \}$$

é compacto para todo  $x \in \mathbb{R}^n$ ,  $\lambda \in \mathbb{R}$  e  $k \in [0, N-1]$ , uma política ótima para  $N$  estágios existirá.

**Teorema 3.4.1.** *Suponha que para cada  $0 \leq k < N$  tem-se que  $x \rightarrow C_k(x, u)$ ,  $u \rightarrow C_k(x, u)$  e  $x \rightarrow C_N(x)$  sejam funções convexas e limitadas para todo  $x \in \mathbb{R}^n$  e  $u \in \mathbb{R}^m$ . Para o sistema em (2.3) avaliado através da função custo (2.4), a política ótima pode ser obtida recursivamente através do seguinte procedimento.*

1. Definir  $J_N^*(x) = C_N(x)$ ,  $x \in \mathbb{R}^n$  e fazer  $k = N - 1$ ;

2. Definir  $J_k(x, u)$  para cada  $x \in \mathbb{R}^n$ , como

$$J_k(x, u) = C_k(x, u) + E \left[ J_{k+1}^* \left( A_k x + B_k u + (\bar{\Sigma}_k + \sum_{i=1}^m \Sigma_i \phi_i(u^i)) \varepsilon_k \right) \right],$$

3. Para cada  $x \in \mathbb{R}^n$ , determine o sinal da ação ótima  $u_k^*$  usando

$$\begin{cases} u_k^{i*} > 0, & \text{se } x \in \mathcal{R}_1^i(J_k), \\ u_k^{i*} < 0, & \text{se } x \in \mathcal{R}_2^i(J_k), \\ u_k^{i*} = 0, & \text{se } x \in \mathcal{R}_3^i(J_k). \end{cases}$$

Se  $\nabla_u C_k(x, u)$  for monótona em  $x$  e se  $C_k$  for coerciva para todo  $k$ , a região  $\mathcal{R}_3$  dada por (3.16) será um hipervolume. Caso  $u_k^{*i} \in \mathcal{R}_1^i(J_k)$  ou  $u_k^{*i} \in \mathcal{R}_2^i(J_k)$  para algum  $i \in \{1, \dots, m\}$ , determine  $u_k^*$  tal que  $J_k(x, u_k^*) \leq J_k(x, u)$  para cada  $u \in \mathbb{R}^m$ . Isso é equivalente a exigir que

$$0 \in \partial_u C_k(x, u_k^*) + \partial_u E \left[ J_{k+1}^* \left( A_k x + B_k u_k^* + \left( \bar{\Sigma}_k + \sum_{i=1}^m \Sigma_i \phi_i(u_k^{*i}) \right) \varepsilon_k \right) \right]$$

4. Defina a função  $J_k^*(x)$  como

$$J_k^*(x) = J_k(x, u_k^*) = C(x, u_k^*) + E[J_{k+1}^*(x_{k+1})]$$

com  $x_{k+1} = A_k x + B_k u_k^* + \left( \bar{\Sigma}_k + \sum_{i=1}^m \Sigma_i \phi_i(u_k^{*i}) \right) \varepsilon_k$ . A função  $u_k^*(x)$  que atender a condição

$$J_k^*(x) = J_k(x, u_k^*) \leq J_k(x, u), \quad \forall u \in \mathbb{R}^m$$

tal que  $u_k^*(x) = u_k^*$ ,  $\forall x \in \mathbb{R}^n$ , será a política ótima. Se  $k = 0$ , pare. Senão, retorne ao passo 2.

Desta forma, a política ótima  $u_k^*(x)$  é obtida para cada  $0 \leq k < N$  e, em particular,  $J^*(x) = J_0(x, u_0^*(x)) \leq J_0(x, u)$ ,  $\forall (x, u)$ .

*Demonstração.* A prova segue diretamente do Lema 3.2.1, Lema 3.3.1 e da equação de Bellman. Observe da Proposição 3.4.1 que, sendo  $C_k$  convexa para todo  $k \in [0, N]$ , as políticas ótimas de realimentação  $x \rightarrow u_k^*(x)$  sempre existirão.  $\square$

Observe que, a cada instante  $k$ , o tomador de decisões poderá aumentar, diminuir ou manter a sua política dependendo do valor do estado. Uma possível interpretação para a região onde  $u_k^* = 0$  é que, em face do risco gerada pela variação da entrada do sistema, é melhor manter a entrada constante. Intuitivamente, esta estratégia está de acordo com problemas reais onde um tomador de decisão, em face da incerteza que a sua decisão pode gerar, decide manter a mesma política que ele possuía anteriormente.

### 3.4.1 Um Exemplo Bidimensional

Para representar graficamente a região de não-variação, ou seja, a região no espaço de estados onde a política ótima é de não variação do controle ( $v^* = 0$ ), considere um sistema bidimensional descrito por:

$$x_{k+1} = A x_k + B u_k + (1 + |u_k|) \varepsilon_k, \quad \varepsilon_k \sim N(0, I_{n \times n})$$

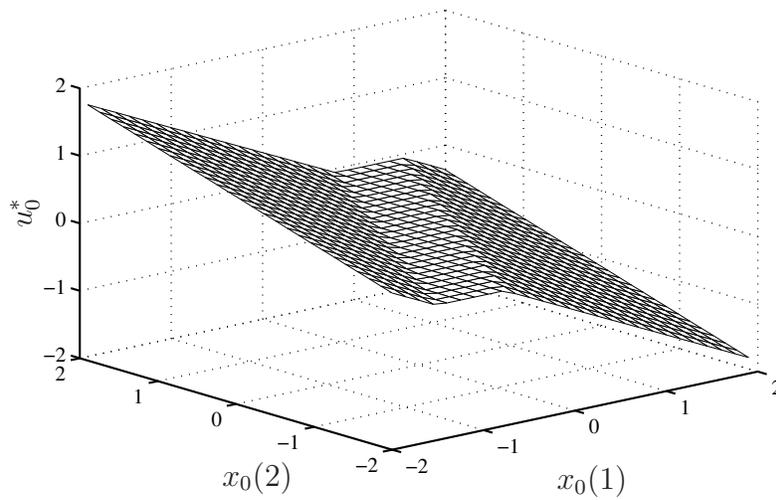


Figura 3.2: Entrada ótima  $u^*$  como uma função do valor do estado.

onde

$$A = \begin{bmatrix} 2 & 1 \\ 1 & -2 \end{bmatrix}; \quad b = \begin{bmatrix} 1 \\ 2 \end{bmatrix}.$$

Suponha que o sistema é avaliado através da função de custo quadrática:

$$J(x_0, \pi) = \sum_{k=0}^4 E[(x_k^T x_k + 2u_k P^T x + 0.5u_k^2) + x_5^T x_5]$$

com  $P = [0 \ 1]^T$ . A Figura 2 mostra  $u_0^*$  como uma função do estado inicial  $x_0$ . A região onde  $u_0^* = 0$  pode ser vista claramente no centro do gráfico.

## Desafios Numéricos

Nos capítulos anteriores, um modelo para sistemas onde a variação da magnitude do controle aumenta a incerteza sistêmica foi proposto e, utilizando Programação Dinâmica, a política ótima de controle para este tipo de sistema avaliado através de uma função custo positiva e convexa foi caracterizada. Foi mostrado, também, que existirá uma região no espaço de estados onde a política ótima será de não variação do controle, chamada de “região de não-variação”. Entretanto, esta caracterização teórica não possui importância prática se não existirem algoritmos numéricos eficientes para a obtenção da política ótima. Além disso, ainda não foi apresentado nenhum exemplo prático de um sistema real modelado como um sistema VCAI. Existe o risco de concluir que o modelo proposto não descreve adequadamente os problemas de tomada de decisão que motivaram este trabalho, confinando este modelo à classe de sistemas dinâmicos de razoável interesse teórico, mas de pouca utilidade prática. Este capítulo e o próximo tratam destas questões.

Neste capítulo, serão apresentadas algumas estratégias numéricas para o cálculo da política ótima em sistemas VCAI envolvendo aproximações da função custo para a Programação Dinâmica. Inicialmente, será feito um estudo sobre erros de aproximação em algoritmos de PD. Em seguida, será analisado o caso onde o sistema é mono-entrada com o desempenho avaliado através de uma função de custo quadrática e duas aproximações para a função custo serão propostas. Um algoritmo para a obtenção da política ótima para um horizonte infinito também será apresentado. Finalmente, uma estratégia de aproximação da função custo utilizando Máquinas de Vetores de Suporte será exposta como uma forma geral para lidar com problemas de controle estocástico ótimo em horizonte finito.

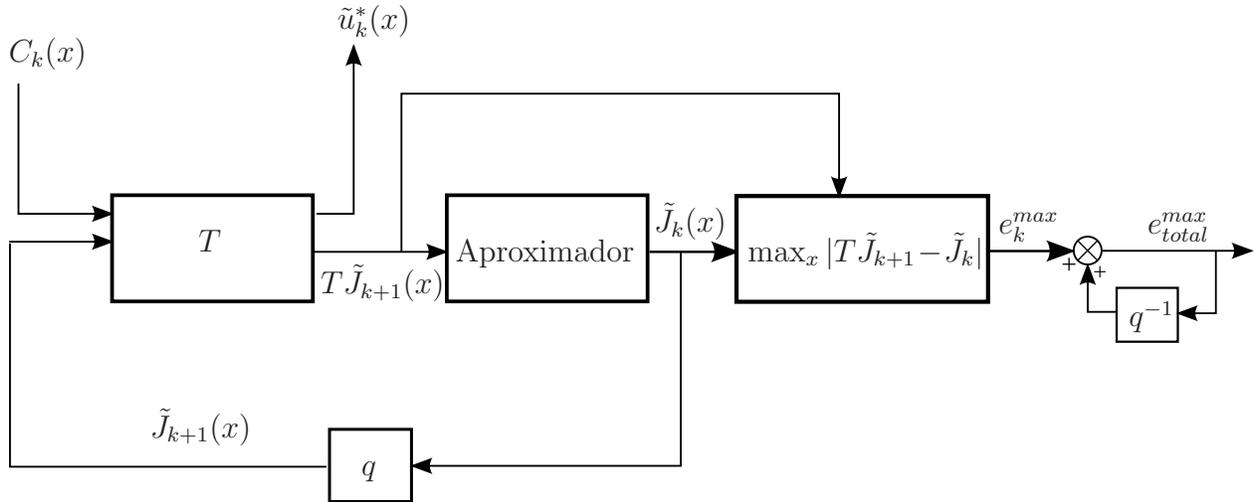


Figura 4.1: Programação Dinâmica Aproximada. O operador  $q$  representa um avanço unitário,  $q^{-1}$  um atraso unitário e  $T$  é o operador da programação dinâmica.

## 4.1 Programação Dinâmica Aproximada

Conforme expresso na Equação (2.13), o operador da programação dinâmica  $T$  sobre uma função  $J : S \rightarrow \mathbb{R}$ , sendo  $S \subset \mathbb{R}^n$  um conjunto compacto e convexo, é

$$(T J)(x) = \min_{u \in U(x)} E\{C(x, u) + J(x_1(x, u))\}, \quad x \in S.$$

Muitas vezes, não é possível obter uma expressão analítica para  $(T J)(x)$ , sendo necessário recorrer a uma aproximação  $\tilde{J}(x) \cong (T J)(x)$ . Ao realizar esta aproximação a cada passo da programação dinâmica será gerada uma política de controle sub-ótima, ou seja, com custo esperado superior à política ótima. Entretanto, uma aproximação adequada poderá levar a um algoritmo de baixo custo computacional e a um controlador de fácil implementação, com um custo satisfatório. Algoritmos que envolvem aproximações da função custo acumulado são chamados de **Algoritmos de Programação Dinâmica Aproximada**.

A Figura 4.1 ilustra como funciona este tipo de algoritmo para um horizonte finito. A cada passo do algoritmo, será feita uma aproximação do custo acumulado e o erro máximo será estimado. O bloco fundamental para obter um algoritmo eficiente é a escolha do aproximador, que será responsável pela aproximação da função custo acumulado. Esta escolha deverá ser feita levando em conta particularidades do problema de controle que está sendo resolvido. Note que o erro total máximo será simplesmente a soma do erro máximo de aproximação em cada passo. A Proposição 4.1.1 e o Lema 4.1.1 tratam do controle de erro na programação dinâmica aproximada.

**Proposição 4.1.1** ([1]). *Seja o operador  $T$  da programação dinâmica dado por*

$$(T J)(x) = \min_{u \in U(x)} E\{C(x, u) + J(x_1(x, u))\}.$$

*Além disso, seja*

$$(T^k J)(x) = (T(T^{k-1} J))(x)$$

*e seja o conjunto não-vazio e convexo  $S \subset \mathbb{R}^n$ . Logo, para duas funções limitadas  $J : S \rightarrow \mathbb{R}$  e  $\tilde{J} : S \rightarrow \mathbb{R}$ ,*

$$\max_{x \in S} |(T^k J)(x) - (T^k \tilde{J})(x)| \leq \max_{x \in S} |J(x) - \tilde{J}(x)|$$

**Lema 4.1.1.** *Sejam  $J' : S \rightarrow \mathbb{R}$ ,  $J : S \rightarrow \mathbb{R}$  e  $\tilde{J} : S \rightarrow \mathbb{R}$ . Suponha*

$$\max_{x \in S} |J(x) - (T J')(x)| \leq \epsilon_1$$

*e*

$$\max_{x \in S} |\tilde{J}(x) - (T J)(x)| \leq \epsilon_2.$$

*Sendo assim,*

$$\max_{x \in S} |\tilde{J}(x) - (T^2 J')(x)| \leq \epsilon_1 + \epsilon_2.$$

*Demonstração.* Observe que

$$\begin{aligned} \max_{x \in S} |\tilde{J}(x) - (T^2 J')(x)| &= \max_{x \in S} |\tilde{J}(x) - (T J)(x) + (T J)(x) - (T^2 J')(x)| \\ &\leq \max_{x \in S} |\tilde{J}(x) - (T J)(x)| + \max_{x \in S} |(T J)(x) - (T^2 J')(x)| \end{aligned}$$

Da Proposição 4.1.1, note que

$$\max_{x \in S} |(T J)(x) - (T^2 J')(x)| \leq \max_{x \in S} |J(x) - (T J')(x)| \leq \epsilon_1$$

Além disso, por hipótese,

$$\max_{x \in S} |\tilde{J}(x) - (T J)(x)| \leq \epsilon_2.$$

Consequentemente,

$$\max_{x \in S} |\tilde{J}(x) - (T^2 J')(x)| \leq \epsilon_1 + \epsilon_2$$

e a prova está completa.  $\square$

A seguir, serão mostrados dois tipos de aproximação que serão utilizados para obter políticas sub-ótimas para sistemas VCAI monoentrada e associados a funções de custo quadráticas.

## 4.2 O Caso Quadrático Mono-Entrada

Como funções de custo quadráticas são muito utilizadas devido a sua conveniência analítica, este caso será estudado detalhadamente. Desta forma, considere o seguinte sistema de controle estocástico:

$$\min J(x_0, \pi) = E \left[ \sum_{k=0}^{N-1} (x_k^T Q_k x_k + 2v_k P_k^T x + v_k^T R_k v_k) + x_N^T Q_N x_N \right] \quad (4.1)$$

$$\text{s.a. } x_{k+1} = A_k x_k + b_k v_k + (\bar{\sigma}_k + \sigma_k |v_k|) \xi_k, \quad x_0 = x \in \mathbb{R}^n \quad (4.2)$$

em que  $Q_k$  e  $R_k$  são matrizes semi-definidas positivas para todo  $k$  e  $J(x_0, \pi)$  é uma função não-negativa.

Para obter a política ótima, o Teorema 3.4.1 poderia ser utilizado diretamente. Entretanto, devido a presença da não-linearidade na porção estocástica da equação de estado, o custo acumulado será não-linear e minimizá-lo exigiria um alto custo computacional. Conseqüentemente, é interessante buscar uma aproximação adequada para o custo acumulado. Suponha uma situação onde o custo acumulado em um dado instante  $k$  fosse dado por

$$J_k^*(x) = \inf_{v \in \mathbb{R}} x^T Q x + 2v P^T x + R v^2 + E[x_1^T Q_1^* x_1 + R_1^{*T} x_1 + c_1] \quad (4.3)$$

com  $x_1 = Ax + bv + (\bar{\sigma} + \sigma|v|)$ . Observe que:

$$\begin{aligned} E[x_1^T Q_1^* x_1] &= x^T A^T Q_1^* A x + 2x^T A^T Q_1^* b v + (b^T Q_1^* b + \sigma^2 \text{tr}(Q_1^*)) v^2 \\ &\quad + (\bar{\sigma}^2 + 2\bar{\sigma}\sigma|v|) \text{tr}(Q_1^*) \\ E[R_1^{*T} x_1] &= R_1^{*T} A x + R_1^{*T} b v \end{aligned}$$

Logo, a função custo pode ser reescrita como:

$$\begin{aligned} J_k^*(x) &= \inf_{v \in \mathbb{R}} K v^2 + 2v L^T x + 2M|v| + N v \\ &\quad + x^T (Q + A^T Q_1^* A) x + R_1^{*T} A x + c_1 + \bar{\sigma}^2 \text{tr}(Q_1^*) \end{aligned}$$

onde

$$\begin{aligned} K &:= R + b^T Q_1^* b + \sigma^2 \text{tr}(Q_1^*), & L &:= A^T Q_1^* b + P, \\ M &:= \bar{\sigma} \sigma \text{tr}(Q_1^*), & N &:= R_1^{*T} b \end{aligned}$$

Utilizando o Lemma 3.3.1, a variação ótima do controle,  $v^*(x)$ , será uma função linear por partes do estado, sendo dada por:

$$v^* = \begin{cases} -K^{-1}(L^T x - M + N/2), & \text{if } Lx > M - N/2 \\ -K^{-1}(L^T x + M + N/2), & \text{if } Lx < -M - N/2 \\ 0, & \text{caso contrário} \end{cases}$$

Essa expressão pode ser reescrita como

$$v^* = -K^{-1}(L^T x + N/2 + Ms(x))$$

onde

$$s(x) = \begin{cases} -1, & \text{if } L^T x > M - N/2 \\ 1, & \text{if } L^T x < -M - N/2 \\ -M^{-1}(L^T x + N/2), & \text{caso contrário} \end{cases}$$

Fazendo  $v = v(x)$  em (4.3) a função custo mínima torna-se:

$$J_k^*(x) = -K^{-1}(L^T x + Ms(x) + N/2)^2 + x^T(Q + A^T Q_1^* A)x \\ + R_1^{*T} Ax + c_1 + \bar{\sigma}^2 \text{tr}(Q_1^*)$$

Rearranjando os termos, tem-se que

$$J_k^*(x) = x^T(Q + A^T Q_1^* A - L^T K^{-1} L)x + (R_1^{*T} A - K^{-1} N L^T)x + \bar{\sigma}^2 \text{tr}(Q_1^*) - K^{-1} N^2/4 + c_1 \\ - K^{-1} M(Ms(x))^2 + 2x^T Ls(x) + Ns(x)$$

Note que a segunda linha da expressão anterior é não-linear, não quadrática e não constante.

Definindo a função  $g : \mathbb{R}^n \rightarrow \mathbb{R}$  como

$$g(x) = Ms(x)^2 + 2x^T Ls(x) + Ns(x)$$

ou seja,

$$g(x) = \begin{cases} 2L^T x - M + N, & \text{if } L^T x > M - N/2 \\ -(2L^T x + M + N), & \text{if } L^T x < -M - N/2 \\ M^{-1}(L^T x + N/2)^2, & \text{caso contrário} \end{cases}$$

tem-se:

$$J_k^*(x) = x^T(Q + A^T Q_1^* A - L^T K^{-1} L)x + (R_1^{*T} A - K^{-1} N L^T)x \\ + c_1 + \bar{\sigma}^2 \text{tr}(Q_1^*) - K^{-1} N^2/4 + K^{-1} M g(x)$$

Neste caso, uma aproximação adequada para  $g(x)$  levaria a um controlador que seria uma função linear por partes do estado, evitando a necessidade de calcular valores esperados complicados. A tarefa de aproximar  $g(x)$  torna-se ainda mais simples fazendo uma mudança de variáveis

$$y(x) = L^T x + N/2,$$

resultando na função  $h(y) = g(x)$  dada por

$$h(y) = \begin{cases} 2y - M & \text{if } y > M \\ -2y - M, & \text{if } y < -M \\ M^{-1}y^2, & \text{caso contrário} \end{cases} .$$

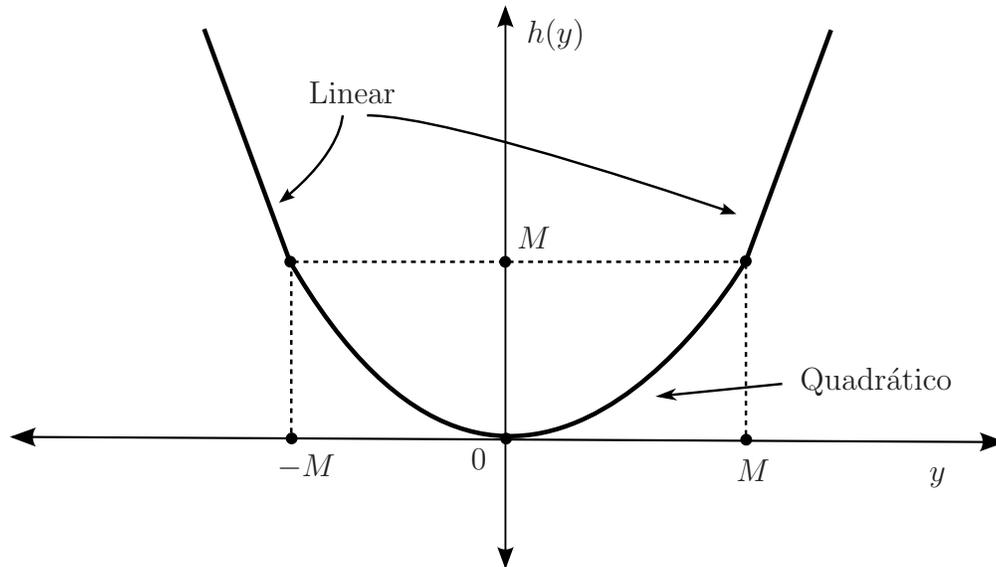


Figura 4.2: Comportamento de  $h(y)$ . Note que para  $|y| \leq M$  a função é quadrática e para  $|y| > M$  a função é linear.

O comportamento de  $h(y)$  está ilustrado na Figura 4.2. Conseqüentemente, o problema de aproximar a porção não quadrática do custo ótimo torna-se o problema de aproximar a função  $h : \mathbb{R} \rightarrow \mathbb{R}$ . Tendo isto em mente, serão propostas duas estratégias para aproximar  $g(x)$ . Uma primeira aproximação consistirá em simplesmente ignorar a porção não-linear da função custo e a segunda será uma aproximação quadrática por mínimos quadrados em um dado intervalo. As principais vantagens destas duas aproximações é o baixo esforço computacional que elas exigem para o cálculo do controlador ótimo e a possibilidade de avaliar explicitamente o erro da aproximação.

### 4.2.1 Aproximação por Mínimos Quadrados

Uma aproximação quadrática para  $g(x)$  em uma dada região  $\mathcal{X} \subseteq \mathbb{R}$  pode ser obtida da aproximação quadrática de  $h(y)$  em um intervalo  $[l_1, l_2]$ ,  $l_2 > l_1$  onde  $y : \mathcal{X} \rightarrow [l_1, l_2]$ . O erro quadrático médio  $e$  no intervalo  $[l_1, l_2]$  entre a função quadrática e  $h$  será

$$e(a, l_1, l_2) = \frac{1}{2(l_2 - l_1)} \int_{l_1}^{l_2} (a_0 y^2 + a_1 y + a_2 - h(y))^2 dy,$$

onde

$$a = [a_0 \quad a_1 \quad a_2]^T$$

A aproximação quadrática ótima por mínimos quadrados para o intervalo  $[l_1, l_2]$ , denotado por  $h^*(y)$ , será

$$h^*(y) = [y^2 \quad y \quad 1]^T a^*$$

onde  $a^*$  é a solução de

$$\nabla_a e(a^*, l_1, l_2) = 0$$

Além disso, o erro máximo da aproximação, denotado por  $e_a^{\max}$ , será

$$e_a^{\max}(l_1, l_2) = \max_{y \in [l_1, l_2]} |a_0^* y^2 + a_1^* y + a_2^* - h(y)|$$

Observe que  $a^*$  pode ser facilmente calculado uma vez que o intervalo de aproximação  $[l_1, l_2]$  for determinado. Como  $h$  é uma função simétrica, é interessante escolher um intervalo de aproximação que também seja simétrico. Sendo assim, suponha que deseja-se aproximar  $h$  para um dado intervalo  $y \in [-l, l]$ ,  $l > M > 0$ . O vetor  $a^*$  pode ser calculado analiticamente, obtendo:

$$\begin{aligned} a_0^* &= -\frac{-15l^4 + 10l^2M^2 - 3M^4}{8l^5} \\ a_1^* &= 0 \\ a_2^* &= -\frac{(l - M)^3(3l + M)}{8l^3} \end{aligned}$$

Consequentemente, o erro máximo da aproximação é:

$$e_a^{\max}(-l, l) = \max \left\{ \frac{(l - M)^3(3l + M)}{8l^3}, \frac{|19l^2 - 3M^2|(l^2 - M^2)^3}{8l^3|15l^4 - 10l^2M^2 + 3M^4|} \right\}$$

Para  $l \gg M$ , o primeiro termo irá dominar enquanto o segundo termo dominará quando  $l \approx M$ . Sendo assim, a aproximação para  $J_k^*$ , denotada por  $\tilde{J}_k^*$ , será

$$\begin{aligned} \tilde{J}_k^*(x) &= x^T(Q + A^T Q_1^* A - L^T K^{-1} L)x + (R_1^{*T} A - K^{-1} N L^T)x + c_1 + \bar{\sigma}^2 \text{tr}(Q_1^*) \\ &\quad - K^{-1} N^2/4 + K^{-1} M(a_0^*(L^T x + N/2)^2 + a_1^*(L^T x + N/2) + a_2^*) \end{aligned}$$

Além disso, o erro máximo da aproximação em cada passo da programação dinâmica será dado por

$$e_k^{\max}(-l, l) = K^{-1} M e_a^{\max}(-l, l)$$

### 4.2.2 Aproximação por Truncamento

Uma segunda estratégia para a aproximação de  $J_k^*(x)$  é ignorar a porção não quadrática desta expressão. Neste caso, tem-se

$$\tilde{J}_k^*(x) = x^T(Q + A^T Q_1^* A - L^T K^{-1} L)x + (R_1^{*T} A - K^{-1} N L^T)x + c_1 + \bar{\sigma}^2 \text{tr}(Q_1^*) - K^{-1} N^2/4$$

e, para  $L^T x + N/2 \in [l_1, l_2], \forall x \in \mathcal{X} \subseteq \mathbb{R}$ , o erro de aproximação máximo em cada passo será

$$e_k^{\max}(l_1, l_2) = K^{-1}M \max\{|2l_1 - M|, |-2l_2 - M|\}$$

A vantagem da aproximação por truncamento em relação a aproximação por mínimos quadrados é que ela não requer a escolha de um intervalo de aproximação  $[l_1, l_2]$ . Escolher um intervalo adequado não é necessariamente uma tarefa simples. Por exemplo, se um intervalo grande de aproximação for escolhido, isto será equivalente a supor que pelo menos algumas componentes do estado estarão distantes da origem. Este intervalo “largo” pode ser adequado se o estado inicial do sistema estiver longe da origem mas, a medida que o sistema é regulado, tendendo ao ponto de equilíbrio, esta aproximação pode rapidamente tornar-se imprecisa. Este problema torna-se ainda mais grave considerando que o sistema possui uma componente estocástica. Idealmente, o intervalo de aproximação deveria ser dependente do tempo e do valor do estado, levando a um erro pequeno em regiões onde o estado possui uma probabilidade alta de estar localizado.

### 4.3 Programação Dinâmica Linear

Para um sistema dinâmico avaliado através de um funcional de custo convexo, existem diversas estratégias para aproximação da política ótima. Nesta Seção será apresentada uma aproximação para o custo ótimo em um horizonte infinito, proposta por D. P. Farias e B.V. Roy em [8], que consiste em transformar o problema de obtenção do custo ótimo em um problema de Programação Linear. Como apresentado no Capítulo 2, o problema de controle estocástico para um horizonte de otimização infinito é:

$$\begin{aligned} \min_{\pi \in \Pi} J(x_0, \pi) &= \lim_{N \rightarrow \infty} E \left[ \sum_{k=0}^N \alpha^k C(x_k, u_k) \right] \\ \text{s.a. } x_{k+1}(x_k, u_k) &= f(x_k, u_k, \varepsilon_k), \quad 0 < \alpha < 1 \end{aligned}$$

O Custo Ótimo para este problema é obtido resolvendo a Equação de Bellman  $J^* = TJ^*$ . Conforme discutido em [1], como

$$\lim_{N \rightarrow \infty} T^N J = J^*,$$

se o espaço de estados,  $S$ , ações,  $U$ , e distúrbios,  $D$ , forem finitos e limitados. Para obter a política ótima, basta resolver o problema de otimização

$$\begin{aligned} \max_J \sum_{x \in S} \kappa(x) J(x), \quad \kappa : S \rightarrow \mathbb{R}; \kappa(x) \geq 0, \forall x \in S \\ \text{s.a. } TJ \geq J, \quad \forall x \in S \end{aligned} \tag{4.4}$$

A função  $\kappa$  é chamada de *função de relevância*. Observe que este problema é um programa não-linear. Para transformá-lo em um problema de otimização linear, suponha que a Função Custo possa ser aproximada por uma combinação linear de funções (não necessariamente ortogonais),  $\phi_i : S \rightarrow \mathbb{R}$ , de modo que

$$\tilde{J}(x, r) = \phi(x)^T r, \quad r \in \mathbb{R}^q \quad (4.5)$$

onde

$$\phi(x) = [\phi_1(x) \quad \phi_2(x) \quad \dots \quad \phi_q(x)]^T$$

Para este caso, o problema descrito na Equação 4.4 torna-se

$$\begin{aligned} \max_r \quad & \sum_{x \in S} \kappa(x) \phi(x)^T r \\ \text{s.a.} \quad & T\phi(x)^T r - \phi(x)^T r \geq 0, \quad \forall x \in S \end{aligned} \quad (4.6)$$

e o custo aproximado mínimo será  $\tilde{J}^*(x)$ . Como o espaço de estados,  $S$ , o de ações,  $U$ , e o de distúrbios,  $D$ , são finitos, para uma política  $u$ , a probabilidade do sistema ir do estado  $x \in S$  para o estado  $y \in S$  será denotada por  $p_{xy}(u)$ . Conseqüentemente, o sistema será um grande processo de decisão markoviano e o operador da programação dinâmica torna-se

$$\begin{aligned} TJ &= \min_{u \in U(x)} E\{C(x, u) + \alpha J(f(x, u, \varepsilon))\} \\ &= \min_{u \in U(x)} C(x, u) + \alpha \sum_{y \in S} p_{xy}(u) J(y) \end{aligned}$$

Da Equação 4.5, observe que

$$T\tilde{J} = \min_{u \in U(x)} C(x, u) + \alpha \sum_{y \in S} p_{xy}(u) \phi(y)^T r \quad (4.7)$$

Substituindo a expressão acima na Equação (4.6), obtém-se

$$\begin{aligned} \max_r \quad & \left( \sum_{x \in S} \kappa(x) \phi(x) \right)^T r \\ \text{s.a.} \quad & \min_{u \in U(x)} \left\{ C(x, u) + \alpha \sum_{y \in S} p_{xy}(u) \phi(y)^T r \right\} - \phi(x)^T r \geq 0, \quad \forall x \in S \end{aligned} \quad (4.8)$$

Desta forma, o problema de obtenção do custo mínimo torna-se um problema de Programação Linear. Entretanto, neste problema ainda há a necessidade de calcular a minimização em  $T\tilde{J}$ . Isto pode ser contornado notando que

$$\min_{u \in U(x)} \left\{ C(x, u) + \alpha \sum_{y \in S} p_{xy}(u) \phi(y) \right\} \leq C(x, u) + \sum_{y \in S} p_{xy}(u) J(y), \quad \forall u \in U$$

Logo, podemos finalmente reescrever o problema do cálculo de  $\tilde{J}^*$  como o programa linear

$$\begin{aligned} \max_r \quad & \left( \sum_{x \in S} \kappa(x) \phi(x) \right)^T r \\ \text{s.a.} \quad & C(x, u) + \left( \alpha \sum_{y \in S} p_{xy}(u) \phi(y) - \phi(x) \right)^T r \geq 0, \quad \forall x \in S, u \in U \end{aligned} \quad (4.9)$$

Apesar do grande número de restrições, este problema pode ser resolvido de forma eficiente através de métodos numéricos para programação linear em grande escala, como o método dos planos cortantes. Uma vez determinada a solução,  $r^*$ , do problema acima, para encontrar a ação ótima  $u^*(x)$  para cada estado  $x$ , basta resolver para todo  $x \in X$  o problema

$$u^*(x) = \min_{u \in U} \left\{ C(x, u) + \alpha \sum_{y \in S} p_{xy}(u) \phi(y)^T r^* \right\}$$

### 4.3.1 Exemplo escalar

Considere o seguinte problema de controle estocástico de um sistema VCAI escalar:

$$\begin{aligned} \min_{\pi \in \Pi} J(x_0, \pi) &= \lim_{N \rightarrow \infty} E \left[ \sum_{k=0}^N 0.7^k (x_k^2 + u_k^2) \right] \\ \text{s.a.} \quad x_{k+1} &= 1.2x_k + u_k + (0.5 + |u_k|)\varepsilon_k, \quad x_0 = 0 \end{aligned}$$

em que  $\varepsilon_k \sim N(0, 1)$ . O primeiro passo para resolver este problema, é discretizar o espaço de estados e de ações. Desta forma, suponha para  $l_1 < l_2$

$$(\tilde{u}_k, \tilde{x}_k) \in \bar{U} \times \bar{S} = \{(u, x) : l_1 < u < l_2, l_1 < x < l_2, \text{ tal que } 10u, 10x \in \mathbb{Z}\}$$

Observe que estamos discretizando o espaço para múltiplos inteiros de  $10^{-1}$ . Além disso, considere a variável aleatória discreta  $w_k \in \{w : 10w \in \mathbb{Z}\}$ , com distribuição de probabilidade

$$\Pr\{a < w_k < b\} = \Pr\{a < \varepsilon_k < b\}, \quad a < b$$

Como está sendo considerado um número finito de estados, suponha a equação de transição

$$\tilde{x}_{k+1} = f(\tilde{x}_k, \tilde{u}_k, w_k)$$

em que

$$f(\tilde{x}_k, \tilde{u}_k, w_k) = \begin{cases} 1.2\tilde{x}_k + \tilde{u}_k + (0.5 + |\tilde{u}_k|)w_k, & \text{se } l_1 < 1.2\tilde{x}_k + \tilde{u}_k + (0.5 + |\tilde{u}_k|)w_k < l_2 \\ l_2, & \text{se } 1.2\tilde{x}_k + \tilde{u}_k + (0.5 + |\tilde{u}_k|)w_k > l_2 \\ l_1, & \text{se } 1.2\tilde{x}_k + \tilde{u}_k + (0.5 + |\tilde{u}_k|)w_k < l_1 \end{cases}$$

É interessante notar que, com a discretização, o sistema torna-se uma grande Cadeia de Markov controlada. Conseqüentemente, será resolvido o problema de otimização discreto

$$\begin{aligned} \min_{\tilde{\pi} \in \tilde{\Pi}} J(\tilde{x}_0, \tilde{\pi}) &= \lim_{N \rightarrow \infty} E \left[ \sum_{k=0}^N 0.7^k (\tilde{x}_k^2 + \tilde{u}_k^2) \right] \\ \text{s.a. } \tilde{x}_{k+1} &= f(\tilde{x}_k, \tilde{u}_k, w_k), \quad \tilde{x}_0 = 0 \end{aligned}$$

Por simplicidade, será considerado  $l_1 = -5$  e  $l_2 = 5$ . O próximo passo é escolher a aproximação para a função custo acumulado. Como a função custo original é um polinômio de segunda ordem, será feita uma aproximação por um polinômio de sexto grau. Sendo assim, utilizando (4.5),  $\tilde{J}$  pode ser escrita como

$$\tilde{J}(x, r) = \phi(x)^T r, \quad r \in \mathbb{R}^q$$

em que

$$\phi(x) = [1 \quad x \quad x^2 \quad x^3 \quad x^4 \quad x^5 \quad x^6]^T$$

Para obter  $r^*$ , é necessário resolver o problema (4.9) que, fazendo  $\kappa(x) = 1$ , será

$$\begin{aligned} \max_{r \in \mathbb{R}^7} & \left( \sum_{\tilde{x} \in U} \phi(\tilde{x}) \right)^T r \\ \text{s.a. } & \tilde{x}^2 + \tilde{u}^2 + 0.7 \left( \sum_{y \in U} p_{\tilde{x}y}(\tilde{u}) \phi(y) - \phi(\tilde{x}) \right)^T r \geq 0, \quad \forall \tilde{x}, \tilde{u} \in U \end{aligned}$$

Neste exemplo, as probabilidades de transição foram estimadas utilizando uma simulação de Monte-Carlo. A solução,  $r^*$ , do problema foi determinada computacionalmente e resultou em

$$r^* = [1.83 \quad 0 \quad 3.1726 \quad 0 \quad -7.73 \cdot 10^{-2} \quad 0 \quad 1.3 \cdot 10^{-4}]$$

Logo,

$$\tilde{J}^*(x) = 1.83 + 3.1726x^2 - 7.73 \cdot 10^{-2}x^4 + 1.3 \cdot 10^{-4}x^6$$

A política ótima pode ser calculada através da minimização

$$\tilde{u}^*(x) = \min_{u \in U} \left\{ \tilde{x}^2 + \tilde{u}^2 + 0.7 \sum_{y \in U} p_{xy}(u) \tilde{J}^*(y) \right\}$$

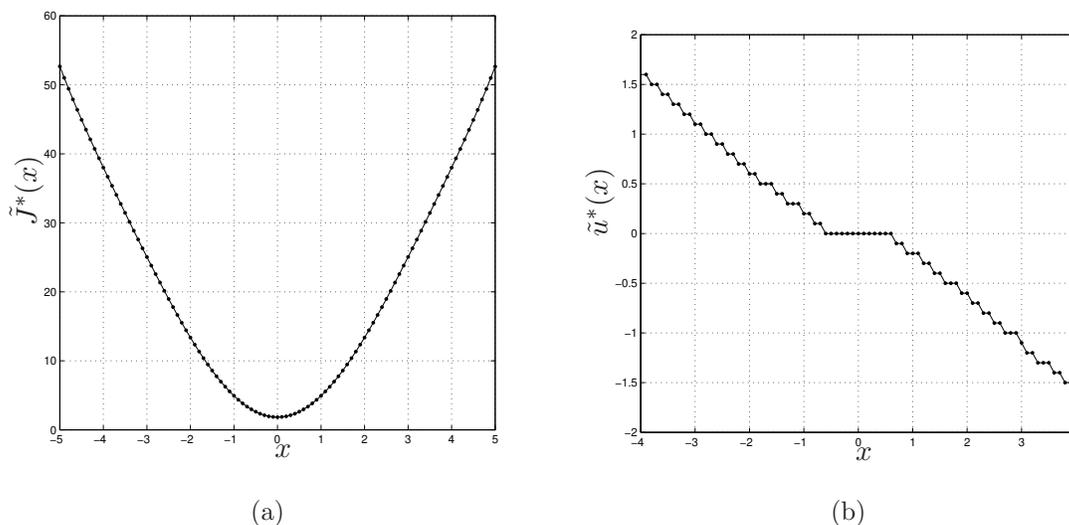


Figura 4.3: (a) Custo mínimo aproximado e (b) política ótima aproximada obtidos utilizando programação linear aproximada. A região de variação será para  $|x| \leq 0.5$

O comportamento da política ótima aproximada e do custo mínimo aproximado estão expostos na Figura 4.3. Neste caso, a região de não-variação será para  $|x| \leq 0.5$ . Note que  $\tilde{u}^*(x)$  pode ser aproximado por uma função linear por partes do estado. Fazendo  $u^*(x) \approx \tilde{u}^*(x)$ , tem-se

$$u^*(x) \approx \begin{cases} -0.55x - 0.5, & \text{se } x < -0.5 \\ 0, & \text{se } |x| \leq 0.5 \\ -0.55x + 0.5, & \text{se } x > 0.5 \end{cases}$$

# Aplicação em um problema macroeconômico <sup>1</sup>

Segundo o economista Alan Greenspan, “*uncertainty is not just an important feature of the monetary policy landscape; it is the defining characteristic of that landscape*”. Existem diversas fontes de incertezas que influenciam a definição da política econômica por um Banco Central (BC) como, por exemplo, transparência, transmissão de informação, expectativas, eficiência, dados, etc.

Na próxima seção será mostrado, a partir dos resultados obtidos nos últimos capítulos, como o problema de definição da política monetária por um Banco Central ou pelo Federal Reserve Bank (FED) pode ser modelado como um sistema VCAI. Feito isso, utilizando simulações de Monte Carlo, será analisado como o comportamento do sistema varia com mudanças no impacto gerado pela variação da magnitude do controle, observando se a política ótima torna-se mais agressiva ou conservadora. Finalmente, será mostrado que a política gerada pelo modelo VCAI representa melhor a política monetária efetivamente adotada pelo FED do que os modelos tradicionais, baseados no Regulador Linear Quadrático (RLQ) clássico.

## 5.1 O Modelo

Considere um modelo auto-regressivo da economia dado por:

$$\pi_{k+1} = \sum_{j=0}^{n_1} \alpha_j \pi_{k-j} + \delta y_k + \epsilon_{\pi,k}, \quad (5.1)$$

$$y_{k+1} = \sum_{j=0}^{n_2} \beta_j \pi_{k-j} + \sum_{j=0}^{n_3} \lambda_j y_{k-j} + \sum_{j=0}^{n_4} \gamma_j i_{k-j} + \epsilon_{y,k}, \quad (5.2)$$

com  $\pi_k$  é a inflação trimestral,  $y_k$  é o hiato do produto, i.e., a diferença entre o Produto Interno Bruto (PIB) potencial e obtido e  $i_k$  é a taxa de juros trimestral definida pelo Federal

<sup>1</sup>Este capítulo foi feito em colaboração com o Prof. Thomas Vallee do Laboratório de Economia da Universidade de Nantes

Reserve (FED), sendo medida em pontos percentual anuais. Os dois distúrbios,  $\epsilon_{\pi,k}$  e  $\epsilon_{y,k}$ , são responsáveis pela incerteza sobre o estado. Este modelo é uma generalização do modelo em [17], [19] e [20].

É interessante observar que grandes variações da taxa de juros levam a uma grande incerteza sobre o estado. Conseqüentemente, este sistema se encaixa no modelo proposto de modo que  $\epsilon_{y,k}$  será uma função da variação da taxa de juros. Denotando estas variações do instante  $k$  para o instante  $k + 1$  por  $v_k$ , obtém-se

$$\dot{i}_k = \dot{i}_{k-1} + v_k$$

Utilizando um modelo similar ao apresentado na Seção 2.1, a incerteza sobre o sistema será modelada como:

$$\epsilon_{k,\pi} = (\bar{\sigma}_\pi + \sigma_\pi |v_k|) \xi_{k,\pi} \quad (5.3)$$

$$\epsilon_{k,y} = (\bar{\sigma}_y + \sigma_y |v_k|) \xi_{k,y} \quad (5.4)$$

onde  $\xi_{k,\pi} \sim N(0, 1)$  e  $\xi_{k,y} \sim N(0, 1)$ .

Para obter a política de controle ótima, precisamos colocar o sistema em uma forma adequada. Neste sentido, o sistema descrito pela Equação (5.2) pode ser reescrito como

$$\pi_{k+1} = \sum_{j=0}^{n_1} \alpha_j \pi_{k-j} + \delta y_k + \epsilon_{\pi,k} \quad (5.5)$$

$$y_{k+1} = \sum_{j=0}^{n_2} \beta_j \pi_{k-j} + \sum_{j=0}^{n_3} \lambda_j y_{k-j} + \gamma_0 (v_k + \dot{i}_{k-1}) + \sum_{j=1}^{n_4} \gamma_j \dot{i}_{k-j} + \epsilon_{y,k} \quad (5.6)$$

Este sistema pode ser reescrito na forma de espaço de estados como:

$$x_{k+1} = Ax_k + Bv_k + \mathcal{E}_k(v_k), \quad (5.7)$$

onde

$$\begin{aligned} n_{max} &= \max\{n_1, n_2\} \\ x_k &\equiv [\pi_k \ \dots \ \pi_{k-n_{max}} \ y_k \ \dots \ y_{k-n_3} \ \dot{i}_{k-1} \ \dots \ \dot{i}_{k-n_4}]^T, \\ \mathcal{E}_k(v_k) &\equiv [\epsilon_{\pi,k+1} \ \mathbf{0}_{1 \times n_{max}} \ \epsilon_{y,k+1} \ \mathbf{0}_{1 \times (n_3+n_4)}]^T. \end{aligned}$$

Definindo:

$$\begin{aligned} a &:= [\alpha_0 \ \dots \ \alpha_{n_1}]^T, \quad b := [\beta_0 \ \dots \ \beta_{n_2}]^T, \\ c &:= [\lambda_0 \ \dots \ \lambda_{n_3}]^T, \quad d := [\gamma_0 + \gamma_1 \ \gamma_2 \ \dots \ \gamma_{n_4}]^T, \end{aligned}$$

a matriz de transição  $A$  será dada por:

$$A \equiv \begin{bmatrix} A_{11} & A_{12} & \mathbf{0}_{n_{max}+1 \times n_4} \\ A_{21} & A_{22} & A_{23} \\ \mathbf{0}_{n_4 \times n_{max}+1} & \mathbf{0}_{(n_4) \times (n_3+1)} & A_{33} \end{bmatrix}$$

onde

$$A_{11} = \begin{cases} \begin{bmatrix} a^T \\ [\mathbf{I}_{n_1 \times n_1} & \mathbf{0}_{n_1 \times 1}] \end{bmatrix}, & \text{se } n_{max} = n_1 \\ \begin{bmatrix} a^T & \mathbf{0}_{1 \times n_2 - n_1} \\ [\mathbf{I}_{n_2 \times n_2} & \mathbf{0}_{n_2 \times 1}] \end{bmatrix}, & \text{se } n_{max} = n_2 \end{cases}$$

e

$$A_{21} = \begin{cases} \begin{bmatrix} [b^T & \mathbf{0}_{1 \times n_1 - n_2}] \\ \mathbf{0}_{n_3 \times n_1 + 1} \end{bmatrix}, & \text{se } n_{max} = n_1 \\ \begin{bmatrix} b^T \\ \mathbf{0}_{n_3 \times n_2 + 1} \end{bmatrix}, & \text{se } n_{max} = n_2 \end{cases}$$

Os outros termos de  $A$  serão:

$$A_{12} = \begin{bmatrix} [\delta & \mathbf{0}_{1 \times n_3}] \\ \mathbf{0}_{n_1 \times n_3 + 1} \end{bmatrix}, \quad A_{22} = \begin{bmatrix} c^T \\ [\mathbf{I}_{n_3 \times n_3} & \mathbf{0}_{n_3 \times 1}] \end{bmatrix}, \\ A_{23} = \begin{bmatrix} d^T \\ \mathbf{0}_{n_3 \times n_4} \end{bmatrix}, \quad A_{33} = \begin{bmatrix} [1 & \mathbf{0}_{1 \times n_4 - 1}] \\ [\mathbf{I}_{n_4 - 1 \times n_4 - 1} & \mathbf{0}_{n_4 - 1 \times 1}] \end{bmatrix}$$

A matriz  $B$  será dada por

$$B \equiv [\mathbf{0}_{1 \times n_{max}+1} \quad \gamma_0 \quad \mathbf{0}_{1 \times n_3} \quad 1 \quad \mathbf{0}_{1 \times (n_4-1)}]^T$$

Neste exemplo, os objetivos do tomador de decisões são reduzir a inflação e o hiato do produto em um período de  $N$  trimestres. Conforme colocado em [17] e [19], supõe-se que o seguinte critério deva ser minimizado a cada período  $k$

$$L = \pi_k^2 + y_k^2 + \frac{1}{2}(i_k - i_{k-1})^2 \quad (5.8)$$

A função custo pode ser reescrita em função de  $v$ , levando a

$$J(x) = E \left[ x'_N Q x_N + \sum_{k=0}^{N-1} x'_k Q x_k + 2P x_k v_k + r v_k^2 \right] \quad (5.9)$$

A partir de (5.7) e (5.9) podemos utilizar o Teorema 1 diretamente para obter a política de controle ótimo.

### 5.1.1 O Modelo no Espaço de Estados

Com o objetivo de simular o modelo, suponha que os parâmetros são dados pela Tabela 1, com  $n_1 = 2$ ,  $n_2 = 2$ ,  $n_3 = 0$  e  $n_4 = 1$ . Estes parâmetros foram obtidos de [16]. Foi utilizado um modelo trimestral para a economia americana, do primeiro trimestre de 1960 até o quarto trimestre de 2006. A taxa de juros ( $i_k$ ) é a média de quatro trimestres do conselho de governantes (Board of Governors). A inflação ( $\pi_k$ ) é medida pelo índice encadeado em um ano, i.e.,  $400(\ln p_k - \ln p_{k-1})$ . O hiato do produto ( $y_k$ ) é definido como  $100(q_k - q_k^*)/q_k^*$ , onde  $q_k$  é o PIB real e  $q_k^*$  é o PIB potencial. Os dados utilizados foram obtidos do “Bureau of Economic Analysis”, órgão do Departamento de Comércio do governo dos EUA. A média de todas as variáveis foram removidas antes da estimação.

Tabela 5.1: Estimativas para o modelo trimestral da economia americana.

Parâmetros	Estimativa	Parâmetros	Estimativa
$\alpha_0$	0.621	$\beta_0$	0.041
$\alpha_1$	0.091	$\beta_1$	-0.083
$\alpha_2$	0.239	$\beta_2$	0.011
$\delta$	0.108	$\gamma_0$	0.117
$\lambda$	0.900	$\gamma_1$	-0.192

Utilizando os parâmetros da tabela, o sistema considerado é:

$$x_{k+1} = Ax_k + Bv_k + (\bar{\Sigma} + \Sigma|v_k|)\xi_k, \quad \xi_k \sim N(0, I_{5 \times 5}) \quad (5.10)$$

$$x_k = [\pi_k \quad \pi_{k-1} \quad \pi_{k-2} \quad y_k \quad i_{k-1}]^T, \quad v_k = [i_k - i_{k-1}],$$

$$A = \begin{bmatrix} 0.621 & 0.091 & 0.239 & 0.108 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0.041 & -0.083 & 0.011 & 0.900 & (0.117 - 0.192) \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix},$$

$$B = [0 \quad 0 \quad 0 \quad 0.117 \quad 1]^T,$$

onde  $\bar{\sigma}$  e  $\sigma$  são matrizes diagonais:

$$\bar{\Sigma}_k = \text{diag}(\bar{\sigma}_\pi, 0, 0, \bar{\sigma}_y, 0)$$

$$\Sigma_k = \text{diag}(\sigma_\pi, 0, 0, \sigma_y, 0)$$

Finalmente, o sistema é avaliado através da função custo (5.9) com

$$Q = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix};$$

$$P = [0 \ 0 \ 0 \ 0 \ 0 \ 0]^T;$$

$$r = 0.5.$$

## 5.2 Experimentos Numéricos

Serão realizados dois experimentos numéricos. No primeiro, será analisado o impacto das variações de  $\sigma$  para um dado estado inicial  $x_0$ . Os resultados obtidos serão comparados com o Regulador Linear Quadrático (RLQ) tradicional. O segundo experimento consistirá em comparar os resultados numéricos com dados reais.

Para todas as simulações, será imposta a condição  $i(k) \geq 0, \forall k$ . Desta forma, se em algum instante  $k$  a variação de  $v_k$  é tal que  $i_k = v_k + i_{k-1} < 0$  então  $v_k = -i_{k-1}$ . Logo, o decréscimo máximo da taxa de juros é fixado como a diferença entre a taxa de juros anterior e 0. A cada passo da programação dinâmica, será utilizada uma aproximação por mínimos quadrados da função custo, conforme descrito na Seção 4.2.1. Este tipo de aproximação permite uma fácil implementação computacional e um bom controle do erro.

### 5.2.1 LQR versus VCAI

O estado inicial para as simulações será  $x_0 = [2.2781 \ 1.0313 \ 1.0528 \ -0.6582 \ 2.3375]^T$ , que são os valores históricos para o primeiro trimestre de 1962. Será observado o impacto do aumento da incerteza relacionada a variações da magnitude do controle sobre o comportamento do sistema. Como colocado anteriormente, quando a incerteza gerada pela variação da entrada (representada por  $\sigma$ ) aumenta, a região no espaço de estados onde  $v_k^* = 0$  também aumentará. Analogamente, se  $\sigma = 0$  o problema torna-se o caso do RLQ tradicional, pois a variação do controle não influenciará a incerteza sobre o sistema, e a política ótima de variação do controle será uma função linear do valor do estado. De forma a analisar a influência do aumento da incerteza, supõe-se que os distúrbios afetam igualmente as equações (5.5) e (5.6), ou seja,

$$\bar{\sigma}_k = \text{diag}(\bar{\sigma}, 0, 0, \bar{\sigma}, 0)$$

$$\sigma_k = \text{diag}(\sigma, 0, 0, \sigma, 0)$$

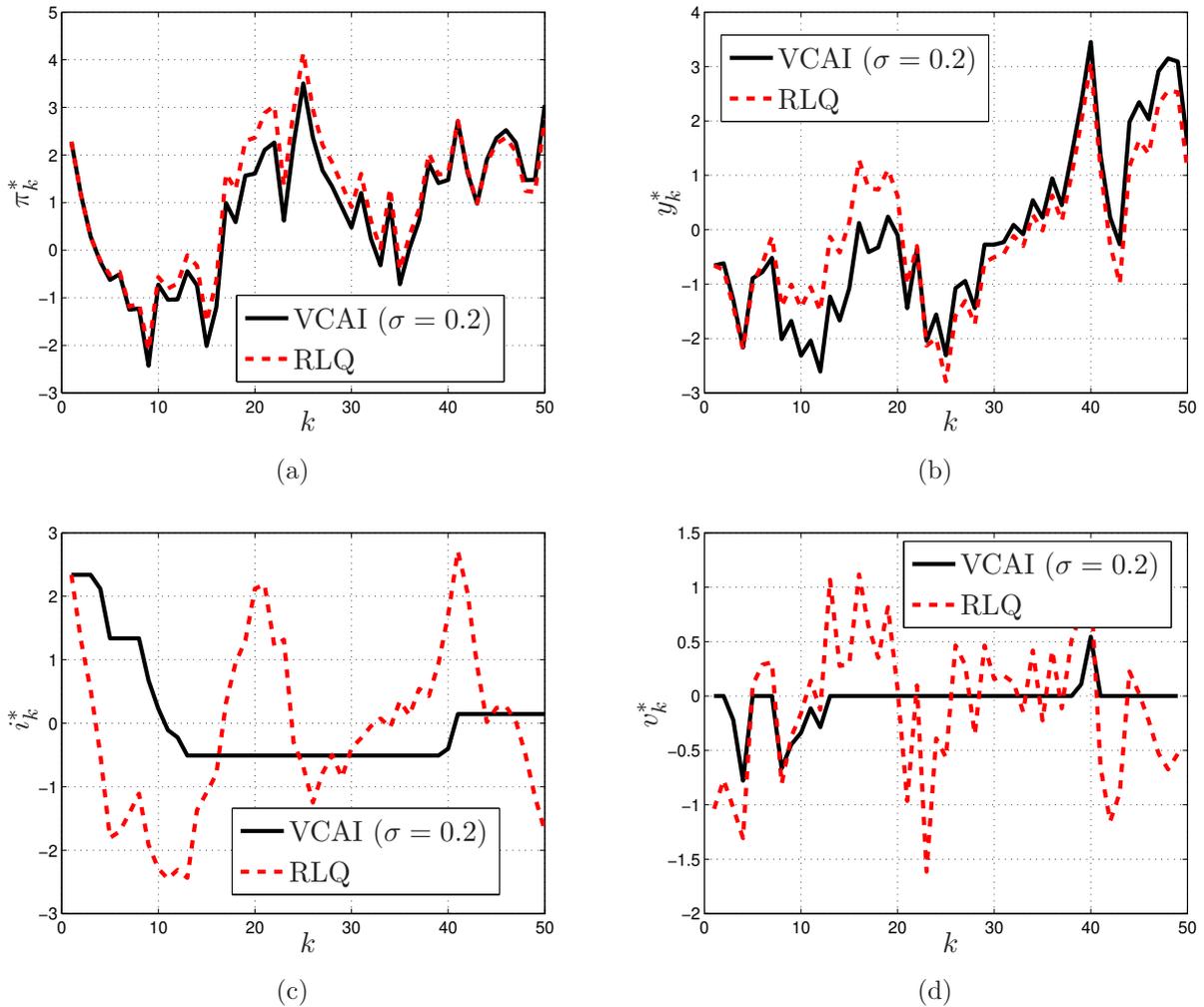


Figura 5.1: Comportamento (a) da inflação,  $\pi_k$ , (b) do PIB,  $y_k$ , da taxa de juros ótima,  $i_k^*$  e (c) da variação ótima da taxa de juros,  $v_k^*$ , para o modelo VCAI e para o RLQ No modelo VCAI,  $\sigma_\pi = \sigma_y = 0.2$ . Em ambos os casos,  $\bar{\sigma}_\pi = \bar{\sigma}_y = 0.2$ .

onde  $\sigma$  e  $\bar{\sigma}$  são escalares constantes.

Realizações do sistema foram simuladas para  $\sigma = 0.2$  e  $\sigma = 0$  considerando  $\bar{\sigma} = 0.2$ . Os resultados estão expostos na Figura 5.1. No caso onde variações da magnitude do controle aumentam a incerteza sobre o sistema ( $\sigma = 0.2$ ) existirão vários instantes de tempo onde  $v_k^* = 0$ , levando a um comportamento mais suave da taxa de juros. Além disso, a política monetária ótima para o sistema VCAI consiste em realizar apenas pequenas variações da taxa de juros ( $(|v_k^*| \leq 0.8)$ ), o que leva a um decréscimo suave da taxa de juros na direção da origem. Por outro lado, a simulação do regulador linear quadrático indica grandes variações da taxa de juros. Desta forma, para  $\sigma = 0$ , a taxa de juros foi alterada em quase todos os instantes. Fazendo  $\sigma = 0.1$ , como ilustrado na Figura 5.2, a taxa de juros é alterada mais

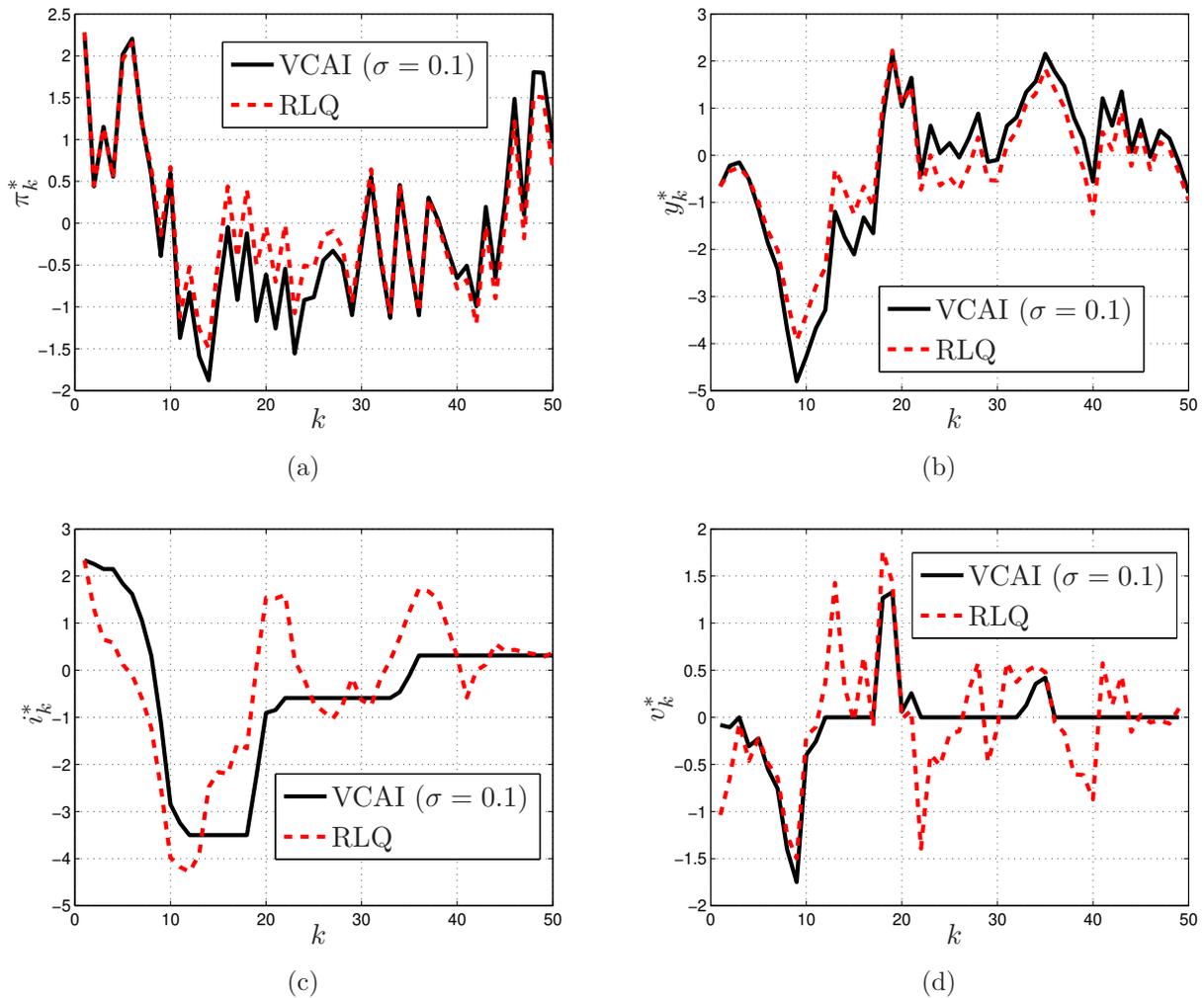


Figura 5.2: Comportamento (a) da inflação,  $\pi_k$ , (b) do PIB,  $y_k$ , da taxa de juros ótica,  $i_k^*$  e (c) da variação ótica da taxa de juros,  $v_k^*$ , para o modelo VCAI e para o RLQ. No modelo VCAI,  $\sigma_\pi = \sigma_y = 0.1$ . Em ambos os casos,  $\bar{\sigma}_\pi = \bar{\sigma}_y = 0.2$ .

frequentemente, pois a região de não-variação será menor.

### 5.2.2 Comparação com Dados Reais

Considerando um horizonte entre 1960 e 2006, serão comparadas para cada instante de tempo,  $k$ , a estimativa da variação da taxa de juros pelo modelo VCAI e pelo modelo RLQ tradicional. Além disso, as duas estimativas serão comparadas com dados históricos. Para tanto, foram considerados dois modelos VCAI: o primeiro com  $\sigma = 0.3$  e o segundo com  $\sigma = 0.8$ . Para os dois modelos VCAI e para o RLQ, supõe-se  $\bar{\sigma} = 0.8$ . Todas as simulações utilizaram a mesma sequência de vetores aleatórios gaussianas  $\{\varepsilon_k\}$ . Os resultados estão expostos na Figura 5.3.

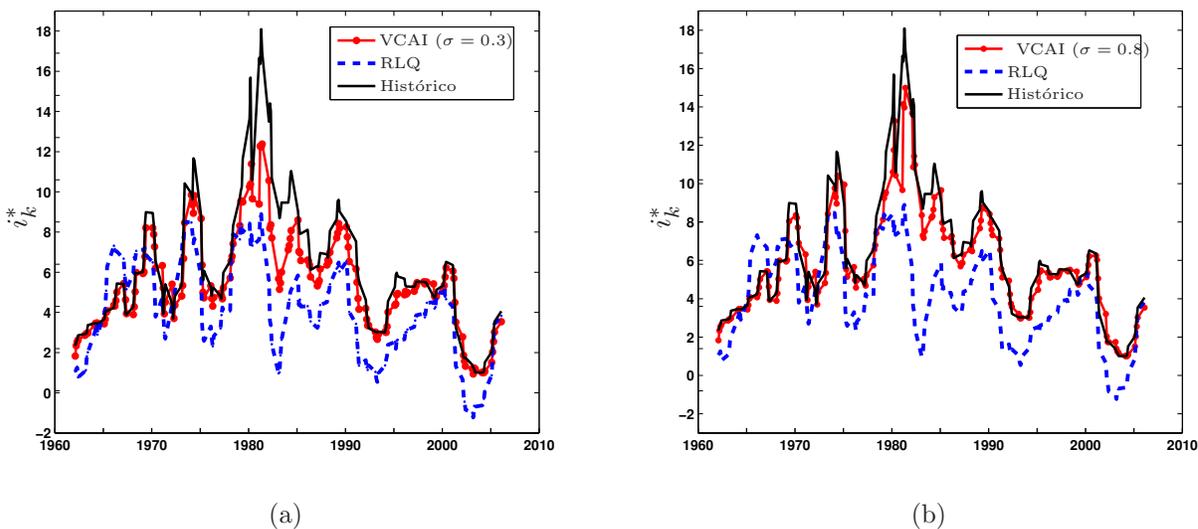


Figura 5.3: Curvas do valor histórico da taxa de juros, da taxa de juros gerada pelo modelo RLQ tradicional ( $\bar{\sigma} = 0.8$  e  $\sigma = 0$ ), e a taxa de juros obtida utilizando um modelo VCAI com  $\bar{\sigma} = 0.8$  e (a)  $\sigma = 0.3$ ; (b)  $\sigma = 0.8$ .

Note que a política obtida utilizando o RLQ possui desempenho pior do que o sistema VCAI para os dois valores de  $\sigma$ . O erro quadrático médio (EQM) entre a taxa de juros histórica e a taxa de juros gerada pelo RLQ foi de 10.62. Já o sistema VCAI com  $\sigma = 0.3$  apresentou um EQM de 2.38. Além disso, a taxa de juros gerada pelo modelo VCAI com  $\sigma = 0.8$  é o que melhor reflete as decisões reais do FED, possuindo um EQM de 1.19. Este resultado indica que de fato existe uma grande incerteza associada a variações da política monetária conduzida pelo Federal Reserve Bank e que, conseqüentemente, o FED tende a ser cauteloso ao variar a sua política monetária.

## Conclusões e Perspectivas

Neste trabalho foi desenvolvido um arcabouço teórico e numérico para o tratamento de sistemas estocásticos onde a variação da magnitude da entrada de controle aumenta a incerteza (sistemas VCAI).

Primeiramente, foi proposto um modelo estocástico em tempo discreto com uma componente linear determinística e uma componente estocástica para esta classe de sistemas. A porção estocástica do sistema foi modelada como uma variável aleatória cuja variância é uma função da variação da entrada controle. A influência da variação do controle sobre a incerteza do sistema foi dimensionada através de uma função contínua, positiva e não diferenciável na origem, chamada de “função de magnitude”. O modelo foi criado da forma mais geral possível, permitindo uma grande flexibilidade para a sua utilização. Uma vez desenvolvido o modelo, o problema de controle ótimo para um sistema VCAI avaliado através de uma função custo convexa em um horizonte de tempo finito foi definido e foram apresentadas diversas ferramentas teóricas úteis para auxiliar na determinação da política ótima de controle. Basicamente, foram utilizados resultados da teoria de funções contínuas porém não diferenciáveis, além de resultados clássicos da análise convexa. Estas ferramentas forneceram a flexibilidade necessária para caracterizar diversas propriedades da função custo, mesmo ela possuindo um termo não diferenciável.

O próximo passo foi extrair características da função custo e da política ótima que se preservavam ao longo do horizonte da Programação Dinâmica. Provou-se que a convexidade da função custo acumulado se preserva ao longo de todo horizonte de otimização. Além disso, foi mostrado que se a função custo a cada instante for diferenciável e estritamente convexa, a função custo acumulado também será diferenciável e convexa. Foi proposto, também, a existência de uma *região de não-variação*, ou seja, uma região de medida não nula no espaço de estados onde a ação ótima é de não variação do controle. Todos estes resultados foram sintetizados no Teorema 3.4.1, que apresenta um algoritmo de programação dinâmica para

sistemas VCAI.

Baseado nas propriedades analíticas dos sistemas VCAI, foram discutidas estratégias numéricas para o cálculo da política ótima. Através do Lema 4.1.1 foi constatado que o erro máximo de aproximação seria simplesmente a soma do erro máximo de aproximação a cada passo. O sistema VCAI avaliado através de uma função custo quadrática foi analisado em detalhes, e dois tipos de aproximação para a função custo foram propostas. Além disso, foi obtida uma expressão analítica para o erro máximo em cada passo para ambas as aproximações, permitindo um controle direto do erro da aproximação. Foi apresentado, também, uma estratégia para aproximação da política ótima para um problema de controle com horizonte infinito utilizando Programação Linear Aproximada. Esta estratégia foi ilustrada através de um exemplo numérico baseado em um sistema VCAI escalar.

Finalmente, a teoria desenvolvida foi utilizada para modelar o problema da condução da política monetária pelo Banco Central. Neste caso, a entrada do sistema é a taxa de juros e o objetivo do BC é minimizar a taxa de inflação e a diferença entre o PIB potencial e o real. A partir de estimativas reais, o processo decisório foi descrito como um sistema VCAI avaliado através de uma função custo convexa. Através de simulações, foi observado o efeito do aumento da incerteza associada à variação do controle sobre o comportamento do sistema e à política ótima. Notou-se, também, que uma formulação através de sistemas VCAI representa melhor as decisões reais do Federal Reserve (FED) do que os modelos baseados no Regulador Linear Quadrático tradicionalmente encontrados na literatura.

## Perspectivas

O propósito deste trabalho foi apenas lançar as bases para o estudo de sistemas VCAI. Como acredita-se ser um trabalho que traz uma contribuição inovadora no tratamento do tipo de incerteza considerada, ainda existem diversas questões em aberto que seriam contribuições relevantes para a área. Dentre as questões teóricas que comporiam essa agenda, destacam-se:

- Análise da variação da região de não-variação com o horizonte de otimização: não está claro se, para um sistema e função custo invariantes no tempo, a região de não-variação aumenta ou diminui com o horizonte de otimização;
- Comportamento da ação ótima de controle fora da região de não-variação: é interessante estudar se um aumento na incerteza associada a variações do controle leva à políticas mais agressivas fora da região de não-variação;
- Estabilidade: analisar a estabilidade do sistema VCAI sujeito às políticas obtidas utilizando as aproximações para a função custo acumulado descritas no Capítulo 4.

Os desafios numéricos para a obtenção da política ótima em sistemas VCAI fazem parte de um contexto mais geral: o problema do cálculo da política ótima para sistemas estocásticos com espaço de estados e ação contínuos. Devido a enorme importância desta questão, ela possui o potencial de se tornar no principal foco de trabalhos futuros. Neste sentido, para um problema de otimização em horizonte finito, a linha de pesquisa mais promissora consiste em aproximar a função custo utilizando *Support Vector Regression*, que é um método conhecido há vários anos na área de aprendizagem estatística.

Também serão exploradas mais aplicações práticas de sistemas VCAI. A principal e mais promissora aplicação desta classe de sistemas é em problemas macroeconômicos. Esta linha continuará sendo explorada de forma mais aprofundada, com o objetivo de permitir a fácil adoção de modelos VCAI por pesquisadores, analistas e gestores de política macroeconômica, em especial a política monetária. Entretanto, também serão estudadas outras aplicações deste tipo de sistema, como na medicina.

## Publicações

O estudo de sistemas VCAI mono entrada resultou em uma publicação no *American Control Conference - ACC 2009* [5]. Estão sendo preparadas duas publicações para periódicos internacionais. A primeira será submetida para um periódico na área de controle e a segunda será submetida para um periódico na área de econometria.

# Bibliografia

- [1] Dimitri P. Bertsekas. *Dynamic Programming & Optimal Control, Vol. 2*. Athena Scientific, 3rd edition, 2005.
- [2] Dimitri P. Bertsekas. *Dynamic Programming & Optimal Control, Vol. I*. Athena Scientific, 3rd edition, 2005.
- [3] Dimitri P. Bertsekas and Steven E. Shreve. *Stochastic Optimal Control: The Discrete-Time Case*. Athena Scientific, 1 edition, 2007.
- [4] Stephen Boyd and Lieven Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.
- [5] A. Calmon and J.B.R. do Val. Control variation as a source of uncertainty: Single input case. In *ACC 2009 - American Control Conference*, Accepted.
- [6] Frank H. Clarke. Generalized gradients and applications. *Transactions of the American Mathematical Society*, 205:247–262, 1975.
- [7] Frank H. Clarke. *Optimization and Nonsmooth Analysis*. Society for Industrial Mathematics, 1987.
- [8] D. P. de Farias and B. Van Roy. The linear programming approach to approximate dynamic programming. *OPERATIONS RESEARCH*, 51(6):850–865, 2003.
- [9] Joao B. R. do Val and Tamer Basar. Receding horizon control of jump linear systems and a macroeconomic policy problem. *Journal of Economic Dynamics and Control*, 23(8):1099–1131, 1999.
- [10] O.L. Jacobs, Y.P. Liu, and H.J. McQuay. Modeling and estimation for patient controlled analgesia of chronic pain. 42(5):477–485, 1995.

- 
- [11] A.M. Jeffrey, Xiaohua Xia, and I.K. Craig. When to initiate HIV therapy: a control theoretic approach. *50(11):1213–1220*, 2003.
- [12] A.A. Julius, A. Halasz, M.S. Sakar, H. Rubin, V. Kumar, and G.J. Pappas. Stochastic modeling and control of biological systems: The lactose regulation system of escherichia coli. *53(Special Issue):51–65*, 2008.
- [13] Hiroaki Kitano. Systems biology: A brief overview. *Science*, 295(5560):1662–1664, 2002.
- [14] Alexey S. Matveev and Andrey V. Savkin. Application of optimal control theory to analysis of cancer chemotherapy regimens. *Systems & Control Letters*, 46(5):311–321, 2002.
- [15] Alexei Onatski and James H. Stock. Robust monetary policy under model uncertainty in a small model of the U.S. economy. *Macroeconomic Dynamics*, 6(01):85–110, 2002.
- [16] Sophie Pardo, Nicolas Rautureau, and Thomas Vallée. Optimal versus realized policy rules in a regime-switching framework. Workgin Paper, LEMNA, IEMN-IAE, University of Nantes, France., 2008.
- [17] Glenn D. Rudebusch and Lars E. O. Svensson. Policy rules for inflation targeting. *SSRN eLibrary*, April 1998.
- [18] J.A.M. Felipe De Souza, M.A.L. Caetano, and T. Yoneyama. Optimal control theory applied to the anti-viral treatment of AIDS. In *Proceedings of the IEEE Conference on Decision and Control*, volume 5, pages 4839–4844, 2000.
- [19] Lars Svensson and Noah Williams. Monetary policy with model uncertainty: Distribution forecast targeting. *National Bureau of Economic Research Working Paper Series*, (11733), 2005.
- [20] Lars E.O. Svensson and Noah Williams. Optimal monetary policy under uncertainty in dsge models: A markov jump-linear-quadratic approach. *National Bureau of Economic Research Working Paper Series*, No. 13892, 2008.