



INSTITUTO
SUPERIOR
TÉCNICO

UNIVERSIDADE TÉCNICA DE LISBOA
INSTITUTO SUPERIOR TÉCNICO

Métodos de Controlo Probabilístico

Miguel José Simões Barão
(Mestre)

Dissertação para obtenção do Grau de Doutor em
Engenharia Electrotécnica e de Computadores

Orientador: Doutor João Manuel Lage de Miranda Lemos

Júri

Presidente: Reitor da Universidade Técnica de Lisboa

Vogais: Doutora Maria de Fátima da Silva Leite
Doutor João Manuel Lage de Miranda Lemos
Doutor Jorge dos Santos Salvador Marques
Doutor Luís Miguel de Mendonça Rato
Doutor João Manuel de Freitas Xavier

Setembro de 2008

Resumo

Esta tese aborda o problema do controlo estocástico não linear de um ponto de vista puramente probabilístico em tempo discreto. O controlo de sistemas estocásticos é geralmente modelado por um sistema estocástico e um controlador determinístico. Embora este tipo de modelação seja apropriado em muitas situações, outras existem em que esta não é a melhor abordagem.

Os sistemas de controlo com perturbações à entrada podem ser modelados por um controlador probabilístico, em que o sinal de actuação é uma variável aleatória. A natureza da perturbação pode ser diversa (*e.g.* ruído aditivo, ganho aleatório, *etc.*). No controlo probabilístico, a variável manipulada gerada pelo controlador não é um valor numérico mas uma distribuição de probabilidade.

Os sistemas de controlo *randomizados* têm também a sua representação natural como controladores probabilísticos. Este tipo de controladores tem a sua aplicação natural nos problemas em que existem vários actuadores num mesmo sistema, sem comunicação entre si, e onde a informação necessária à actuação é difundida de igual modo para todos. Nesta situação, uma decisão determinística onde todos os agentes são forçados a actuarem da mesma maneira, pode não gerar a melhor solução.

As contribuições desta tese dizem respeito à formulação e resolução de problemas de optimização e de controlo, num contexto puramente probabilístico.

São introduzidas ferramentas de geometria de informação que permitem resolver os problemas de optimização usando a geometria intrínseca aos espaços considerados. Nesta geometria, uma família paramétrica de distribuições de probabilidade é vista como uma variedade diferencial, onde os pontos são distribuições de probabilidade. Esta variedade é equipada com uma métrica Riemanniana definida pela matriz de informação de Fisher. São introduzidas duas conexões (não Riemannianas) que permitem definir geodésicas de forma muito simples em famílias mistura e famílias exponenciais de distribuições de probabilidade. É introduzida a divergência de Kullback-Leibler, e mostra-se que esta satisfaz um teorema de Pitágoras generalizado. Os resultados anteriores são então aplicados à resolução de problemas de optimização nas famílias mistura e exponencial.

É desenvolvido um algoritmo iterativo de optimização para distribuições discretas que faz uso das propriedades geométricas do espaço para, intrinsecamente, garantir a satisfação das restrições. São fornecidas fórmulas de cálculo com complexidade linear no tempo e no espaço. Em determinados funcionais de custo, como é o caso da divergência de Kullback-Leibler, o algoritmo proposto comporta-se como um algoritmo *quasi*-Newton. É ainda proposto um critério de paragem do algoritmo e são obtidos os intervalos admissíveis para os passos das iterações. São apresentados exemplos ilustrativos do método em problemas de minimização e maximização.

É formulado o controlo probabilístico em famílias paramétricas de distribuições de probabilidade. São consideradas variáveis aleatórias contínuas e discretas. No primeiro caso, mostra-se que formulação probabilística linear Gaussiana é consistente com a solução de um problema linear quadrático determinístico. No caso discreto, é aplicado o algoritmo iterativo ao controlo probabilístico de cadeias de Markov em problemas de dimensão muito elevada, mostrando-se a viabilidade do método.

Palavras chave: Controlo Estocástico, Geometria da Informação, Probabilidade, Optimização, Geometria Diferencial, Teoria da Informação.

Abstract

This thesis studies the discrete-time stochastic nonlinear control problem from a pure probabilistic viewpoint. Stochastic control is classically modeled by a stochastic system and a deterministic controller. While this formulation is appropriate in several situations, there are others in which it is not the best approach.

Control systems with input disturbances can be modeled by a probabilistic controller, where the actuating signal is a random variable. Disturbances can be quite diverse (*e.g.* additive noise, random gain, *etc.*). Probabilistic control has probability distributions as manipulated variable instead of a deterministic value.

Randomized control systems also have their natural representation as probabilistic controllers. This kind of control has its natural application when many controllers act simultaneously on the same system, without communication/synchronization between them, but having shared access to information (*e.g.* broadcasted information). In situations like these, a deterministic decision would produce the same response by all agents, and not achieve the best results.

The contributions of this thesis concern the formulation and solution of optimization and control problems in a purely probabilistic context.

The optimization problems are solved using tools from information geometry, which make use of the intrinsic geometry of the underlying spaces. There, a parametric family of probability distributions is viewed as a differential manifold, whose points are specific probability distributions. It becomes a Riemannian manifold by defining the Fisher information matrix as its metric. Two connections are introduced, which allow the definition of geodesics in a simple way for mixture and exponential families. It is shown that the Kullback-Leibler divergence satisfies a generalized Pythagorean relation. The previous results are applied to solving optimization problems in mixture and exponential families.

An iterative algorithm for discrete distributions is developed, which makes use of geometrical properties to satisfy the probability constraints. Computational formulas are given and shown to have linear complexity in both space and time. For some cost functions, such as the Kullback-Leibler, the proposed algorithm behaves as a *quasi*-Newton algorithm. A stopping criteria is proposed and admissible intervals for the iteration step size are given. Illustrative examples are presented both for minimization and maximization problems.

The probabilistic control is formulated in parametric families of probability distributions. Both continuous and discrete random variables are considered. In the first case, it is shown that the linear Gaussian formulation is consistent with the solution of a deterministic linear quadratic problem. In the discrete case, the iterative algorithm is applied to find the solution to the control of controlled Markov chains in highly dimensional problems.

Keywords: Stochastic Control, Information Geometry, Probability, Optimization, Differential Geometry, Information Theory.

Agradecimentos

Gostaria de agradecer profundamente ao meu orientador, o Professor João Miranda Lemos, pelo seu apoio e amizade, e pelas condições de trabalho que me proporcionou no INESC-ID sem as quais este trabalho não teria sido realizado.

Este trabalho foi realizado no âmbito do projecto “*DynaMo - Dynamical modeling, control and optimization of metabolic networks PTDC/EEA-ACR/69530/2006*”.

Nomenclatura

| | |
|--------------------------------------|---|
| t | Índice temporal. |
| $t_1 : t_2$ | Conjunto de índices $(t_1, t_1 + 1, t_1 + 2, \dots, t_2 - 1, t_2)$. |
| x | Sinal (de qualquer dimensão). |
| x_t | Valor do sinal x no instante t . |
| $x_{t_1:t_2}$ | Sinal x restrito à janela temporal de t_1 a t_2 , <i>i.e.</i> $x_{t_1:t_2} = (x_{t_1}, \dots, x_{t_2})$. |
| $\stackrel{\text{def}}{=}$ | Igualdade por definição. |
| $s(\cdot \cdot)$ | Modelo do sistema definido por uma distribuição de probabilidade condicional. |
| $c(\cdot \cdot)$ | Controlador definido por uma distribuição de probabilidade condicional. |
| $S(\cdot \cdot)$ | Modelo de referência (ideal) do sistema. |
| $C(\cdot \cdot)$ | Modelo de referência (ideal) do controlador. |
| \mathcal{X} | Conjunto, variedade ou família paramétrica de modelos. |
| $D(\cdot \cdot)$ | Divergência de Kullback-Leibler. |
| $h(X)$ ou $h(p)$ | Entropia diferencial da variável aleatória contínua $X \sim p(x)$. |
| $H(X)$ ou $H(p)$ | Entropia da variável aleatória discreta $X \sim p(x)$. |
| $I(X; Y)$ | Informação mútua entre as variáveis aleatórias X e Y . |
| $\frac{\partial}{\partial \theta^i}$ | Derivada parcial relativamente à i -ésima coordenada; i -ésimo vector tangente. |
| df | Diferencial da função f . |
| $T_\theta \mathcal{C}$ | Espaço tangente a \mathcal{C} no ponto θ . |
| $\langle a, b \rangle_g$ | Produto interno relativamente à métrica g_{ij} , ou Euclidiano por defeito. |
| $ \mathcal{X} , M $ | Cardinalidade do conjunto \mathcal{X} ou determinante da matriz M . |
| $E_p[f(x)]$ | Valor esperado relativamente à distribuição p , <i>i.e.</i> $\int p(x)f(x)dx$. |
| $\mathcal{N}(\mu, \Sigma)$ | Gaussiana com média μ e matriz de covariância Σ . |

Conteúdo

| | | |
|----------|---|-----------|
| 1 | Introdução | 1 |
| 1.1 | Motivação | 1 |
| 1.2 | Enquadramento da tese no estado da arte | 2 |
| 1.3 | Estrutura da tese | 4 |
| 1.4 | Principais contribuições da tese | 4 |
| 2 | Geometria da informação | 7 |
| 2.1 | Introdução | 7 |
| 2.2 | Variedade estatística | 8 |
| 2.3 | Vectores tangentes e métrica de Fisher | 11 |
| 2.4 | As conexões- α de Amari | 14 |
| 2.4.1 | A conexão métrica | 15 |
| 2.4.2 | A conexão exponencial | 15 |
| 2.4.3 | A conexão mistura | 17 |
| 2.5 | Divergências entre distribuições de probabilidade | 18 |
| 2.5.1 | Teorema de Pitágoras | 19 |
| 2.5.2 | Minimização da divergência de Kullback-Leibler em famílias mistura | 20 |
| 2.5.3 | Minimização da divergência de Kullback-Leibler em famílias exponenciais | 22 |
| 2.6 | Conclusão | 22 |
| 3 | Otimização de distribuições de probabilidade | 23 |
| 3.1 | Introdução | 23 |
| 3.2 | Método do gradiente natural | 24 |
| 3.2.1 | Definição de gradiente natural | 24 |
| 3.2.2 | Método do gradiente natural | 25 |
| 3.3 | Cálculo rápido para distribuições discretas | 25 |
| 3.4 | Propriedades do gradiente natural em famílias discretas | 27 |
| 3.4.1 | Pontos de estacionariedade | 27 |
| 3.4.2 | Satisfação das restrições | 28 |
| 3.5 | Versão discreta do método do gradiente natural | 32 |
| 3.5.1 | Intervalos admissíveis para o passo | 32 |
| 3.5.2 | Condição de paragem do algoritmo | 33 |
| 3.6 | Minimização da divergência de Kullback-Leibler | 34 |
| 3.7 | Gradiente natural da divergência K-L em famílias mistura | 36 |
| 3.8 | Capacidade de um canal discreto sem memória | 39 |
| 3.8.1 | Diferencial da informação mútua | 40 |
| 3.8.2 | Método do gradiente natural da informação mútua | 41 |

| | | |
|----------|---|-----------|
| 4 | Controlo probabilístico | 45 |
| 4.1 | Introdução | 45 |
| 4.2 | Modelos probabilísticos | 46 |
| 4.2.1 | Exemplos | 48 |
| 4.3 | Formulação do problema de controlo | 49 |
| 4.4 | Solução óptima do controlo probabilístico | 52 |
| 4.4.1 | Caso discreto | 52 |
| 4.4.2 | Caso contínuo | 53 |
| 4.4.3 | Extensão para horizonte temporal infinito | 54 |
| 4.4.4 | Considerações de implementação da solução discreta | 56 |
| 4.5 | Solução paramétrica do controlo probabilístico | 57 |
| 4.6 | Formulação probabilística do controlo linear Gaussiano | 60 |
| 4.7 | Interpretação geométrica | 62 |
| 4.8 | Optimização de famílias mistura de controladores | 64 |
| 4.9 | Optimização de famílias discretas de controladores | 65 |
| 4.9.1 | Solução pelo método do gradiente natural | 66 |
| 4.10 | Exemplos | 68 |
| 4.10.1 | Problema de baixa complexidade | 68 |
| 4.10.2 | Problema de complexidade elevada | 70 |
| 4.11 | Aplicação à genómica | 72 |
| 4.11.1 | Introdução | 72 |
| 4.11.2 | Redes Reguladoras de Genes | 75 |
| 4.11.3 | Modelação de Redes Reguladoras de Genes | 75 |
| 5 | Conclusão | 81 |
| 5.1 | Problemas em aberto e trabalho futuro | 82 |
| A | Apêndice ao capítulo: Geometria da informação | 83 |
| A.1 | Reparametrização e métrica de Fisher | 83 |
| A.2 | Conexão-0 de Amari | 85 |
| A.3 | Métrica induzida pela divergência K-L | 86 |
| A.4 | Conexão induzida pela divergência K-L | 86 |
| A.5 | Relação Pitagórica | 87 |
| A.6 | Projectão ortogonal sobre famílias mistura | 88 |
| B | Apêndice ao capítulo: Optimização de distribuições de probabilidade | 91 |
| B.1 | Sobre a definição de gradiente natural | 91 |
| B.2 | Amplificação de uma direcção numa transformação linear | 92 |
| B.3 | Intervalos admissíveis para o passo η no método do gradiente natural | 93 |
| B.4 | Notas sobre distribuições e vectores não normalizados | 94 |
| C | Apêndice ao capítulo: Controlo probabilístico | 97 |
| C.1 | Controlo probabilístico paramétrico | 97 |
| C.2 | Solução da formulação linear Gaussiana paramétrica | 99 |
| C.3 | Alguns resultados úteis | 102 |
| C.3.1 | Cálculo matricial | 102 |
| C.3.2 | Função densidade de probabilidade Gaussiana | 102 |
| C.3.3 | Entropia de uma distribuição Gaussiana | 102 |
| C.3.4 | Divergência de K-L entre distribuições Gaussianas | 103 |

Capítulo 1

Introdução

1.1 Motivação

A generalidade dos problemas tratados habitualmente em controlo puramente determinístico considera que as variáveis existentes – variável manipulada, saída a controlar, estado, perturbações e referência – são descritas por funções reais da variável tempo. Em cada instante, a variável manipulada assume um valor real bem definido, que corresponde ao valor a aplicar ao actuador nesse instante (ou a um vector de valores a aplicar a vários actuadores no caso de sistemas multivariáveis).

Embora a descrição determinística tenha dado origem a teorias ricas e elaboradas que encontraram aplicações de grande sucesso, a consideração de novos problemas sugeriu a necessidade de outras formulações que incluíssem a descrição da incerteza nos sistemas de controlo. Um exemplo de sistemas de controlo deste tipo inclui o uso de filtros de Kalman ou mínimos quadrados recursivos, onde as variáveis ou parâmetros estimados são frequentemente descritos por um vector de médias e uma matriz de covariância. Nestes sistemas admite-se uma descrição probabilística de parte do sistema, restrita frequentemente a modelos Gaussianos. Como caso extremo, pode admitir-se uma descrição probabilística para toda a cadeia fechada e libertar a restrição Gaussianiana nos modelos. Embora de um ponto de vista conceptual esta abordagem seja elegante, na realidade, uma boa parte dos problemas habitualmente tratados em controlo exige um valor determinístico para o actuador. Existe no entanto uma classe de problemas cuja abordagem sugere uma formulação probabilística. Nestes problemas, um número elevado actuadores cooperam para um resultado final, sem comunicação ou com comunicação degradada entre si. Para recorrer a um exemplo inspirado pela natureza, o transporte de um grão por um conjunto de formigas é feito, não através de uma sincronização perfeita dos esforços das formigas, mas como resultado de um conjunto de acções aleatórias, que somadas acabam por produzir o resultado desejado.

Um outro exemplo é o da tecnologia emergente dos microrobots, em que estes são libertados em grandes quantidades para regular uma variável num determinado ambiente (por exemplo um corpo humano). Como ilustração de um sistema deste tipo, admita-se que os microrobots são todos idênticos e cada um deles tem a capacidade de actuar ON/OFF. Não havendo comunicação entre eles e não sendo possível endereçá-los individualmente, não é possível forçar a percentagem de ON/OFFs de todo o conjunto de forma determinística. Uma maneira de o conseguir consiste em transmitir para todos a percentagem desejada p , enquanto cada um deles se liga com probabilidade p . Assim, e tal como no problema das formigas, a acção aleatória somada acaba por se aproximar do objectivo (com grande probabilidade).

Uma outra área de potencial aplicação é a da biologia celular, em particular em problemas de diferenciação de células estaminais e na pesquisa na área do cancro. As células estaminais são células sem função definida, mas que podem dar origem a células com funções específicas, como células musculares ou outros tecidos do organismo. A função de uma célula está definida no seu código genético. Assim, para que uma célula estaminal dê origem a células com funções específicas é necessário activar/inibir genes específicos. Sabe-se que esta influência se dá pelo ambiente externo a que uma célula está submetida, como contacto com células vizinhas, químicos segregados por outras células e certas moléculas no seu ambiente. Um problema que surge nesta área é precisamente produzir células com funções específicas. Este problema pode ser visto como um problema de controlo, onde se pretende levar o código genético (o estado) a uma determinada configuração desejada, actuando no seu ambiente. Dada a natureza estocástica deste tipo de problemas, estes surgem como possíveis áreas de aplicação do controlo probabilístico.

Além dos problemas anteriores, em que é considerado um número elevado de actuadores não determinísticos, também os problemas tradicionais de controlo com perturbações na entrada podem ser formulados como controladores probabilísticos. A formulação, neste caso, consiste apenas em mover a perturbação da entrada do processo para a saída do controlador.

O conjunto destes problemas sugere que todas variáveis num sistema de controlo sejam descritas, não por funções reais de variável real, mas por variáveis aleatórias cujas distribuições de probabilidade variam em função do tempo. Nesta formulação, a variável manipulada calculada pelo controlador é uma variável aleatória com uma dada distribuição de probabilidade. Esta distribuição de probabilidade tanto modela a “nuvem” de actuadores nos problemas das formigas, microrobots e células, como modela a variável manipulada somada às perturbações à entrada do processo.

Por forma a obter controladores deste tipo, é necessária uma descrição matemática conveniente. A descrição adoptada nesta tese consiste em descrever cada subsistema da cadeia fechada por uma distribuição de probabilidade condicionada. Mais precisamente, as variáveis à saída de um bloco são descritas por uma distribuição de probabilidade condicionada, onde as variáveis condicionantes são as entradas e as variáveis de estado. A evolução destes sistemas ao longo do tempo é governada pelas leis básicas das probabilidades tais como a regra da cadeia e a equação de Chapman-Kolmogorov.

Dispondo de uma descrição probabilística da cadeia fechada, o problema de controlo é formulado de modo a que as variáveis aleatórias – usualmente as entradas e saídas de um sistema – apresentem distribuições de probabilidade próximas de distribuições predefinidas e consideradas como ideais. Este objectivo é perseguido minimizando a divergência de Kullback-Leibler entre a distribuição de probabilidade obtida com os modelos probabilísticos e a distribuição de probabilidade ideal.

1.2 Enquadramento da tese no estado da arte

A linha de trabalho probabilística, que surge inicialmente em [21], formula o problema de optimização para variáveis aleatórias contínuas num horizonte temporal finito. A solução encontrada é uma equação explícita envolvendo operadores integrais. Nesse trabalho, mostra-se que a formulação probabilística do caso linear Gaussiano tem solução analítica e que esta é coincidente com a solução LQ clássica. Exceptuando alguns casos particulares, como o LQ, as fórmulas apresentadas não têm solução analítica.

Em [27] é apresentada uma versão do controlo probabilístico para cadeias de Markov finitas. Nesse trabalho, as soluções obtidas são análogas às do caso anterior, com variáveis aleatórias contínuas, onde as funções densidade de probabilidade são substituídas por funções

de probabilidade e os integrais são substituídos por somatórios. O facto de as variáveis poderem tomar valores de um conjunto finito, garante a existência de soluções analíticas tornando o problema tratável. Para além da obtenção do controlo óptimo, é também considerada a existência de incerteza paramétrica nos modelos e proposto um mecanismo de estimação Bayesiano, combinado com uma estratégia de controlo de horizonte temporal recidivo, e com o princípio de equivalência certa [28].

A abordagem seguida nesta tese para o controlo probabilístico consiste em efectuar o procedimento de optimização sobre uma família paramétrica de controladores. A restrição paramétrica gera uma solução subóptima mas computacionalmente viável *online*, o que não acontece em [21]. A motivação para considerar famílias paramétricas surge nos trabalhos de Kulhavý [19, 20]. Kulhavý debruça-se sobre o problema da estimação recursiva não linear, cuja solução exacta é computacionalmente intratável. Vários métodos têm surgido na literatura para aproximar a solução óptima (EKF, métodos de grelha ou filtros de partículas [8]). No entanto, estes métodos falham no sentido em que não são geralmente coerentes com a lei de Bayes. Em contrapartida, nos trabalhos de Kulhavý, a estimação é efectuada sobre famílias paramétricas em que os parâmetros são uma estatística reduzida dos dados. Esta estatística define uma classe de equivalência de distribuições de probabilidade *a posteriori*. Esta estatística é actualizada de maneira a manter a coerência com a lei de Bayes. A consideração de famílias paramétricas reduz problemas computacionalmente intratáveis, que requerem memória infinita, em problemas computacionalmente tratáveis com memória constante. Esta abordagem trás no entanto a desvantagem de a determinação dos parâmetros não ser em geral simples de efectuar, requerendo muitas vezes a utilização de métodos iterativos. Outro facto interessante em [20] é o formalismo usado, baseado na geometria da informação. Este formalismo permite dar uma interpretação geométrica e intuitiva aos procedimentos de actualização dos parâmetros usados.

A geometria da informação é um nome introduzido por Amari [4, 6] para designar a análise dos procedimentos estatísticos usando ferramentas de geometria diferencial. As ideias base são a identificação de uma família paramétrica de distribuições de probabilidade com uma variedade diferencial, onde os parâmetros funcionam como coordenadas, e da matriz de informação de Fisher com a métrica Riemanniana.

Nesta tese é desenvolvido um programa de trabalho semelhante a [20] e [21] para o controlo probabilístico. O projecto de um controlador probabilístico é formulado como um problema de optimização em espaços de distribuições de probabilidade, parametrizados por um número finito de parâmetros. Estes espaços, vistos como variedades estatísticas, admitem uma interpretação geométrica em que os pontos óptimos são obtidos por projecção ortogonal sobre a variedade.

Consideram-se duas famílias de distribuições de probabilidade, a família mistura e a família exponencial. Estas duas famílias têm propriedades especiais que justificam a sua utilização em contextos de optimização, como são os casos da estimação e controlo óptimo. Em particular, o teorema de Pitágoras pode ser generalizado para estas famílias substituindo a usual distância Euclidiana pela divergência de Kullback-Leibler.

Um caso particular de uma família exponencial de distribuições de probabilidade é o da família Gaussiana. Para esta família o controlo probabilístico é equivalente ao controlo linear quadrático (LQ), onde as matrizes de covariância da entrada e do estado, no controlo probabilístico, correspondem às inversas das matrizes de penalização da entrada e do estado no funcional de custo LQ. Este resultado, originalmente apresentado em [21], permite obter a mesma solução de um problema essencialmente determinístico (o controlo LQ) numa perspectiva probabilística.

Outro caso estudado diz respeito ao controlo de cadeias de Markov. O formalismo usado para o controlo probabilístico é igualmente aplicável a este tipo de sistemas. A aborda-

gem paramétrica apresenta a possibilidade de considerar restrições nos parâmetros de forma natural, o que não acontece com a solução mais geral [27].

Além dos casos anteriores, em cuja solução ótima surge explicitamente, outras situações existem em tal não acontece sendo necessário recorrer a métodos aproximados. Nesta tese, são propostos métodos de aproximação baseados na geometria das famílias de distribuições de probabilidade usadas em cada problema.

1.3 Estrutura da tese

A tese está organizada do seguinte modo:

Capítulo 1 No presente capítulo o problema a resolver é enquadrado na literatura e formulado em linhas gerais. Descreve-se a estrutura da tese e enumeram-se as principais contribuições.

Capítulo 2 Neste capítulo revêm-se os principais conceitos de geometria da informação que são usados nos desenvolvimentos subsequentes. Os espaços de distribuições de probabilidade (paramétricas) são interpretados como variedades diferenciais, com uma métrica dada pela matriz de informação de Fisher. São introduzidas as conexões de Amari e a divergência de Kullback-Leibler. Mostra-se uma versão do teorema de Pitágoras e a sua relação com problemas de otimização da divergência de Kullback-Leibler em famílias mistura e exponenciais.

Capítulo 3 Neste capítulo descrevem-se métodos de otimização de distribuições de probabilidade. É apresentado o método do gradiente natural e é feita a sua aplicação a famílias mistura. Deduzem-se as suas propriedades e é proposta uma expressão computacionalmente eficiente para o seu cálculo. São apresentados exemplos de aplicação.

Capítulo 4 Neste capítulo formula-se o problema de controlo probabilístico e propõem-se soluções para vários casos particulares. Apresenta-se um algoritmo recursivo que fornece a solução paramétrica. Esta solução é interpretada geometricamente e é comparada com as soluções existentes na literatura. O algoritmo recursivo é resolvido explicitamente para os casos particulares do controlo linear Gaussiano e para cadeias de Markov controláveis. É aplicado o algoritmo desenvolvido no capítulo 3 para obter as soluções em problemas de dimensão elevada.

Capítulo 5 Finalmente, apresentam-se conclusões e descrevem-se os problemas em aberto passíveis de investigação futura.

1.4 Principais contribuições da tese

O capítulo 2 introduz um conjunto de resultados de geometria da informação usados nos capítulos posteriores. Não são apresentados resultados teóricos novos. São, no entanto, fornecidas algumas formulações alternativas às apresentadas em [6, 26], e que se consideraram oportunas para o trabalho desenvolvido nos restantes capítulos.

Os capítulos 3 e 4 apresentam contribuições originais das quais se destacam as seguintes:

- Otimização em espaços de distribuições de probabilidade discretas pelo método do gradiente natural. Em particular,
 - formulação para o caso de distribuições discretas;

- obtenção de uma expressão computacionalmente eficiente para o cálculo do gradiente natural;
 - algoritmo de optimização, incluindo condições de paragem e tamanhos admissíveis dos passos;
 - derivação das propriedades, em particular no que diz respeito à satisfação das restrições, velocidade de convergência e complexidade computacional.
- Definição do funcional de custo para horizonte temporal infinito.
 - Formulação paramétrica do controlo probabilístico na perspectiva da geometria da informação. Sobre esta formulação foi:
 - obtida a solução para horizonte temporal finito;
 - dada uma interpretação geométrica da solução, assim como a sua relação com a solução óptima não paramétrica;
 - deduzida a solução paramétrica para o caso linear Gaussiano;
 - efectuada a aplicação do método do gradiente natural à optimização.

Capítulo 2

Geometria da informação

2.1 Introdução

Neste capítulo apresenta-se um resumo das ferramentas de geometria da informação usadas nesta tese. O tratamento de espaços de distribuições de probabilidades do ponto de vista geométrico tem tido um desenvolvimento relativamente recente, descrito nas monografias [4, 6, 26], e que consiste essencialmente na criação de uma moldura de trabalho baseada na geometria diferencial aplicada aos espaços de distribuições de probabilidade e à estatística.

A vantagem dos métodos da geometria da informação tem a ver com o facto de terem uma base matemática sólida, e apresentarem uma nova perspectiva sobre os problemas que envolvem distribuições de probabilidade.

A geometria da informação parte da definição de variedade estatística \mathcal{S} como um conjunto de modelos de distribuições de probabilidade $p_\theta(x)$, parametrizadas por um conjunto finito de parâmetros $\theta = \{\theta^1, \dots, \theta^n\}$, onde os parâmetros $\{\theta^i\}$ são vistos como um sistema de coordenadas global para \mathcal{S} .

Duas variedades tomam especial relevo e têm um tratamento especial na literatura e neste trabalho. Estas são as famílias mistura e exponencial definidas em seguida.

Dado um conjunto de distribuições $\{p_0(x), p_1(x) \dots, p_n(x)\}$, define-se uma família mistura por

$$p_\theta(x) = \sum_{i=1}^n \theta_i p_i(x) + \left(1 - \sum_{i=1}^n \theta_i\right) p_0(x), \quad (2.1)$$

onde os n parâmetros $\{\theta_1, \dots, \theta_n\}$ são as coordenadas da distribuição $p_\theta(x)$.

Por outro lado, e usando o conjunto de n distribuições $\{p_1(x) \dots, p_n(x)\}$, define-se a família exponencial por

$$p_\theta(x) = \frac{p_1^{\theta_1}(x) p_2^{\theta_2}(x) \dots p_n^{\theta_n}(x)}{\gamma(\theta)} = \exp\left(\sum_{i=1}^n \theta_i \log p_i(x) - \varphi(\theta)\right), \quad (2.2)$$

onde $\gamma(\theta)$ e $\varphi(\theta)$ são factores de normalização que garantem $\int p_\theta(x) dx = 1$.

Um vector tangente à variedade estatística num ponto θ tem também duas representações preferidas: as representações mistura e exponencial. Um vector tangente na representação exponencial escreve-se como combinação linear de vectores da base

$$\sum_{i=1}^n a^i \frac{\partial \log p}{\partial \theta^i} \quad (2.3)$$

onde a^i são as componentes do vector e $\frac{\partial \log p}{\partial \theta^i}$ são os vectores da base. Na representação mistura, os vectores da base tomam a forma de

$$\frac{\partial p}{\partial \theta^i}, \quad i = 1, \dots, n. \quad (2.4)$$

O espaço tangente à variedade \mathcal{S} num ponto θ é o espaço gerados por todos os vectores tangentes nesse ponto e representa-se por $T_\theta \mathcal{S}$. A secção 2.3 introduz estes conceitos de forma mais detalhada.

Tendo definido as noções de vector tangente e espaço tangente passa a ser possível introduzir os conceitos geométricos de ângulo e ortogonalidade. O produto interno entre vectores tangentes num ponto de \mathcal{S} é definido na secção 2.3. O produto interno permite definir o ângulo entre vectores ou curvas, norma (induzida) de um vector e distância entre dois pontos. Permite também a passagem de um vector ao seu dual e *vice-versa*, uma operação necessária para a definição de vector gradiente.

Na secção 2.4 são introduzidas as conexões- α de Amari. Uma conexão é um operador diferencial que permite calcular derivadas direccionais de campos vectoriais. Tradicionalmente, em variedades Riemannianas, é definida uma conexão que surge naturalmente da métrica. No entanto, Amari mostra que em famílias de distribuições de probabilidade existe uma família de conexões às quais se pode atribuir significado estatístico. Desta família, duas conexões tomam especial importância, pois estão intimamente ligadas a famílias de distribuições mistura e exponencial e à divergência de Kullback-Leibler.

Na secção 2.5 é definida a divergência de Kullback-Leibler. Em geral, uma divergência entre duas distribuições de probabilidade é uma medida de discrepância semelhante a uma distância, mas que não respeita todos os axiomas de distância (em particular não é simétrica). A divergência de Kullback-Leibler é um exemplo de uma divergência com vasta aplicação. Esta é definida pela fórmula

$$D(p||q) \stackrel{\text{def}}{=} \int p(x) \log \frac{p(x)}{q(x)} dx$$

onde $p(x)$ e $q(x)$ são duas distribuições de probabilidade e $D(p||q)$ é a divergência entre elas.

No contexto da geometria da informação, os problemas de optimização envolvendo a divergência de Kullback-Leibler e famílias mistura e exponencial, têm soluções que podem ser interpretadas como generalização do teorema de Pitágoras e da projecção ortogonal.

2.2 Variedade estatística

Considere-se um modelo probabilístico $p(x|\theta)$, onde $x \in \mathcal{X}$ e $\theta = (\theta^1, \dots, \theta^n)$ é um vector de parâmetros.

A função $p(x|\theta)$ é uma função de probabilidade ou uma função densidade de probabilidade consoante a cardinalidade de \mathcal{X} , isto é, consoante x é uma variável discreta ou contínua. Independentemente da natureza de \mathcal{X} , ir-se-á referir $p(x|\theta)$ como uma distribuição de probabilidade, sendo o sentido geralmente subentendido do contexto.

Considere-se agora a família de distribuições de probabilidade $\mathcal{S} = \{p(x|\theta)|\theta \in \Theta\}$ e suponha-se que a relação entre θ e $p(x|\theta)$ é bi-unívoca, isto é,

$$\theta^A = \theta^B \quad \Leftrightarrow \quad p(x|\theta^A) = p(x|\theta^B),$$

onde a igualdade entre distribuições é entendida como igual em quase todo o lado. Nestas condições θ pode ser considerado um sistema de coordenadas global de \mathcal{S} . Além de θ , qualquer

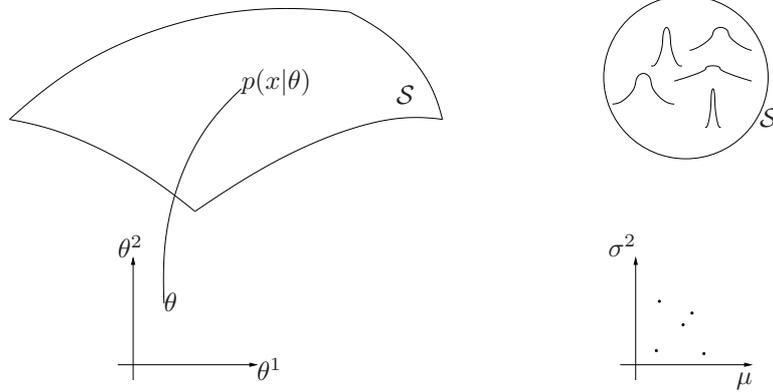


Figura 2.1: (Esq.) Variedade estatística genérica. (Dir.) Família gaussiana e correspondente sistema de coordenadas.

outro vector $\xi = F(\theta)$, com F e F^{-1} diferenciáveis é também um sistema de coordenadas, uma vez que a relação entre ξ e $p(x|F^{-1}(\xi))$ é também bi-unívoca.

A família de distribuições de probabilidade \mathcal{S} pode ser vista como uma entidade para a qual vários sistemas de coordenadas são possíveis. A esta entidade mais abstracta \mathcal{S} dá-se o nome de *variedade estatística*. Assim, uma variedade estatística é uma família paramétrica de distribuições de probabilidade, em que cada ponto é uma distribuição de probabilidade diferente. A figura 2.1 ilustra esta ideia de forma abstracta no desenho da esquerda.

Em seguida dão-se dois exemplos de variedades estatísticas que surgem de dois modelos comuns.

A família Gaussiana é uma variedade estatística de dimensão 2, em que os parâmetros (coordenadas) são a média μ e a variância σ^2 (ver figura 2.1 do lado direito). A família Gaussiana multivariável é também uma variedade estatística cujos parâmetros são o vector das médias μ e a matriz de covariância Σ . Esta variedade tem dimensão $(n^2 + 3n)/2$ onde $n = \dim(x)$.

A família de distribuições finitas sobre o alfabeto $\mathcal{X} = \{x_0, \dots, x_n\}$ é uma variedade estatística de dimensão $n = |\mathcal{X}| - 1$. Os parâmetros são as probabilidades $(\theta^1, \dots, \theta^n) = (p(x_1), \dots, p(x_n))$. Note-se que a restrição $\sum_i p(x_i) = 1$ implica que a probabilidade $p(x_0)$ seja automaticamente determinada por

$$p(x_0) = 1 - \sum_{i=1}^n p(x_i) = 1 - \sum_{i=1}^n \theta^i. \quad (2.5)$$

De um ponto de vista mais geral existem duas famílias de distribuições que têm um papel de relevo. Estas são as famílias mistura e exponencial, construídas a partir de um conjunto de funções de base $C(x)$ e $F_i(x)$ como se segue:

$$p(x|\theta) = C(x) + \sum_{i=1}^n \theta^i F_i(x) \quad (\text{família mistura}) \quad (2.6)$$

$$p(x|\theta) = \exp\left(C(x) + \sum_{i=1}^n \theta^i F_i(x) - \varphi(\theta)\right) \quad (\text{família exponencial}) \quad (2.7)$$

As distribuições Gaussiana e discreta finita referidas acima são dois casos particulares das famílias exponencial e mistura, respectivamente: Defina-se

$$F_1(x) = x, \quad F_2(x) = x^2 \quad (2.8)$$

então a família Gaussiana pode escrever-se na forma (2.7) fazendo

$$C(x) = 0, \quad \theta^1 = \frac{\mu}{\sigma^2}, \quad \theta^2 = -\frac{1}{2\sigma^2}, \quad \varphi(\theta) = -\frac{(\theta^1)^2}{4\theta^2} + \frac{1}{2} \log\left(-\frac{\pi}{\theta^2}\right). \quad (2.9)$$

Para a distribuição discreta, com $|\mathcal{X}| = n + 1$, defina-se o conjunto de distribuições $\{p_0, p_1, \dots, p_n\}$ em que $p_i(x_j) = \delta_{ij}$ são funções delta de Kronecker. Como qualquer distribuição discreta pode ser escrita na forma

$$\begin{aligned} p(x|\theta) &= \sum_{i=1}^n \theta^i p_i(x) + \left(1 - \sum_{i=1}^n \theta^i\right) p_0(x) \\ &= p_0(x) + \sum_{i=1}^n \theta^i (p_i(x) - p_0(x)), \end{aligned} \quad (2.10)$$

então, definindo

$$C(x) = p_0(x), \quad F_i(x) = p_i(x) - p_0(x), \quad (2.11)$$

obtém-se a precisamente a forma da equação (2.6) para a família mistura.

A distribuição discreta goza também da particularidade invulgar de ser simultaneamente uma família exponencial. Com efeito, definindo

$$C(x) = 0, \quad F_i(x_j) = \delta_{ij}, \quad \theta^i = \frac{p(x_i)}{1 - \sum_{j=1}^n p(x_j)}, \quad \varphi(\theta) = \log\left(1 + \sum_{i=1}^n \exp \theta^i\right) \quad (2.12)$$

obtém-se uma representação como a da equação (2.7)

Exemplo 1 (Mistura de Gaussianas). *Considerem-se duas distribuições Gaussianas $p_A(x)$ e $p_B(x)$ com parâmetros $(\mu_A, \sigma_A^2) = (-5, 1)$ e $(\mu_B, \sigma_B^2) = (5, 3)$. A família mistura de dimensão 1 que contém estas duas Gaussianas é definida por*

$$p(x|\theta) \stackrel{\text{def}}{=} (1 - \theta)p_A(x) + \theta p_B(x) = \underbrace{p_A(x)}_{C(x)} + \theta \underbrace{(p_B(x) - p_A(x))}_{F(x)} \quad (2.13)$$

A figura 2.2 mostra, em cima, as distribuições obtidas para vários valores de θ no intervalo $[0, 1]$.

Exemplo 2 (Família exponencial de Gaussianas). *Considerando as mesmas distribuições $p_A(x)$ e $p_B(x)$ do exemplo anterior, a família exponencial de dimensão 1 que contém estas duas Gaussianas é definida por*

$$p(x|\theta) \stackrel{\text{def}}{=} \frac{p_A(x)^{1-\theta} p_B(x)^\theta}{\gamma(\theta)} = \exp\left(\underbrace{\log p_A(x)}_{C(x)} + \theta \underbrace{(\log p_B(x) - \log p_A(x))}_{F(x)} - \underbrace{\log \gamma(\theta)}_{\varphi(\theta)}\right) \quad (2.14)$$

A figura 2.2 mostra, em baixo, as distribuições obtidas para vários valores de θ no intervalo $[0, 1]$. Neste caso, todas as distribuições são Gaussianas, o que não acontece no caso da família mistura do exemplo 1.

Tomando como fixas as distribuições Gaussianas $p_A(x)$ e $p_B(x)$, as famílias mistura e exponencial definidas nos dois exemplos anteriores são variedades estatísticas de dimensão 1. Os sistemas de coordenadas destas variedades consistem no parâmetro θ usado na mistura e no parâmetro θ usado na combinação exponencial.

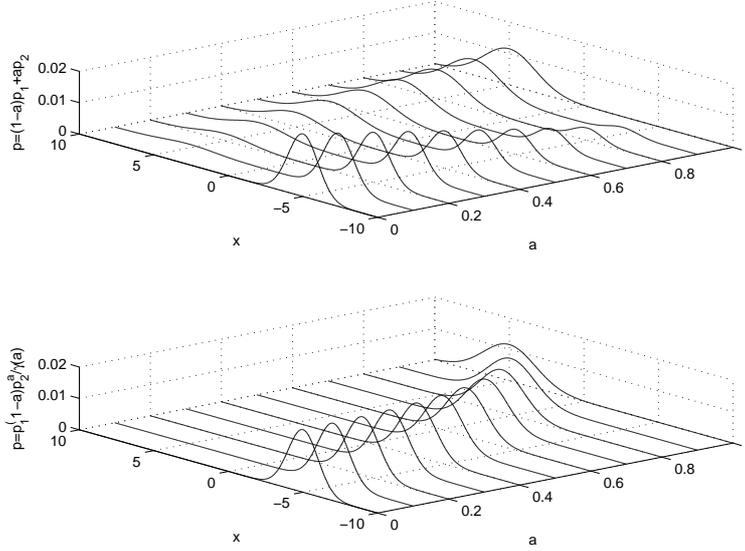


Figura 2.2: Família mistura de Gaussianas (em cima), e família exponencial de Gaussianas (em baixo).

2.3 Vectores tangentes e métrica de Fisher

Considere-se uma variedade estatística \mathcal{S} de dimensão n . Para cada ponto θ , correspondente a uma distribuição $p(x|\theta) \in \mathcal{S}$, a matriz de informação de Fisher $G(\theta)$ é uma matriz simétrica com componentes reais $g_{ij}(\theta)$ definidas por

$$g_{ij}(\theta) \stackrel{\text{def}}{=} E_{p(x|\theta)} \left[\frac{\partial \log p(x|\theta)}{\partial \theta^i} \frac{\partial \log p(x|\theta)}{\partial \theta^j} \right]. \quad (2.15)$$

A matriz G é sempre semidefinida positiva uma vez que, para todo $v \in \mathbb{R}^n$ e $v \neq 0$, tem-se

$$\begin{aligned} v^T G(\theta) v &= \sum_{i,j} v^i v^j E_{p(x|\theta)} \left[\frac{\partial \log p(x|\theta)}{\partial \theta^i} \frac{\partial \log p(x|\theta)}{\partial \theta^j} \right] \\ &= E_{p(x|\theta)} \left[\sum_{i,j} v^i v^j \frac{\partial \log p(x|\theta)}{\partial \theta^i} \frac{\partial \log p(x|\theta)}{\partial \theta^j} \right] \\ &= E_{p(x|\theta)} \left[\left(\sum_i v^i \frac{\partial \log p(x|\theta)}{\partial \theta^i} \right)^2 \right] \\ &\geq 0. \end{aligned} \quad (2.16)$$

Se os n termos

$$\frac{\partial \log p(x|\theta)}{\partial \theta^i}, \quad i = 1, \dots, n \quad (2.17)$$

são funções de x linearmente independentes, então G é definida positiva. Admite-se daqui em diante que esta última situação se verifica.

O facto das n funções (2.17) serem linearmente independentes relativamente a x , permite associar um espaço vectorial de dimensão n a cada ponto θ , em que (2.17) são os vectores

da base. A este espaço dá-se o nome de *espaço tangente* e representa-se por $T_\theta\mathcal{S}$, o espaço tangente a \mathcal{S} no ponto θ .

Note-se que se (2.17) são linearmente independentes, então

$$\frac{\partial p(x|\theta)}{\partial \theta^i}, \quad i = 1, \dots, n \quad (2.18)$$

são também funções linearmente independentes em x . Chama-se à representação (2.17) a *representação exponencial* de um vector tangente, e à representação (2.18) a *representação mistura* desse mesmo vector tangente. A passagem de uma representação para a outra consiste unicamente em multiplicar ou dividir, conforme apropriado, os vectores da base por $p(x|\theta)$, uma vez que $\frac{\partial \log p}{\partial \theta^i} = \frac{1}{p} \frac{\partial p}{\partial \theta^i}$.

Esta representação de vector tangente contrasta com a representação tradicional, em geometria diferencial, de um vector tangente como um operador de derivada parcial. Dada uma variedade diferencial \mathcal{M} com um sistema de coordenadas θ e uma função suave f definida em \mathcal{M} , a derivada direccionada de f num ponto q tem a forma

$$\sum_{i=1}^n a^i(q) \frac{\partial}{\partial \theta^i} \Big|_q f,$$

em que a direcção da derivação é especificada pelas componentes a^i . Os operadores $\frac{\partial}{\partial \theta^i}$ são usualmente interpretados como os vectores da base (do espaço tangente).

Em variedades estatísticas, é possível dar uma interpretação análoga como se segue. Suponha-se definido um funcional F como

$$F(p) \stackrel{\text{def}}{=} \int_{\mathcal{X}} f(p(x)) dx,$$

onde $f : \mathbb{R} \rightarrow \mathbb{R}$ é uma função diferenciável (exemplos: $f(p) = p^2$, $f(p) = -p \log p$) e $p(x)$ uma função densidade de probabilidade arbitrária com $p(x) > 0$ para $x \in \mathcal{X}$. Restringindo $p(x)$ a uma variedade \mathcal{S} com coordenadas θ , então, o funcional F passa a ser uma função de θ . Isto é, $F : \Theta \rightarrow \mathbb{R}$ é uma função na variedade e as suas derivadas nas direcções das coordenadas θ^i , $i = 1, \dots, n$, são

$$\frac{\partial F}{\partial \theta^i} = \int f'(p(x|\theta)) \frac{\partial p}{\partial \theta^i} dx = \int p(x|\theta) f'(p(x|\theta)) \frac{\partial \log p}{\partial \theta^i} dx.$$

Nas três expressões anteriores observam-se as três representações dos vectores da base do espaço tangente, nomeadamente a representação tradicional como operador derivada parcial $\frac{\partial}{\partial \theta^i}$, a representação mistura $\frac{\partial p}{\partial \theta^i}$ e a representação exponencial $\frac{\partial \log p}{\partial \theta^i}$.

Da mesma forma, a derivada direccionada de F numa direcção arbitrária v com componentes v^i é

$$\underbrace{\sum_{i=1}^n v^i \frac{\partial}{\partial \theta^i}}_{\text{tradicional}} F = \int f'(p) \underbrace{\left(\sum_{i=1}^n v^i \frac{\partial p}{\partial \theta^i} \right)}_{\text{mistura}} dx = \int p f'(p) \underbrace{\left(\sum_{i=1}^n v^i \frac{\partial \log p}{\partial \theta^i} \right)}_{\text{exponencial}} dx. \quad (2.19)$$

Assim, as três representações estão naturalmente identificadas pela equação anterior. Note-se que a representação exponencial (termo da direita) da equação anterior se pode escrever como o valor esperado

$$E_p \left[f'(p) \sum_{i=1}^n v^i \frac{\partial \log p}{\partial \theta^i} \right].$$

Esta expressão vai surgir mais tarde nesta tese em problemas de otimização, onde o funcional F será substituído pela divergência de Kullback-Leibler $D(\cdot\|\cdot)$.

Tendo definido o conceito de vector tangente e espaço tangente, torna-se possível a introdução de um produto interno no espaço tangente a cada ponto da variedade. Para o efeito é usada a matriz de informação de Fisher (2.15). Assim, define-se o produto interno em $T_\theta\mathcal{S}$ pela fórmula

$$\left\langle \frac{\partial}{\partial\theta^i}, \frac{\partial}{\partial\theta^j} \right\rangle(\theta) \stackrel{\text{def}}{=} g_{ij}(\theta). \quad (2.20)$$

A métrica que se obtém é uma métrica Riemanniana [11] a que se dá o nome de *métrica de Fisher*.

Para dois vectores arbitrários a e b em $T_\theta\mathcal{S}$ com componentes a^i e b^j , respectivamente, o seu produto interno calcula-se da seguinte maneira:

$$\left\langle \sum_{i=1}^n a^i \frac{\partial}{\partial\theta^i}, \sum_{j=1}^n b^j \frac{\partial}{\partial\theta^j} \right\rangle(\theta) = \sum_{i,j} a^i b^j g_{ij}(\theta),$$

que não é mais que a forma quadrática $a^T G b$, onde a e b são representados na forma de vectores coluna, e G é a matriz de informação de Fisher.

A escolha da matriz de informação de Fisher como a métrica natural nestes espaços não é arbitrária. Os espaços de distribuições de probabilidade são, em certa medida, espaços peculiares com características próprias. Nestes espaços, a métrica de Fisher é vista como uma métrica natural. Existem vários argumentos que sustentam esta visão:

1. A métrica de Fisher é invariante a mudanças de coordenadas, no sentido em que, as componentes se calculam de acordo com a equação (2.15), independentemente do sistema de coordenadas θ usado. Por outras palavras, após a transformação de coordenadas $\theta \mapsto \xi$, a métrica é

$$\tilde{g}_{ij}(\xi) = E_\xi \left[\frac{\partial \log p(x|\xi)}{\partial \xi^i} \frac{\partial \log p(x|\xi)}{\partial \xi^j} \right].$$

Naturalmente, as componentes de $g_{ij}(\theta)$ e $\tilde{g}_{ij}(\xi)$ diferem, mas os produtos internos entre pares de vectores, representados nos dois sistemas de coordenadas, coincidem (ver apêndice A.1).

2. A métrica de Fisher é invariante a mapas que transformam as variáveis aleatórias, desde que satisfaçam um critério de suficiência. Isto é, suponha-se que ao modelo $p(x|\theta)$, $x \in \mathcal{X}$, corresponde a métrica g_{ij} ; então, se o mapa $t : \mathcal{X} \rightarrow \mathcal{Y}$ é uma estatística suficiente para este modelo, a métrica de Fisher G_t do modelo induzido $q(y|\theta)$ coincide com a original, $G(\theta) = G_t(\theta)$. Intuitivamente, não havendo perda de informação na aplicação da estatística t , a métrica mantém-se (ver [6] pags. 29–30).

A invariância relativamente à parametrização garante que a geometria obtida é intrínseca, e não devida a uma parametrização particular. A invariância não ocorreria caso fosse usada uma métrica Euclidiana, por exemplo.

A matriz de informação pode ainda ser vista intuitivamente como indicador do nível de independência entre vectores da base do espaço tangente. Se a matriz G é singular, então existem um ou mais vectores linearmente dependentes, conforme se observa em (2.16). Nessa situação, o espaço tangente, e consequentemente a própria variedade, têm na realidade uma dimensão inferior à considerada.

2.4 As conexões- α de Amari

Uma conexão é um operador diferencial que permite calcular derivadas direccionais de campos vectoriais numa variedade. A existência de um operador deste tipo permite definir o que se entende por geodésica, ou seja, definir o que é uma “linha recta” numa variedade.

Em \mathbb{R}^n , uma curva é uma linha recta se o vector tangente à curva se mantém constante ao longo da curva. Como os vectores tangentes formam um campo vectorial ao longo da curva, se a derivada direccional deste campo vectorial se anular na direcção da curva, então esta é uma geodésica. Numa variedade mantém-se o mesmo princípio, mas torna-se necessário saber calcular a “derivada direccional de um campo vectorial”, *i.e.*, a conexão.

Genericamente, uma conexão representa-se pelo operador $\nabla_v w$ querendo significar derivada de w na direcção v . O resultado desta operação é um novo campo vectorial que se pode escrever, em cada ponto, como uma combinação linear dos vectores tangentes $\partial/\partial\theta^k$. A fórmula

$$\nabla_v w = \sum_k \underbrace{\left(\sum_i v^i \frac{\partial w^k}{\partial \theta^i} + \sum_{i,j} v^i w^j \Gamma_{ij}^k \right)}_{\text{componentes do vector resultante}} \frac{\partial}{\partial \theta^k} \quad (2.21)$$

mostra como se obtêm as componentes do novo campo vectorial em função das componentes de w e v , e dos *símbolos de Christoffel* Γ_{ij}^k .

Tomando como ilustração o caso de um campo vectorial w em \mathbb{R}^n , a sua derivada na direcção v determina-se calculando a matriz Jacobiana $\frac{\partial w}{\partial \theta}$ e multiplicando-a pela direcção v . O vector resultante tem as componentes

$$\sum_{i=1}^n v^i \frac{\partial w^k}{\partial \theta^i}, \quad k = 1, \dots, n.$$

Comparando esta expressão com (2.21) observa-se que a derivação em \mathbb{R}^n é um caso particular de (2.21), em que os símbolos de Christoffel são identicamente nulos, $\Gamma_{ij}^k(\theta) = 0$.

Em geral, pode definir-se uma infinidade de conexões numa variedade, uma vez que existem muitas possibilidades para a definição dos símbolos de Christoffel. No entanto, estes não são definidos arbitrariamente devendo ser tais que a conexão verifique certas propriedades desejadas (ver *e.g.* [11]).

Para as variedades estatísticas, Amari [4] define uma família a um parâmetro de conexões, conhecidas como *conexões- α* e representadas por $\nabla^{(\alpha)}$, onde para cada $\alpha \in [-1, 1]$ se obtém uma conexão diferente. Os símbolos de Christoffel que caracterizam as conexões- α variam com α e calculam-se pela fórmula

$$\Gamma_{ij,k}^{(\alpha)} = E_\theta \left[\left(\frac{\partial^2 \log p}{\partial \theta_i \partial \theta_j} + \frac{1 - \alpha}{2} \frac{\partial \log p}{\partial \theta_i} \frac{\partial \log p}{\partial \theta_j} \right) \frac{\partial \log p}{\partial \theta_k} \right]. \quad (2.22)$$

onde

$$\Gamma_{ij,k} \stackrel{\text{def}}{=} \left\langle \nabla_{\partial/\partial\theta^i} \frac{\partial}{\partial\theta^j}, \frac{\partial}{\partial\theta^k} \right\rangle = \sum_l \Gamma_{ij}^l g_{lk}. \quad (2.23)$$

Das conexões- α de Amari, três delas merecem destaque. Estas são as conexões *mistura*, *métrica* e *exponencial* correspondentes a $\alpha = -1$, $\alpha = 0$ e $\alpha = 1$, respectivamente. As secções seguintes apresentam algumas das suas propriedades.

2.4.1 A conexão métrica

A conexão métrica, também conhecida como conexão Riemanniana ou de Levi-Civita, é a que se obtém directamente do tensor métrico g_{ij} pela fórmula

$$\Gamma_{ij,k} = \frac{1}{2} \left(\frac{\partial g_{ik}}{\partial \theta^j} + \frac{\partial g_{jk}}{\partial \theta^i} - \frac{\partial g_{ij}}{\partial \theta^k} \right). \quad (2.24)$$

Usando o tensor g_{ij} dado pela métrica de Fisher (2.15), obtém-se a conexão-0 de Amari, correspondente a fazer $\alpha = 0$ em (2.22) (ver cálculos em A.2).

Embora na literatura sobre geometria Riemanniana se trate exclusivamente esta conexão, em variedades estatísticas esta encontra relativamente pouca aplicação.

2.4.2 A conexão exponencial

A definição de conexão exponencial pode ser feita recorrendo à noção de derivada covariante. Assim, considere-se um vector tangente com a representação exponencial $\frac{\partial \log p}{\partial \theta^i}$. Derivando este vector numa direcção j , obtém-se uma expressão

$$\frac{\partial}{\partial \theta^j} \frac{\partial \log p}{\partial \theta^i} \quad (2.25)$$

que não é exprimível apenas como combinação linear dos vectores tangentes $\frac{\partial \log p}{\partial \theta^i}$, $i = 1, \dots, n$, e consequentemente não pertence ao espaço tangente. No entanto, a expressão resultante pode ser escrita como a soma de uma componente tangente com uma ortogonal ao espaço tangente relativamente à métrica de Fisher. Aqui coloca-se uma questão técnica: A métrica de Fisher está definida apenas para o espaço tangente; e a componente “ortogonal” não é sequer um vector tangente a nenhuma outra variedade, uma vez que o seu valor esperado não se anula. Pode, no entanto, definir-se $\tilde{\pi}^\perp \stackrel{\text{def}}{=} \pi^\perp - E[\pi^\perp]$, cujo valor esperado se anula. Entende-se em seguida que o significado do termo *ortogonalidade* é estendido para significar que $E_p[\pi^\perp \frac{\partial \log p}{\partial \theta^k}] = E_p[\tilde{\pi}^\perp \frac{\partial \log p}{\partial \theta^k}] = 0$, para qualquer expressão π^\perp .

Definindo $\pi(\cdot)$ como sendo a projecção no espaço tangente e $\pi^\perp(\cdot)$ a componente ortogonal restante, *i.e.*

$$\frac{\partial^2 \log p}{\partial \theta^j \partial \theta^i} = \pi \left(\frac{\partial^2 \log p}{\partial \theta^j \partial \theta^i} \right) + \pi^\perp \left(\frac{\partial^2 \log p}{\partial \theta^j \partial \theta^i} \right), \quad (2.26)$$

podem definir-se os símbolos de Christoffel $\Gamma_{ij,k}$ da conexão exponencial de maneira a que esta coincida com a derivação covariante que se obtém por derivação seguida de projecção no espaço tangente. Assim,

$$\begin{aligned} \Gamma_{ij,k} &= \left\langle \nabla_j \frac{\partial}{\partial \theta^i}, \frac{\partial}{\partial \theta^k} \right\rangle \\ &= E \left[\pi \frac{\partial \log p}{\partial \theta^k} \right] \\ &= E \left[(\pi + \pi^\perp) \frac{\partial \log p}{\partial \theta^k} \right] \\ &= E \left[\frac{\partial^2 \log p}{\partial \theta^j \partial \theta^i} \frac{\partial \log p}{\partial \theta^k} \right]. \end{aligned} \quad (2.27)$$

O resultado (2.27) coincide com a expressão (2.22) para $\alpha = 1$. À conexão cujos símbolos de Christoffel são dados por (2.27) dá-se o nome de conexão exponencial ou conexão-(1) de Amari, realçando-se assim o facto de pertencer à família de conexões- α . Esta conexão

tem propriedades importantes quando aplicada a famílias exponenciais de distribuições de probabilidade, como se mostra em seguida.

Da definição de família exponencial

$$p(x|\theta) = \exp \left(C(x) + \sum_{i=1}^n \theta^i F_i(x) - \varphi(\theta) \right) \quad (2.28)$$

$$\varphi(\theta) = \log \int \exp \left(C(x) + \sum_{i=1}^n \theta^i F_i(x) \right) dx \quad (2.29)$$

obtém-se que

$$\frac{\partial \log p(x|\theta)}{\partial \theta^i} = F_i(x) - \frac{\partial \varphi(\theta)}{\partial \theta^i} \quad (2.30)$$

e

$$\frac{\partial^2 \log p(x|\theta)}{\partial \theta^i \partial \theta^j} = - \frac{\partial^2 \varphi(\theta)}{\partial \theta^i \partial \theta^j}. \quad (2.31)$$

Nesta família, a conexão exponencial ($\alpha = 1$) tem símbolos de Christoffel identicamente nulos em toda a variedade, pois

$$\begin{aligned} \Gamma_{ij,k}^{(1)} &= E_\theta \left[\frac{\partial^2 \log p(x|\theta)}{\partial \theta^i \partial \theta^j} \frac{\partial \log p(x|\theta)}{\partial \theta^k} \right] \\ &= - \frac{\partial^2 \varphi(\theta)}{\partial \theta^i \partial \theta^j} \underbrace{E_\theta \left[\frac{\partial \log p(x|\theta)}{\partial \theta^k} \right]}_{=0} \\ &= 0. \end{aligned} \quad (2.32)$$

Assim a família exponencial é um variedade plana¹ relativamente à conexão $\nabla^{(1)}$ e a curva exponencial que une duas distribuições é uma geodésica, como se mostra em seguida.

Seja \mathcal{S} uma variedade estatística que é uma família exponencial como em (2.28), e considerem-se dois pontos em \mathcal{S} com coordenadas θ_a e θ_b conhecidas. Chama-se *curva exponencial* à família a um parâmetro de distribuições que une $p_a \stackrel{\text{def}}{=} p(x|\theta_a)$ e $p_b \stackrel{\text{def}}{=} p(x|\theta_b)$ definida por

$$r(x|t) \stackrel{\text{def}}{=} \frac{p_a^{1-t} p_b^t}{\gamma(t)}, \quad (2.33)$$

em que t é o parâmetro e $\gamma(t)$ uma função de normalização. Tendo em conta que as distribuições envolvidas são de uma família exponencial, o desenvolvimento da expressão anterior resulta em

$$r(x|t) = \exp \left(C(x) + \sum_{i=1}^n ((1-t)\theta_a^i + t\theta_b^i) F_i(x) - (1-t)\varphi(\theta_a) - t\varphi(\theta_b) - \log \gamma(t) \right). \quad (2.34)$$

Defina-se $\theta(t) \stackrel{\text{def}}{=} (1-t)\theta_a + t\theta_b$. Então

$$r(x|t) = \exp \left(C(x) + \sum_{i=1}^n \theta^i(t) F_i(x) - \varphi(\theta(t)) \right), \quad (2.35)$$

donde se observa que todos os pontos desta curva pertencem a \mathcal{S} e se justifica o nome de curva exponencial. Note-se no entanto que, no espaço de parâmetros, a curva $\theta(t)$ é uma recta em \mathbb{R}^n .

¹Numa variedade plana a conexão calcula-se da mesma forma que a derivada direccional de um campo vectorial em \mathbb{R}^n .

Um vector tangente à curva exponencial é dado por

$$\begin{aligned}
\frac{\partial \log r}{\partial t} &= \sum_{i=1}^n \frac{\partial \theta^i(t)}{\partial t} F_i(x) - \sum_{i=1}^n \frac{\partial \varphi}{\partial \theta^i} \frac{\partial \theta^i(t)}{\partial t} \\
&= \sum_{i=1}^n \frac{\partial \theta^i(t)}{\partial t} \left(F_i(x) - \frac{\partial \varphi}{\partial \theta^i} \right) \\
&= \sum_{i=1}^n \frac{\partial \theta^i(t)}{\partial t} \frac{\partial \log p}{\partial \theta^i} \\
&= \sum_{i=1}^n \left(\theta_b^i - \theta_a^i \right) \frac{\partial \log p}{\partial \theta^i}
\end{aligned} \tag{2.36}$$

Assim, as componentes v^i do vector tangente à curva são as constantes

$$v^i = \theta_b^i - \theta_a^i, \quad i = 1, \dots, n, \tag{2.37}$$

e não dependem do ponto t da curva onde o vector tangente é calculado.

Finalmente, para verificar que a curva exponencial numa família exponencial \mathcal{S} é uma geodésica relativamente à conexão exponencial $\nabla^{(1)}$, mostra-se que $\nabla_v^{(1)}v$ é idênticamente nula ao longo da curva. Usando a fórmula (2.21) e tendo em consideração que os símbolos de Christoffel Γ_{ij}^k se anulam conforme (2.32), obtém-se

$$\begin{aligned}
\nabla_v^{(1)}v &= \sum_k \left(\sum_i v^i \frac{\partial v^k}{\partial \theta^i} \right) \frac{\partial}{\partial \theta^k} \\
&= \sum_k \left(\sum_i (\theta_b^i - \theta_a^i) \underbrace{\frac{\partial}{\partial \theta^i} (\theta_b^k - \theta_a^k)}_{=0} \right) \frac{\partial}{\partial \theta^k} \\
&= 0.
\end{aligned} \tag{2.38}$$

2.4.3 A conexão mistura

À semelhança da conexão exponencial, a conexão mistura pode ser definida derivando um vector tangente e projectando o resultado no espaço tangente usando a métrica de Fisher, com a diferença de que se usa a representação mistura $\frac{\partial p}{\partial \theta^i}$ para o vector tangente. Derivando na direcção j , obtém-se que

$$\begin{aligned}
\frac{\partial}{\partial \theta^j} \frac{\partial p}{\partial \theta^i} &= \frac{\partial}{\partial \theta^j} \left(p \frac{\partial \log p}{\partial \theta^i} \right) \\
&= \frac{\partial p}{\partial \theta^j} \frac{\partial \log p}{\partial \theta^i} + p \frac{\partial^2 \log p}{\partial \theta^j \partial \theta^i} \\
&= p \left(\frac{\partial \log p}{\partial \theta^j} \frac{\partial \log p}{\partial \theta^i} + \frac{\partial^2 \log p}{\partial \theta^j \partial \theta^i} \right).
\end{aligned} \tag{2.39}$$

Os símbolos de Christoffel são então

$$\Gamma_{ij,k} = E \left[\left(\frac{\partial \log p}{\partial \theta^j} \frac{\partial \log p}{\partial \theta^i} + \frac{\partial^2 \log p}{\partial \theta^j \partial \theta^i} \right) \frac{\partial \log p}{\partial \theta^k} \right], \tag{2.40}$$

correspondentes a $\alpha = -1$ em (2.22). A conexão-(-1) tem propriedades importantes quando aplicada a famílias mistura, como se mostra em seguida.

A família mistura de $n + 1$ distribuições $\{p_0, p_1, \dots, p_n\}$ parametrizada por n parâmetros θ^i , e definida por

$$p \stackrel{\text{def}}{=} \sum_{i=1}^n \theta^i p_i + \left(1 - \sum_{i=1}^n \theta^i\right) p_0, \quad (2.41)$$

com $\theta^i > 0$ e $\sum_{i=1}^n \theta^i < 1$, tem como primeira e segunda derivadas, em representação exponencial, as expressões

$$\frac{\partial \log p}{\partial \theta^i} = \frac{p_i - p_0}{p} \quad (2.42)$$

e

$$\frac{\partial^2 \log p}{\partial \theta^i \partial \theta^j} = -\frac{(p_i - p_0)(p_j - p_0)}{p^2}, \quad (2.43)$$

respectivamente.

A conexão mistura tem símbolos de Christoffel identicamente nulos em toda a variedade, pois

$$\begin{aligned} \Gamma_{ij,k}^{(-1)} &= E \left[\left(\frac{\partial \log p}{\partial \theta^j} \frac{\partial \log p}{\partial \theta^i} + \frac{\partial^2 \log p}{\partial \theta^j \partial \theta^i} \right) \frac{\partial \log p}{\partial \theta^k} \right] \\ &= E \left[\underbrace{\left(\frac{p_j - p_0}{p} \frac{p_i - p_0}{p} - \frac{(p_j - p_0)(p_i - p_0)}{p^2} \right)}_{=0} \frac{p_k - p_0}{p} \right] \\ &= 0. \end{aligned} \quad (2.44)$$

Assim, a família mistura é uma variedade plana relativamente à conexão $\nabla^{(-1)}$ e a mistura de duas distribuições da variedade é uma geodésica.

2.5 Divergências entre distribuições de probabilidade

Chama-se *divergência* entre duas distribuições p e q a uma função suave $D(p||q)$ que verifica as seguintes propriedades:

$$D(p||q) \geq 0, \quad \text{e} \quad D(p||q) = 0 \quad \text{sse} \quad p = q. \quad (2.45)$$

Assim, $D(p||q)$ é uma medida de separação entre duas distribuições semelhante a uma distância, mas que não verifica necessariamente todos os axiomas de uma distância. Em particular, a simetria e a desigualdade triangular podem não ser satisfeitas.

Embora a divergência não seja em geral uma distância, esta está intimamente ligada à estrutura métrica da variedade estatística. Em particular, é possível induzir uma métrica Riemanniana $g^{(D)}$ e uma conexão $\nabla^{(D)}$ a partir unicamente de uma divergência definida globalmente na variedade, como se segue.

Considere-se uma distribuição p_θ com parâmetros θ e uma distribuição vizinha $p_{\theta+\Delta\theta}$ onde $\Delta\theta = \{\Delta\theta_1, \dots, \Delta\theta_n\}$ representa um pequeno desvio dos parâmetros relativamente ao ponto θ . Define-se a métrica Riemanniana e a conexão induzidas pelas fórmulas

$$g_{ij}^{(D)}(\theta) \stackrel{\text{def}}{=} \frac{\partial^2}{\partial \Delta\theta_i \partial \Delta\theta_j} D(p_{\theta+\Delta\theta} || p_\theta) = \frac{\partial^2}{\partial \Delta\theta_i \partial \Delta\theta_j} D(p_\theta || p_{\theta+\Delta\theta}) \quad (2.46)$$

e

$$\Gamma_{ij,k}^{(D)}(\theta) \stackrel{\text{def}}{=} \frac{\partial}{\partial \Delta\theta'_k} \frac{\partial^2}{\partial \Delta\theta_i \partial \Delta\theta_j} D(p_{\theta+\Delta\theta} || p_{\theta+\Delta\theta'}). \quad (2.47)$$

tomadas em $\Delta\theta = 0$ e $\Delta\theta' = 0$.

Uma divergência que surge habitualmente na literatura é a divergência de Kullback-Leibler definida por

$$D_{\text{KL}}(p\|q) \stackrel{\text{def}}{=} \int p(x) \log \frac{p(x)}{q(x)} dx. \quad (2.48)$$

Esta surge naturalmente em vários problemas de estatística e de teoria da informação. A divergência de Kullback-Leibler está também intimamente ligada à estrutura métrica da variedade estatística. Em particular, é possível provar (ver apêndice A.3) que a métrica induzida pela divergência de Kullback-Leibler é a métrica de Fisher, *i.e.*,

$$\frac{\partial}{\partial \Delta\theta_i} \frac{\partial}{\partial \Delta\theta_j} D(p_{\theta+\Delta\theta}\|p_\theta) = \frac{\partial}{\partial \Delta\theta_i} \frac{\partial}{\partial \Delta\theta_j} D(p_\theta\|p_{\theta+\Delta\theta}) = g_{ij}(\theta). \quad (2.49)$$

A matriz de informação de Fisher é assim a matriz Hessiana da divergência $D(p_{\theta+\Delta\theta}\|p_\theta)$ num ponto θ vista como função do desvio $\Delta\theta$, e pode também ser interpretada como aproximação local de segunda ordem da divergência de Kullback-Leibler, $D(p_{\theta+\Delta\theta}\|p_\theta) \approx \Delta\theta^T G \Delta\theta$.

Adicionalmente, verifica-se que as conexões induzidas pela divergência de Kullback-Leibler $D(p\|q)$ e pela sua dual $D^*(p\|q) \stackrel{\text{def}}{=} D(q\|p)$ são, respectivamente, as conexões mistura $\Gamma_{ij,k}^{(-1)}(\theta)$ e exponencial $\Gamma_{ij,k}^{(1)}(\theta)$ (ver apêndice A.4).

2.5.1 Teorema de Pitágoras

Em espaços Euclidianos, com produto interno $\langle \cdot, \cdot \rangle$, e com a respectiva norma $\|x\| \stackrel{\text{def}}{=} \sqrt{\langle x, x \rangle}$ e distância $d(x, y) \stackrel{\text{def}}{=} \|x - y\|$ induzidas, quaisquer três pontos p, q e r distintos verificam uma equação que relaciona as distâncias entre eles da seguinte maneira:

$$d(p, q)^2 = d(p, r)^2 + d(r, q)^2 - 2\langle \vec{r\bar{p}}, \vec{r\bar{q}} \rangle \quad (2.50)$$

onde $\vec{r\bar{p}} = p - r$, e $\vec{r\bar{q}} = q - r$. Em particular, quando os vectores $\vec{r\bar{p}}$ e $\vec{r\bar{q}}$ são ortogonais, obtém-se o conhecido teorema de Pitágoras

$$d(p, q)^2 = d(p, r)^2 + d(r, q)^2.$$

Em espaços de distribuições de probabilidades é possível construir um resultado análogo recorrendo à divergência de Kullback-Leibler entre distribuições e ao produto interno definido pela métrica de Fisher.

Considerem-se três distribuições de probabilidade p, q e r , e definam-se as curvas mistura $m(a)$ e exponencial $e(b)$ seguintes:

$$m(a) \stackrel{\text{def}}{=} ap + (1 - a)r \quad (2.51)$$

e

$$e(b) \stackrel{\text{def}}{=} \frac{r^{1-b} q^b}{\gamma(b)}, \quad (2.52)$$

onde $a \in [0, 1]$ e $b \in [0, 1]$ parametrizam as curvas mistura e exponencial, respectivamente, e $\gamma(b)$ é um factor de normalização. Então, verifica-se a seguinte igualdade (ver apêndice A.5 e figura 2.3):

$$D(p\|q) = D(p\|r) + D(r\|q) - \left\langle \frac{\partial \log m(a)}{\partial a} \Big|_{a=0}, \frac{\partial \log e(b)}{\partial b} \Big|_{b=0} \right\rangle_r. \quad (2.53)$$

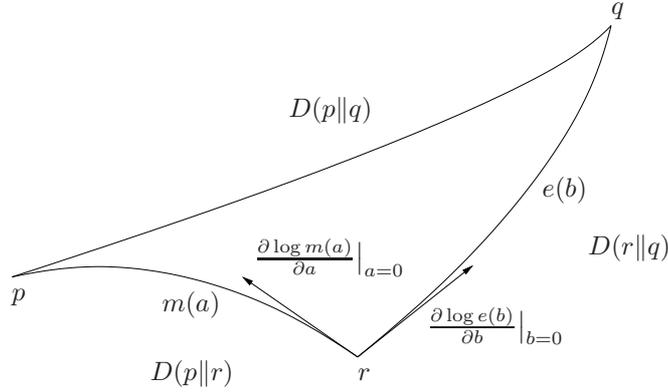


Figura 2.3: Teorema Pitagórico para distribuições de probabilidade.

Em particular, se no ponto r as curvas $m(a)$ e $e(b)$ são ortogonais, então obtém-se a fórmula do teorema de Pitágoras generalizado para distribuições de probabilidade:

$$D(p||q) = D(p||r) + D(r||q). \quad (2.54)$$

Este resultado tem aplicação directa em problemas de optimização envolvendo distribuições de probabilidade e a divergência de Kullback-Leibler. Em particular, observando a equação (2.53) obtém-se imediatamente a solução dos dois problemas que se enunciam em seguida:

1. Dada uma família mistura \mathcal{M} , e uma distribuição $q \notin \mathcal{M}$, qual a distribuição $r \in \mathcal{M}$ que minimiza $D(r||q)$?
2. Dada uma família exponencial \mathcal{E} , e uma distribuição $p \notin \mathcal{E}$, qual a distribuição $r \in \mathcal{E}$ que minimiza $D(p||r)$?

A ideia para a resposta a estas questões consiste em considerar duas situações: Na primeira, a curva mistura $m(a)$ pertence a uma família mistura \mathcal{M} e responde-se ao primeiro problema; Na segunda situação, a curva exponencial $e(b)$ pertence a uma família exponencial \mathcal{E} e responde-se ao segundo problema enunciado. Em ambas, considera-se que a condição de ortogonalidade se verifica no ponto r . As secções 2.5.2 e 2.5.3 apresentam as soluções em maior detalhe.

2.5.2 Minimização da divergência de Kullback-Leibler em famílias mistura

Considere-se uma família mistura \mathcal{M} de distribuições de probabilidade e uma distribuição $q \notin \mathcal{M}$. Pretende-se encontrar a distribuição $r \in \mathcal{M}$ que minimiza $D(r||q)$ (ver figura 2.4(a)).

A solução deste problema consiste em efectuar uma projecção ortogonal de q sobre a família mistura \mathcal{M} ao longo de uma curva exponencial (ver apêndice A.6). A demonstração usando a relação Pitagórica é a seguinte:

Suponhamos que a distribuição r é o resultado da projecção ortogonal de q sobre \mathcal{M} ao longo de uma curva exponencial. Então, o produto interno anula-se,

$$\left\langle \frac{\partial \log r}{\partial \theta^i}, \frac{\partial \log e}{\partial b} \Big|_{b=0} \right\rangle = 0, \quad i = 1, \dots, n \quad (2.55)$$

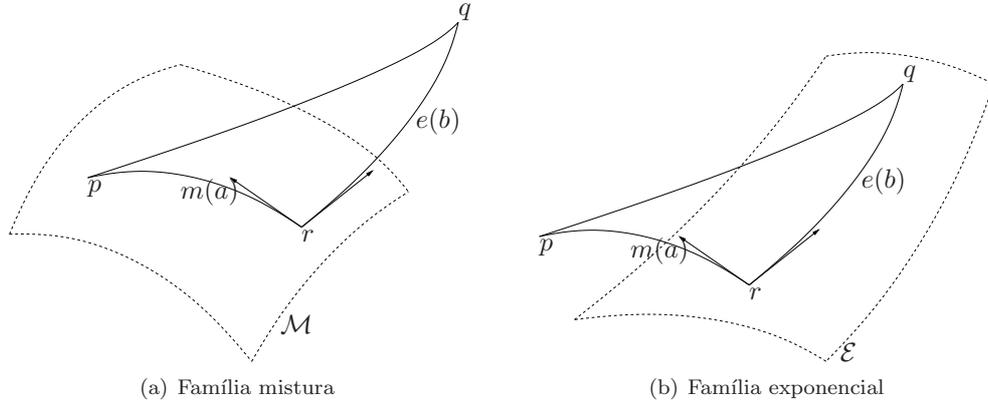


Figura 2.4: Minimização da divergência de Kullback-Leibler.

e verifica-se o teorema de Pitágoras (2.54), para todo $p \in \mathcal{M}$ com $p \neq r$. Como $D(p||r) > 0$, então $D(p||q) > D(r||q)$, ficando assim provado que r é a distribuição em \mathcal{M} mais “perto” de q .

Qualquer outro ponto sobre a curva exponencial $e(b)$ tem como projecção ortogonal o mesmo ponto $r \in \mathcal{M}$. Isto é, $r \in \mathcal{M}$ é o ponto que minimiza a divergência de \mathcal{M} a qualquer ponto $e(b)$ da curva exponencial, como se mostra em seguida.

Proposição 3. *Considere-se a curva exponencial $e(b) \stackrel{\text{def}}{=} \frac{1}{\gamma(b)} r^b q^{1-b}$. Para todo $b \in [0, 1[$ é verdade que*

$$\arg \min_{r \in \mathcal{M}} D(r||q) = \arg \min_{r \in \mathcal{M}} D(r||e(b)). \quad (2.56)$$

Demonstração. Dado um ponto q , o ponto $r \in \mathcal{M}$ que minimiza $D(r||q)$ verifica

$$\frac{\partial D(r||q)}{\partial \theta} = E_r \left[\log \frac{r}{q} \frac{\partial \log r}{\partial \theta} \right] = 0. \quad (2.57)$$

Para um ponto $e(b)$ temos que

$$\begin{aligned} \frac{\partial D(r||e(b))}{\partial \theta} &= E_r \left[\log \frac{r}{e(b)} \frac{\partial \log r}{\partial \theta} \right] \\ &= E_r \left[\log \frac{r \gamma(b)}{r^b q^{1-b}} \frac{\partial \log r}{\partial \theta} \right] \\ &= (1-b) E_r \left[\log \frac{r}{q} \frac{\partial \log r}{\partial \theta} \right] + E_r \left[\log \gamma(b) \frac{\partial \log r}{\partial \theta} \right] \\ &= (1-b) \underbrace{E_r \left[\log \frac{r}{q} \frac{\partial \log r}{\partial \theta} \right]}_{=0} \\ &= 0. \end{aligned} \quad (2.58)$$

□

Embora de um ponto de vista conceptual a solução geométrica seja elegante, a implementação prática da projecção ortogonal requer algum cuidado. Para uma família mistura

$$r(x|\theta) = C(x) + \sum_{i=1}^n \theta^i F_i(x),$$

a condição de ortogonalidade (2.55) resulta em

$$\int F_i(x) \log \frac{q(x)}{r(x|\theta)} dx = 0, \quad i = 1, \dots, n \quad (2.59)$$

não havendo em geral uma solução explícita para os parâmetros θ . Assim, torna-se necessário desenvolver métodos iterativos para obter uma aproximação do ponto óptimo. Como se trata de famílias mistura, os parâmetros θ estão naturalmente sujeitos a restrições. Estas terão de ser consideradas no procedimento de optimização.

2.5.3 Minimização da divergência de Kullback-Leibler em famílias exponenciais

De forma análoga ao problema anterior, considera-se uma família exponencial \mathcal{E} e uma distribuição $p \notin \mathcal{E}$. Pretende-se encontrar a distribuição $r \in \mathcal{E}$ que minimiza $D(p||r)$ (ver figura 2.4(b)).

A solução deste problema consiste em efectuar uma projecção ortogonal de p sobre a família exponencial \mathcal{E} ao longo de uma curva mistura. Segue-se a demonstração.

Suponha-se que a distribuição r é o resultado da projecção ortogonal de p sobre \mathcal{E} . Está-se novamente nas condições de aplicação do teorema de Pitágoras (2.54), para todo $q \in \mathcal{E}$ com $q \neq r$. Como $D(r||q) > 0$, então $D(p||q) > D(p||r)$, ficando assim provado que r é a distribuição em \mathcal{E} mais “perto” de p .

2.6 Conclusão

Neste capítulo foi introduzida uma moldura de trabalho geométrica para o tratamento de distribuições de probabilidade. Uma família paramétrica de distribuições de probabilidade foi interpretada como uma variedade diferencial, com uma métrica natural definida pela matriz de informação de Fisher. Foram introduzidas as conexões de Amari, e deduzidas as geodésicas para os casos particulares das famílias exponencial e mistura. Foi introduzida a divergência de Kullback-Leibler e apresentadas algumas das suas propriedades, em particular o teorema de Pitágoras generalizado. Finalmente foram considerados problemas de optimização da divergência de Kullback-Leibler numa perspectiva geométrica.

Capítulo 3

Optimização de distribuições de probabilidade

3.1 Introdução

Em problemas que envolvem optimização de distribuições de probabilidade, dos quais as secções 2.5.2 e 2.5.3 são um exemplo, é frequente recorrer a métodos iterativos para aproximar a solução ideal.

A literatura [17, 9, 29] é rica em algoritmos de optimização com e sem restrições. Uma boa parte destes algoritmos baseiam-se na informação local do funcional de custo, usualmente o gradiente e a Hessiana, para efectuar a condução em direcção a um mínimo local (ou global, se se reunirem determinadas condições). Quando o gradiente ou a Hessiana não estão disponíveis, são frequentemente aproximados numericamente.

Nos problemas com restrições, é usual tratar separadamente as restrições de igualdade das restrições de desigualdade. O método dos multiplicadores de Lagrange é habitualmente usado no primeiro caso. No segundo caso, distinguem-se ainda as restrições activas das inactivas. Enquanto as activas são tratadas como restrições de igualdade, as inactivas são ignoradas e a geometria Euclidiana é aplicada como habitual. A distinção entre restrições activas e inactivas leva a que possam existir comutações abruptas nas estruturas de dados e descontinuidades nas direcções seguidas.

Os métodos de ponto interior são outra classe de algoritmos onde o ponto óptimo é aproximado pelo interior da região admissível. Nestes algoritmos é definida uma função barreira que toma um custo infinito quando o ponto estimado se aproxima da fronteira, modificando desta forma a direcção seguida.

Todos estes métodos, sendo gerais, não consideram a estrutura particular do espaço de parâmetros em que os problemas são formulados. Dada uma determinada parametrização, a optimização é efectuada tal como se de um espaço Euclidiano se tratasse. Esta abordagem tem como consequência que várias parametrizações possam conduzir a desempenhos muito diferentes dos algoritmos. Embora em muitos casos, esta abordagem genérica seja suficiente, problemas há, em que a estrutura particular do problema permite construções mais ricas. É o caso das variedades estatísticas.

Nesta secção introduz-se o método do gradiente natural sugerido por Amari [1, 5] em substituição do gradiente Euclidiano. São desenvolvidos alguns exemplos de aplicação a distribuições de probabilidade discretas e à minimização da divergência de K-L. Deduzem-se as propriedades exibidas pelo método.

3.2 Método do gradiente natural

3.2.1 Definição de gradiente natural

Dada uma função suave $F(\theta)$, $\theta \in \mathbb{R}^n$, num espaço Euclidiano, é possível definir uma função que meça a taxa de variação de F numa direcção arbitrária dada por um vector v . Esta função designa-se por *diferencial* de F , e representa-se por $dF(v)$. É um operador linear que actua sobre o espaço vectorial onde v vive.

Dado um vector v , representado na forma de uma matriz coluna, então o diferencial dF pode ser representado pela matriz linha das derivadas parciais. A taxa de variação de F na direcção v é obtida pelo produto matricial

$$dF(v) = \begin{bmatrix} \frac{\partial F}{\partial \theta^1} & \cdots & \frac{\partial F}{\partial \theta^n} \end{bmatrix} \begin{bmatrix} v^1 \\ \vdots \\ v^n \end{bmatrix}.$$

Uma questão que se pode colocar em seguida é a determinação da direcção de maior subida de F . Esta pode ser determinada resolvendo um problema de optimização com restrições que se enuncia da seguinte maneira: “Pretende-se encontrar o vector v que maximiza $dF(v)$, sujeito à restrição $\|v\| = 1$ ”. Este problema pode ser resolvido usando o método dos multiplicadores de Lagrange como se mostra em seguida.

Define-se o funcional aumentado

$$L(v, \lambda) \stackrel{\text{def}}{=} dF(v) + \lambda(1 - \langle v, v \rangle),$$

e determinam-se os pontos de estacionariedade de L relativamente a v e λ simultaneamente. A solução depende do produto interno $\langle v, v \rangle$ usado para definir a restrição $\|v\| = 1$. Usando uma métrica $G(\theta)$ e fazendo $dL = \mathbf{0}$ resulta que (ver apêndice B.1)

$$\lambda v^T G = dF,$$

onde $\lambda = \|dL^T\|$ é seleccionado de maneira a satisfazer a restrição. Ao vector não normalizado λv dá-se o nome de *gradiente* e escreve-se ∇F . Usando esta definição de gradiente, a igualdade anterior pode escrever-se alternativamente como

$$\langle \nabla F, w \rangle_G = dF(w), \quad \forall w \neq \mathbf{0}. \quad (3.1)$$

Esta é a definição de gradiente que surge habitualmente em geometria Riemanniana [14]. Como se pode observar, o vector gradiente depende da métrica G do espaço, pelo que, uma mesma função F poderá ter vectores gradientes diferentes, consoante a métrica considerada. No caso Euclidiano, a métrica *standard* é a identidade, $G = I$, obtendo-se o vector gradiente

$$\nabla F = dF^T.$$

No entanto, para uma métrica genérica G , obtém-se

$$\nabla F = G^{-1} dF^T.$$

Para realçar o facto de que foi usada uma estrutura métrica particular, é frequente designar-se o gradiente por *gradiente natural* ou *Riemanniano*, em oposição ao *gradiente standard* ou *Euclidiano*. A notação $\tilde{\nabla} F$ e ∇F é usada para cada um dos casos, respectivamente.

No caso particular das variedades estatísticas, o gradiente natural é calculado relativamente à métrica de Fisher, sendo G^{-1} a inversa da matriz de informação G .

3.2.2 Método do gradiente natural

Dado um espaço de parâmetros Θ , e um funcional de custo $F : \Theta \rightarrow \mathbb{R}$, é frequente a utilização de métodos computacionais iterativos para a obtenção dos parâmetros $\theta \in \Theta$ que optimizam o funcional F . Um dos métodos usados é o método de descida de gradiente, também conhecido por *steepest-descent*, e que consiste em conduzir os parâmetros θ na direcção de maior descida do funcional. Idealmente, este método corresponde à integração da equação diferencial ordinária

$$\dot{\theta} = -\nabla F, \quad (3.2)$$

onde $\dot{\theta}$ representa o vector velocidade da curva percorrida pelos parâmetros.

Na prática, a equação anterior é discretizada grosseiramente pelo método de Euler, obtendo-se a equação mais usual

$$\theta_{\tau+1} = \theta_{\tau} - \eta_{\tau} \nabla F, \quad (3.3)$$

onde η_{τ} representa o passo da discretização e τ indexa a τ -ésima iteração do método.

A utilização do gradiente *standard* ∇F assume implicitamente uma geometria Euclidiana. Nos casos em que a geometria natural não é a Euclidiana, a direcção de *steepest-descent* é dada pelo gradiente natural, como se viu na secção anterior. Abordagens deste tipo têm sido sugeridas na literatura em vários domínios, ver p. ex. [16, 24]. Amari sugere [1, 2, 3, 5, 18] que seja usada esta direcção quando os espaços de pesquisa são variedades estatísticas. O método do gradiente natural tem encontrado sucesso em várias aplicações, nomeadamente no treino de perceptrões multicamada e a separação cega de fontes (ver referências anteriores).

A determinação do gradiente natural em cada iteração suscita uma questão de índole prática: tanto na expressão $\tilde{\nabla} F = G^{-1} \nabla F$, com os problemas numéricos associados à inversão explícita de G , como na solução do sistema linear $G \tilde{\nabla} F = \nabla F$ por métodos numericamente robustos, o custo computacional desta operação, quer em termos de tempo, quer de espaço, pode revelar-se suficientemente elevado para degradar o desempenho ou mesmo impossibilitar a sua realização em problemas de maior dimensão. Em seguida propõe-se um método rápido para calcular o gradiente natural a partir do gradiente *standard*, para espaços de distribuições de probabilidade discretas.

3.3 Cálculo rápido para distribuições discretas

Seja \mathcal{P} o conjunto das distribuições de probabilidade discretas $p(x)$, onde $x \in \{0, 1, 2, \dots, n\}$. A distribuição $p(x)$ admite várias representações (ver secção 2.2). Uma delas consiste em definir os parâmetros θ^i como

$$\theta^i \stackrel{\text{def}}{=} \Pr\{X = i\} = p(i), \quad i = 0, 1, 2, \dots, n. \quad (3.4)$$

Como a soma das probabilidades é um, Os $n + 1$ parâmetros não são todos independentes. No entanto, seleccionando de entre estes n parâmetros distintos, obtém-se uma representação de $p(x)$ bi-unívoca. Seja $(\theta^1, \dots, \theta^n)$ este sistema de coordenadas, enquanto o parâmetro dependente θ^0 é determinado por

$$\theta^0 \stackrel{\text{def}}{=} 1 - \sum_{i=1}^n \theta^i. \quad (3.5)$$

Como as probabilidades são números reais positivos, então $\theta^i > 0$, para $i = 0, 1, 2, \dots, n$. A restrição em θ^0 pode ainda ser escrita em termos das coordenadas independentes como

$$\sum_{i=1}^n \theta^i < 1. \quad (3.6)$$

À região do espaço de parâmetros que satisfaz estas restrições dá-se o nome de *região admissível*. É um conjunto convexo em \mathbb{R}^n . Tem no entanto uma estrutura mais rica que \mathbb{R}^n que vem do facto das probabilidades serem de certa maneira “especiais”. Para ilustrar esta estrutura especial, suponha-se que uma probabilidade θ^i é exactamente nula. Então o espaço tem na realidade uma dimensão menor que n , uma vez que o evento x_i nunca ocorre e, conseqüentemente, pode ser inteiramente removido da definição de $p(x)$. No entanto, se existir a mínima possibilidade de x_i ocorrer, a sua probabilidade é não nula, embora possa ser arbitrariamente pequena. Suponha-se que a probabilidade θ^i correspondente está perto de zero. Então, embora estritamente falando o espaço seja ainda n dimensional, este está de certa forma “perto” de perder uma dimensão. Este tipo de informação acerca da estrutura do espaço é capturada pela matriz de informação de Fisher.

Para distribuições de probabilidade discretas, como as acima descritas, a matriz de informação de Fisher tem componentes $g_{ij}(\theta)$ dadas por

$$\begin{aligned} g_{ij}(\theta) &= E_{\theta} \left[\frac{\partial \log p(x|\theta)}{\partial \theta^i} \frac{\partial \log p(x|\theta)}{\partial \theta^j} \right] \\ &= \sum_{x=0}^n \frac{1}{p(x|\theta)} \frac{\partial p(x|\theta)}{\partial \theta^i} \frac{\partial p(x|\theta)}{\partial \theta^j} \\ &= \frac{1}{p(0|\theta)} \frac{\partial p(0|\theta)}{\partial \theta^i} \frac{\partial p(0|\theta)}{\partial \theta^j} + \sum_{x=1}^n \frac{1}{p(x|\theta)} \frac{\partial p(x|\theta)}{\partial \theta^i} \frac{\partial p(x|\theta)}{\partial \theta^j} \\ &= \frac{1}{1 - \sum_{k=1}^n \theta^k} + \frac{\delta_{ij}}{\theta^i}, \end{aligned} \quad (3.7)$$

onde δ_{ij} é a função delta de Kronecker. Em notação matricial, a matriz de informação de Fisher G com estas componentes é

$$G(\theta) = \frac{1}{1 - \sum_{k=1}^n \theta^k} \begin{bmatrix} 1 & \cdots & 1 \\ \vdots & & \vdots \\ 1 & \cdots & 1 \end{bmatrix} + \begin{bmatrix} 1/\theta^1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & 1/\theta^n \end{bmatrix}. \quad (3.8)$$

Esta matriz é não singular na região admissível, mas contém elementos divergentes quando pelo menos uma das $n + 1$ probabilidades se aproxima de zero. Por este motivo, é mal condicionada nessas regiões do espaço de parâmetros, e a sua inversão explícita levanta problemas numéricos com frequência.

Para lidar com esta situação, a matriz G é reescrita como

$$G = A + bcb^T, \quad (3.9)$$

onde

$$A \stackrel{\text{def}}{=} \begin{bmatrix} 1/\theta^1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & 1/\theta^n \end{bmatrix}, \quad b \stackrel{\text{def}}{=} \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \quad (3.10)$$

e $c \stackrel{\text{def}}{=} 1/(1 - \sum_{i=1}^n \theta^i)$. Usando a identidade de Woodbury¹, a inversa de (3.9) pode ser escrita como

$$G^{-1} = A^{-1} - A^{-1}b(b^T A^{-1}b + c^{-1})^{-1}b^T A^{-1}, \quad (3.11)$$

¹ $(A + BCD)^{-1} = A^{-1} - A^{-1}B(DA^{-1}B + C^{-1})^{-1}DA^{-1}$, também conhecida como lema de inversão de matrizes.

e finalmente simplificada para

$$G^{-1} = \begin{bmatrix} \theta^1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \theta^n \end{bmatrix} - \begin{bmatrix} \theta^1 \\ \vdots \\ \theta^n \end{bmatrix} [\theta^1 \quad \dots \quad \theta^n]. \quad (3.12)$$

Como se pode observar, a inversa G^{-1} pode ser calculada de forma exacta efectuando apenas somas e multiplicações. Ao evitar as inversões presentes na fórmula original, obteve-se uma fórmula numericamente robusta para G^{-1} .

Tendo presente que o objectivo final é a obtenção do gradiente natural $\tilde{\nabla}F = G^{-1}\nabla F$, é ainda possível otimizar a carga computacional usada neste cálculo, tanto em termos de espaço como de tempo.

Escrevendo os parâmetros θ^i numa matriz coluna θ , então o gradiente natural $\tilde{\nabla}F$ pode ser obtido de ∇F fazendo

$$\tilde{\nabla}F = \theta \circ \nabla F - \theta (\theta \cdot \nabla F), \quad (3.13)$$

onde \circ representa o produto de Hadamard², e \cdot representa o produto interno *standard*³. A fórmula (3.13) permite o cálculo do gradiente natural sem construir explicitamente a inversa G^{-1} . Assim, em termos computacionais, a memória usada reduz-se à necessária para representar os três vectores θ , ∇F e $\tilde{\nabla}F$. Em termos de tempo de execução, calculado em número de operações, observa-se que o produto interno *standard* pode ser obtido do produto de Hadamard $\theta \circ \nabla F$ previamente calculado fazendo

$$\theta \cdot \nabla F = \sum_i (\theta \circ \nabla F)^i. \quad (3.14)$$

Assim, o número total de operações efectuadas para obter o gradiente natural $\tilde{\nabla}F$ consiste em $2n$ multiplicações e $2n + 1$ adições.

A fórmula (3.13) permite um escalamento linear do algoritmo no tempo e no espaço. Este resultado tem grande importância na implementação do algoritmo em problemas de dimensão elevada.

3.4 Propriedades do gradiente natural em famílias discretas

3.4.1 Pontos de estacionariedade

O fluxo do gradiente natural, escrito como

$$\dot{\theta} = -G^{-1}\nabla F, \quad (3.15)$$

tem como solução uma curva θ_τ que converge eventualmente para um ponto de equilíbrio da equação anterior. Os pontos de equilíbrio desta equação são os pontos que satisfazem $\dot{\theta} = \mathbf{0}$, isto é,

$$G^{-1}\nabla F = \mathbf{0}. \quad (3.16)$$

verifica-se imediatamente que, na região admissível, G^{-1} é não singular, pelo que o gradiente natural anula-se quando o gradiente standard ∇F se anula. Sobre a fronteira – região onde as restrições estão activas – a matriz G^{-1} é singular. O seu espaço nulo indica as direcções

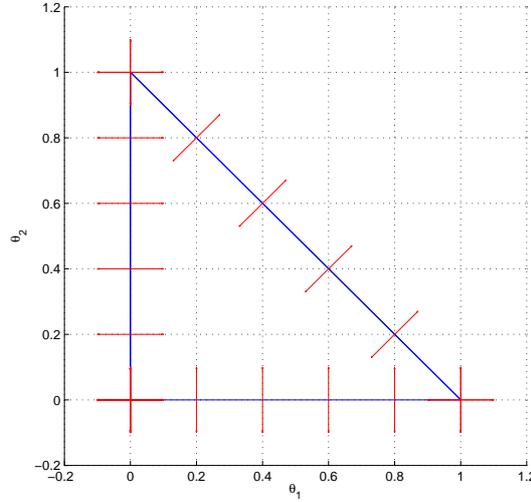


Figura 3.1: Espaço nulo de G^{-1} calculado sobre as restrições.

sobre as quais o gradiente Euclidiano é anulado. A figura 3.1 mostra estas direcções num problema de dimensão dois.

Conclui-se que podem existir pontos (ou regiões) de equilíbrio no método do gradiente natural, que não existem quando é usado o gradiente standard. Estas regiões de equilíbrio vão garantir a satisfação das restrições nos parâmetros θ , como se irá mostrar na secção seguinte.

3.4.2 Satisfação das restrições

A optimização do funcional F pelo método do gradiente *standard*, tem de ser formulado como um problema de optimização com restrições de desigualdade. As restrições são impostas para garantir que os parâmetros evoluem dentro da região admissível. Nesta secção, mostra-se que o mesmo funcional F , quando optimizado com o método do gradiente natural, satisfaz intrinsecamente as restrições

$$\theta^i > 0, \quad i = 1, \dots, n, \quad (3.17)$$

$$\sum_{k=1}^n \theta^k < 1. \quad (3.18)$$

Assim, em optimização de distribuições de probabilidade discretas, o método do gradiente natural pode ser encarado como um método sem restrições.

Para observar de que forma a matriz G^{-1} transforma a direcção seguida pelos parâmetros, definem-se as matrizes

$$\Lambda \stackrel{\text{def}}{=} \begin{bmatrix} \theta^1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \theta^n \end{bmatrix}, \quad \sqrt{\theta} \stackrel{\text{def}}{=} \begin{bmatrix} \sqrt{\theta^1} \\ \vdots \\ \sqrt{\theta^n} \end{bmatrix}, \quad (3.19)$$

²O produto de Hadamard (ou produto de Shur) é o produto elemento a elemento: $(a \circ b)^i = a^i b^i$.

³O produto interno standard (*dot product*) é definido por: $a \cdot b = \sum_i a^i b^i$.

e define-se $\sqrt{\Lambda}$ como a única matriz definida positiva tal que $\Lambda = \sqrt{\Lambda}\sqrt{\Lambda}$, i.e.,

$$\sqrt{\Lambda} = \begin{bmatrix} \sqrt{\theta^1} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \sqrt{\theta^n} \end{bmatrix}. \quad (3.20)$$

Então, como $\theta = \sqrt{\Lambda}\sqrt{\theta}$, a equação (3.12) pode ser factorizada da seguinte maneira:

$$\begin{aligned} G^{-1}(\theta) &= \sqrt{\Lambda}\sqrt{\Lambda}^T - \sqrt{\Lambda}\sqrt{\theta}\sqrt{\theta}^T\sqrt{\Lambda}^T \\ &= \sqrt{\Lambda}(\mathbf{I} - \sqrt{\theta}\sqrt{\theta}^T)\sqrt{\Lambda}^T \\ &= \sqrt{\Lambda}\left(\mathbf{I} - \|\sqrt{\theta}\|^2 \frac{\sqrt{\theta}}{\|\sqrt{\theta}\|} \frac{\sqrt{\theta}^T}{\|\sqrt{\theta}\|}\right)\sqrt{\Lambda}^T, \end{aligned} \quad (3.21)$$

onde $\sqrt{\theta}/\|\sqrt{\theta}\|$ são vectores unitários relativamente à norma Euclidiana. Verifica-se ainda que

$$\|\sqrt{\theta}\|^2 = \sqrt{\theta}^T\sqrt{\theta} = \sum_{i=1}^n \theta^i = 1 - \theta^0. \quad (3.22)$$

Finalmente, definindo as matrizes

$$\mathbf{H}_0 \stackrel{\text{def}}{=} \mathbf{I} - (1 - \theta^0) \frac{\sqrt{\theta}}{\|\sqrt{\theta}\|} \frac{\sqrt{\theta}^T}{\|\sqrt{\theta}\|}, \quad (3.23)$$

$$\mathbf{H}_i \stackrel{\text{def}}{=} \mathbf{I} - (1 - \theta^i) \mathbf{e}_i \mathbf{e}_i^T, \quad i = 1, \dots, n, \quad (3.24)$$

onde $\mathbf{e}_i^T = [0 \ \dots \ 0 \ 1 \ 0 \ \dots \ 0]$ é o i -ésimo vector da base *standard*, então a equação (3.21) pode ser escrita como o produto

$$G^{-1} = (\sqrt{\Lambda})^{-1} \mathbf{H}_n \cdots \mathbf{H}_2 \mathbf{H}_1 \mathbf{H}_0 (\sqrt{\Lambda}), \quad (3.25)$$

ou alternativamente

$$G^{-1} = (\sqrt{\Lambda}) \mathbf{H}_0 \mathbf{H}_1 \mathbf{H}_2 \cdots \mathbf{H}_n (\sqrt{\Lambda})^{-1}. \quad (3.26)$$

As equações (3.25) e (3.26) anteriores permitem uma nova interpretação do papel desempenhado pela métrica no cálculo do gradiente natural. Em primeiro lugar, é efectuada uma mudança de base definida por $(\sqrt{\Lambda})$. A figura 3.2 mostra, para $n = 2$, o efeito desta transformação sobre θ quando se fixa cada um dos três parâmetros $(\theta^0, \theta^1, \theta^2)$, individualmente.

Na nova base, o vector transformado sofre um escalamento definido pelo produto $\mathbf{H}_0 \cdots \mathbf{H}_n$. Cada um destes factores atenua a componente do gradiente na direcção da respectiva restrição. Tomando como exemplo a coordenada θ^1 , então a transformação

$$\mathbf{H}_1 = \begin{bmatrix} \theta^1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{bmatrix}. \quad (3.27)$$

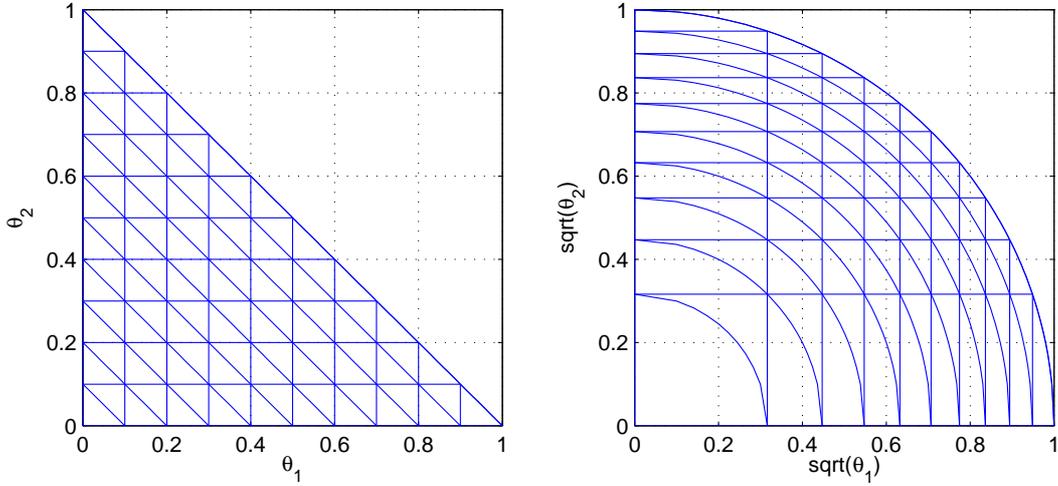


Figura 3.2: Curvas de parâmetros constantes em coordenadas θ e $\sqrt{\theta}$ à esquerda e direita, respectivamente. As diagonais correspondentes a vários valores de θ^0 no gráfico da esquerda, são mapeadas para as circunferências correspondentes a $\|\sqrt{\theta}\|$ constante no gráfico da direita.

atenua a primeira componente do vector gradiente por um factor θ^1 , deixando as restantes componentes inalteradas. Na situação limite, com $\theta^1 = 0$, a transformação

$$\mathbf{H}_1 = \begin{bmatrix} 0 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{bmatrix} \quad (3.28)$$

efectua uma projecção para o subespaço definido por $\theta^1 = 0$. Assim, esta projecção entra suavemente em acção à medida que as coordenadas se aproximam da restrição. As restantes transformações \mathbf{H}_2 até \mathbf{H}_n actuam da mesma maneira nas respectivas restrições.

O factor \mathbf{H}_0 é, contudo, diferente dos outros. Este produz um escalamento θ^0 na direcção $\sqrt{\theta}$ (ver apêndice B.2). Quando $\theta^0 = 0$ o efeito é o de uma projecção sobre a esfera unitária $\|\sqrt{\theta}\| = 1$. Quando $\theta^0 \neq 0$, a projecção é incompleta e a componente ortogonal à superfície não é totalmente anulada. Assim, o factor \mathbf{H}_0 desempenha um papel em tudo semelhante aos restantes factores, mas neste caso relativamente à esfera unitária, que nas coordenadas originais θ corresponde à restrição $\sum_{i=1}^n \theta^i < 1$. Observa-se ainda que, nas novas coordenadas, os escalamentos são efectuados de forma independente relativamente a cada restrição, *i.e.*, as restrições intersectam-se mutuamente em ângulos de 90° .

No cálculo do gradiente natural, a acção combinada dos factores \mathbf{H}_0 a \mathbf{H}_n encurta o gradiente *standard* nas direcções das restrições. Este encurtamento é mais acentuado nas componentes do vector para as quais as restrições estão mais próximas. Nos casos extremos, quando uma restrição é “tocada”, G^{-1} efectua uma projecção do vector sobre essa restrição da mesma forma que o método de projecção de Rosen [17, 9, 29]. Como consequência dessa projecção, as restrições definem conjuntos invariantes sobre os quais o fluxo do gradiente natural evolui.

Lema 4. *As restrições com igualdade $\sum_{i=1}^n \theta^i = 1$ e $\theta^i = 0$, $i = 1, \dots, n$ definem conjuntos invariantes da equação $\dot{\theta} = -G^{-1}\nabla F$.*

Demonstração. Começa-se por mostrar que a afirmação é verdadeira para $\theta^1 = 0$. Tendo em consideração que todas as matrizes são diagonais, à excepção de \mathbf{H}_0 , então pode fazer-se a seguinte reordenação:

$$\dot{\theta} = \mathbf{H}_1 \mathbf{H}_n \cdots \mathbf{H}_2 (\sqrt{\Lambda})^{-1} \mathbf{H}_0 (\sqrt{\Lambda}) \nabla F. \quad (3.29)$$

Torna-se claro que \mathbf{H}_1 anula a primeira componente de $\dot{\theta}$, isto é, $\dot{\theta}^1 = 0$ e portanto $\theta^1 = 0$ é um conjunto invariante. O mesmo raciocínio aplica-se aos restantes factores \mathbf{H}_2 até \mathbf{H}_n . Para provar que a restrição de igualdade $\sum_{i=1}^n \theta^i = 1$ também define um conjunto invariante, mostra-se que

$$\sum_{i=1}^n \dot{\theta}^i = 0. \quad (3.30)$$

Para o efeito, escreve-se

$$\dot{\theta} = (\sqrt{\Lambda}) \mathbf{H}_0 (\sqrt{\Lambda})^{-1} \mathbf{H}_n \cdots \mathbf{H}_1 \nabla F. \quad (3.31)$$

Colapsando os termos $\mathbf{H}_n \cdots \mathbf{H}_1 \nabla F$ num vector \mathbf{v} obtém-se:

$$\begin{aligned} [1 \quad \cdots \quad 1] \dot{\theta} &= [1 \quad \cdots \quad 1] (\sqrt{\Lambda}) \mathbf{H}_0 (\sqrt{\Lambda})^{-1} \mathbf{v} \\ &= [1 \quad \cdots \quad 1] (\sqrt{\Lambda}) (\mathbf{I} - \sqrt{\theta} \sqrt{\theta}^T) (\sqrt{\Lambda})^{-1} \mathbf{v} \\ &= [1 \quad \cdots \quad 1] (\mathbf{I} - \theta [1 \quad \cdots \quad 1]) \mathbf{v} \\ &= [1 \quad \cdots \quad 1] \mathbf{v} - \sum_{i=1}^n \theta^i [1 \quad \cdots \quad 1] \mathbf{v} \\ &= \underbrace{\left(1 - \sum_{i=1}^n \theta^i\right)}_{=0} [1 \quad \cdots \quad 1] \mathbf{v} \\ &= 0. \end{aligned} \quad (3.32)$$

Assim, se a condição inicial do método satisfaz $\sum_{i=1}^n \theta^i = 1$, então a equação (3.32) garante que esta propriedade se mantém na evolução futura de θ . \square

O resultado seguinte garante que, dentro de certas condições, o método do gradiente natural pode ser usado como método sem restrições.

Teorema 5. *Considere-se o espaço das distribuições de probabilidade discretas $p(x|\theta)$ e respectiva região admissível $\Theta \stackrel{\text{def}}{=} \{\theta \mid \theta^i > 0, \sum \theta^i < 1\}$. Seja $F(\theta)$ uma função suave, cujo domínio contém a região admissível, e tal que ∇F é contínua à Lipschitz. Então, partindo de um θ_τ admissível, o fluxo do gradiente natural é admissível (i.e., está contido na região admissível).*

Demonstração. As condições do teorema garantem a existência e unicidade de solução pelo teorema de Picard. Resta provar que condições iniciais no interior da região admissível evoluem na região admissível. O lema 4 mostra que a região admissível Θ tem como fronteira um conjunto invariante. Como o fluxo gradiente é contínuo, então este está limitado à região admissível. \square

Conclui-se que, ao contrário dos métodos tradicionais de optimização com restrições, o método do gradiente natural garante intrinsecamente a não violação das restrições de forma suave. A inibição de certas direcções na evolução depende da proximidade dos parâmetros às restrições.

3.5 Versão discreta do método do gradiente natural

Nas secções anteriores, a análise do método do gradiente natural foi realizada relativamente à trajectória “ideal” obtida por integração da equação diferencial $\dot{\theta} = -\tilde{\nabla}F$. Nas condições em que a trajectória obtida é contínua, provou-se que o algoritmo não viola as restrições. Na versão discretizada

$$\theta_{\tau+1} = \theta_{\tau} - \eta_{\tau} \tilde{\nabla}F, \quad (3.33)$$

os passos $\eta_{\tau} > 0$ podem conduzir a situações em que as restrições nos parâmetros são violadas na actualização dos parâmetros, obtendo-se $\theta \notin \Theta$. Nestas condições o passo η_{τ} deve ser corrigido para que os parâmetros se mantenham na região admissível Θ .

Uma forma de forçar os parâmetros a manterem-se na região admissível Θ consiste em determinar, em cada iteração τ , um intervalo admissível I_{τ} para o passo, tal que $\eta_{\tau} \in I_{\tau}$ garante $\theta_{\tau+1} \in \Theta$.

3.5.1 Intervalos admissíveis para o passo

Para garantir que a actualização dos parâmetros de θ_{τ} para $\theta_{\tau+1}$ não viola as restrições, é necessário impor

$$\theta_{\tau+1}^i > 0, \quad i = 1, \dots, n, \quad (3.34)$$

e

$$\sum_{i=1}^n \theta_{\tau+1}^i < 1, \quad (3.35)$$

obtendo-se, respectivamente,

$$\theta_{\tau}^i - \eta_{\tau} (\tilde{\nabla}F)^i > 0, \quad i = 1, \dots, n \quad (3.36)$$

e

$$\sum_{i=1}^n (\theta_{\tau}^i - \eta_{\tau} (\tilde{\nabla}F)^i) < 1. \quad (3.37)$$

Resolvendo as inequações (3.36) e (3.37) relativamente a η_{τ} positivo, correspondente a um procedimento de minimização, obtém-se (ver apêndice B.3) que

$$\eta_{\tau} < \frac{1}{\max(\alpha) - \theta \cdot \nabla F}, \quad \alpha \stackrel{\text{def}}{=} \{0, \nabla F^1, \dots, \nabla F^n\}. \quad (3.38)$$

Isto é, se em cada iteração do método do gradiente natural o passo η_{τ} satisfaz (3.38), então os parâmetros θ_{τ} evoluem dentro da região admissível Θ .

O teste (3.38) requer duas operações escalares (subtração e divisão) e o cálculo do valor máximo do conjunto α de $n+1$ escalares. O produto interno $\theta \cdot \nabla F$ está disponível do cálculo do gradiente natural efectuado em (3.14). Assim, e tal como o cálculo do gradiente natural, o teste escala linearmente com o número de parâmetros em tempo de cálculo e não requer espaço adicional.

Em problemas de maximização, correspondentes a considerar $\eta_{\tau} < 0$ na equação (3.33), existem condições análogas para o passo. Verifica-se, para esses casos, que o passo deve satisfazer

$$\eta_{\tau} > \frac{1}{\min(\alpha) - \theta \cdot \nabla F}. \quad (3.39)$$

Em suma, na versão discretizada do método do gradiente natural, quer em problemas de minimização quer de maximização, o passo deve satisfazer $\eta_\tau \in I_\tau$, com

$$I_\tau \stackrel{\text{def}}{=} \left] \frac{1}{\min(\alpha) - \theta \cdot \nabla F}, \frac{1}{\max(\alpha) - \theta \cdot \nabla F} \right[\quad (3.40)$$

e

$$\alpha \stackrel{\text{def}}{=} \{0, \nabla F^1, \dots, \nabla F^n\}. \quad (3.41)$$

3.5.2 Condição de paragem do algoritmo

O método do gradiente natural aproxima assintoticamente a solução óptima do problema. À medida que a solução estimada se aproxima da solução óptima, os incrementos nos parâmetros são progressivamente menores, tornando-se negligenciáveis ao fim de algum tempo. Assim, torna-se necessário introduzir um critério de paragem que interrompa a optimização quando se considera a solução estimada como sendo adequada.

Um critério de paragem possível consiste em comparar as componentes $\partial F / \partial \theta^i$ com um limiar ϵ , e parar se todas as derivadas parciais são inferiores (em módulo) a ϵ . Este critério apresenta, no entanto, dois problemas:

1. não toma em consideração a geometria natural do espaço;
2. devido à curvatura da função F , os parâmetros podem ser mais sensíveis numas direcções do que noutras, pelo que o mesmo ϵ para todas as direcções não é apropriado.

Outro critério possível seria testar a norma Euclidiana do gradiente standard $\|\nabla F\|$ em vez da norma do máximo proposta atrás. Também esta norma não se mostra apropriada, uma vez que não toma em consideração a métrica natural do espaço.

Como solução para estes problemas, propõe-se que seja usada a norma induzida pela métrica de Fischer:

$$\|\mathbf{v}\|_g \stackrel{\text{def}}{=} \mathbf{v}^T G \mathbf{v}. \quad (3.42)$$

Ao aplicar esta norma ao gradiente natural $\tilde{\nabla} F$, obtém-se

$$\begin{aligned} \|\tilde{\nabla} F\|_g &= \tilde{\nabla} F^T G \tilde{\nabla} F \\ &= \nabla F^T G^{-1} \nabla F. \end{aligned} \quad (3.43)$$

Observando que $G^{-1} \nabla F = \tilde{\nabla} F$, obtém-se uma fórmula computacionalmente eficiente para o cálculo da norma que não faz intervir as matrizes G ou G^{-1} :

$$\|\tilde{\nabla} F\|_g = \nabla F^T \tilde{\nabla} F. \quad (3.44)$$

Esta fórmula requer apenas um produto interno *standard* entre dois vectores disponibilizados pelo algoritmo.

Tendo em consideração que $\nabla F^T = dF$, então equação (3.44) pode ainda ser interpretada como

$$\|\tilde{\nabla} F\|_g = dF(\tilde{\nabla} F), \quad (3.45)$$

isto é, a maior taxa de variação de F , do ponto de vista Riemanniano. Finalmente, se esta variação for inferior a um limiar ϵ , então o método iterativo é interrompido e os parâmetros actuais θ são considerados a melhor aproximação dos parâmetros óptimos.

3.6 Minimização da divergência de Kullback-Leibler

Nesta secção aplica-se o método do gradiente natural a dois problemas de baixa dimensão. O objectivo é puramente ilustrativo e pretende mostrar a superioridade do método relativamente ao gradiente *standard*.

O primeiro problema a otimizar é formulado da seguinte maneira: Pretende-se encontrar a distribuição de probabilidade discreta $p(x)$ que minimiza a divergência K-L $D(p||q)$ a uma distribuição ideal $q(x)$. Este é um problema artificial, uma vez que a solução $p(x) = q(x)$ já é conhecida. O seu interesse advém de poder servir de banco de ensaio onde os dois métodos são testados.

No segundo exemplo apresentado, a mesma divergência de K-L é minimizada, mas agora a distribuição p não é completamente livre. Assume-se que p é uma distribuição conjunta gerada pela regra da cadeia das probabilidades $p_{X,Y}(x,y) = p_{Y|X}(y|x)p_X(x)$, onde apenas $p_X(x)$ é livre. Este problema está relacionado em parte com o problema de controlo probabilístico tratado no capítulo 4 desta tese.

Exemplo 6. Usando a parametrização usada anteriormente, a distribuição de probabilidade $p(x)$ escreve-se, em notação matricial,

$$\mathbf{p} = \begin{bmatrix} \theta^0 \\ \theta^1 \\ \vdots \\ \theta^n \end{bmatrix}, \quad (3.46)$$

onde $(\theta^1, \dots, \theta^n)$ são parâmetros livres e $\theta^0 = 1 - \sum_{i=1}^n \theta^i$ é um parâmetro dependente. Calculando o gradiente *standard* da divergência $D(p||q)$, obtém-se

$$\nabla D = \begin{bmatrix} -1 & 1 & \mathbf{0} \\ \vdots & & \ddots \\ -1 & \mathbf{0} & 1 \end{bmatrix} (\log \mathbf{p} - \log \mathbf{q}), \quad (3.47)$$

sendo os logaritmos tomados elemento a elemento sobre os vectores de probabilidades \mathbf{p} e \mathbf{q} . A estrutura particular da matriz da esquerda em (3.47) admite uma fórmula de cálculo mais eficiente, sem recurso a multiplicações, que é a seguinte:

$$\mathbf{v} = \log \mathbf{p} - \log \mathbf{q} \quad (3.48)$$

$$\nabla D = \begin{bmatrix} v^1 - v^0 \\ v^2 - v^0 \\ \vdots \\ v^n - v^0 \end{bmatrix}. \quad (3.49)$$

Para concretizar, suponha-se que $\mathbf{q} = [0.2494 \quad 0.0025 \quad 0.7481]^T$ é a distribuição objectivo e $\mathbf{p} = [\frac{1}{3} \quad \frac{1}{3} \quad \frac{1}{3}]^T$ a estimativa inicial de \mathbf{p} .

A utilização do método do gradiente *standard*, com passo fixo $\eta = 0.01$, resulta num comportamento errático e não converge para a solução, como se pode observar na figura 3.3 à esquerda. Este exemplo realça duas fraquezas conhecidas deste método:

1. Convergência lenta quando a função a otimizar exhibe, nas várias direcções, curvaturas muito diferentes;

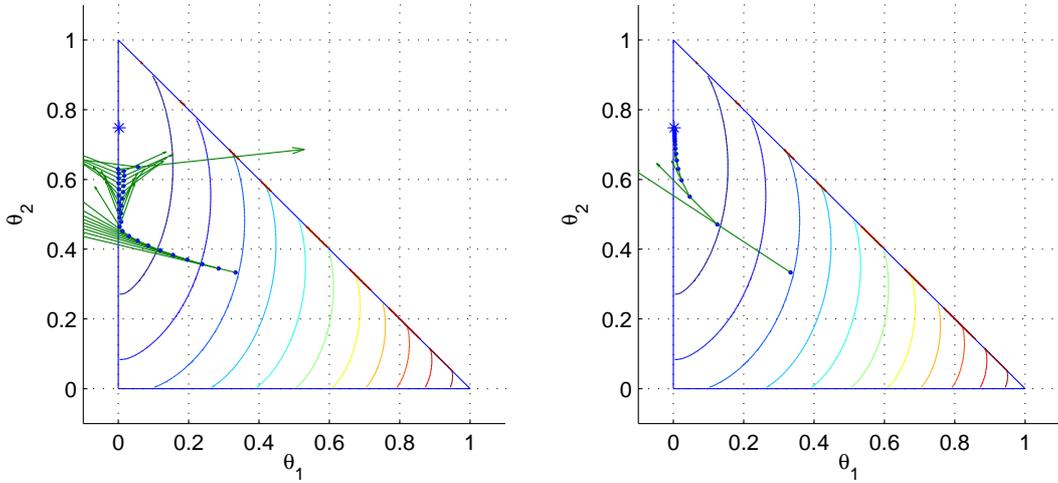


Figura 3.3: Esquerda: método do gradiente standard com passo fixo; Direita: método do gradiente natural com passo fixo.

2. A utilização de um passo fixo não produz bons resultados. O método torna-se instável quando o passo é grande, e converge muito lentamente quando é pequeno.

O mesmo problema é resolvido com o método do gradiente natural, usando um passo fixo de $\eta = 0.18$. A figura 3.3 mostra, do lado direito, como a convergência é muito mais rápida neste caso. A razão para este comportamento, tem a haver com o facto que, perto do ponto óptimo \mathbf{q} , a Hessiana da divergência de K-L é igual à matriz de informação de Fisher (ver secção 2.5). Assim, o método pode ser considerado como um método quasi-Newton, de métrica variável.

Exemplo 7. Neste exemplo, pretende-se otimizar a divergência de K-L, $D(p\|q)$, entre a conjunta $p_{X,Y}(x,y)$ e uma distribuição ideal $q(x,y)$. Assume-se que a distribuição $p_{X,Y}(x,y)$ é gerada pela regra da cadeia

$$p_{X,Y}(x,y) = p_{Y|X}(y|x)p_X(x), \quad (3.50)$$

onde $x \in \{0, 1, \dots, n\}$, $y \in \{0, 1, \dots, m\}$, e onde $p_{Y|X}(y|x)$ é previamente especificada, enquanto $p_X(x)$ é livre. A optimização é efectuada nesta última distribuição. Para facilitar a notação, os subscritos em p_X , $p_{X,Y}$ e $p_{Y|X}$ não são escritos sempre que forem subentendidos do contexto.

O gradiente de $D(p(x,y)\|q(x,y))$ relativamente aos parâmetros θ que especificam $p(x)$ tem componentes dadas por

$$\frac{\partial D}{\partial \theta^i} = \sum_{x=0}^n \sum_{y=0}^m p(y|x) \log \frac{p(y|x)p(x)}{q(x,y)} \frac{\partial p(x)}{\partial \theta^i}. \quad (3.51)$$

Em notação matricial, seja \mathbf{p}_X a matriz coluna das probabilidades $p(x)$, $\mathbf{P}_{Y|X}$ a matriz de transição $m \times n$ de $p(y|x)$, e $\mathbf{P}_{Y,X}$ e $\mathbf{Q}_{Y,X}$ as matrizes das probabilidades conjuntas $p(x,y)$ e $q(x,y)$, respectivamente. As posições dos subscritos X e Y são usados nas matrizes para indicar a ordem dos índices, de maneira que $\mathbf{P}_{X,Y} = (\mathbf{P}_{Y,X})^T$. Usando esta notação, o

gradiente standard ∇D pode ser calculado pelas três equações

$$\mathbf{P}_{Y,X} = \mathbf{P}_{Y|X} \circ (\mathbf{p}_X [1 \ \cdots \ 1])^T, \quad (3.52)$$

$$\mathbf{T}_{Y,X} = \mathbf{P}_{Y|X} \circ (\log \mathbf{P}_{Y,X} - \log \mathbf{Q}_{Y,X}), \quad (3.53)$$

$$\nabla D = \begin{bmatrix} -1 & 1 & \mathbf{0} \\ \vdots & & \ddots \\ -1 & \mathbf{0} & 1 \end{bmatrix} \mathbf{T}_{X,Y} \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \quad (3.54)$$

$$= \sum_{y=0}^m \begin{bmatrix} \mathbf{T}_{1y} - \mathbf{T}_{0y} \\ \mathbf{T}_{2y} - \mathbf{T}_{0y} \\ \vdots \\ \mathbf{T}_{ny} - \mathbf{T}_{0y} \end{bmatrix}, \quad (3.55)$$

onde os logaritmos são tomados elemento a elemento sobre as matrizes $\mathbf{P}_{Y,X}$ e $\mathbf{Q}_{Y,X}$.

A matriz G^{-1} é calculada da mesma forma que no exemplo anterior, uma vez que é calculada apenas para $p(x)$. O gradiente natural é, então, dado pela fórmula

$$\tilde{\nabla} D = \theta \circ \nabla D - \theta (\theta \cdot \nabla D). \quad (3.56)$$

Para ilustrar a evolução dos dois métodos – standard e natural – construiu-se um problema de baixa dimensão onde $\mathbf{P}_{X,Y}$ e $\mathbf{Q}_{X,Y}$ são matrizes de dimensão 3×2 . A figura 3.4 mostra a evolução dos parâmetros de $p(x)$ para ambos os métodos. A interpretação é semelhante à do exemplo anterior. A figura 3.5 mostra a evolução do custo $D(p(x,y)||q(x,y))$. Observa-se que, ao contrário do exemplo 6, o custo não converge para zero. Neste caso não existe nenhuma distribuição $p(x)$ que satisfaz com igualdade

$$p(x)p(y|x) = q(x,y). \quad (3.57)$$

3.7 Gradiente natural da divergência K-L em famílias mistura

Nesta secção apresenta-se um método de optimização equivalente ao método do gradiente natural baseado numa perspectiva geométrica das famílias mistura e exponencial. O problema que se coloca é o seguinte: Dada uma família mistura \mathcal{M} , e uma distribuição $q \notin \mathcal{M}$, o gradiente natural da função $D(m||q)$, tomada no ponto $m \in \mathcal{M}$, é um vector

$$\tilde{\nabla} D(m||q) = \sum_j a^j \frac{\partial \log m}{\partial \theta^j}, \quad (3.58)$$

cujas componentes a^j são dadas pela fórmula

$$\begin{aligned} a^j &= \sum_{i=1}^n g^{ij} \frac{\partial D}{\partial \theta^i} \\ &= \sum_{i=1}^n g^{ij} \int (p_i - p_0) \log \frac{m}{q} dx. \end{aligned} \quad (3.59)$$

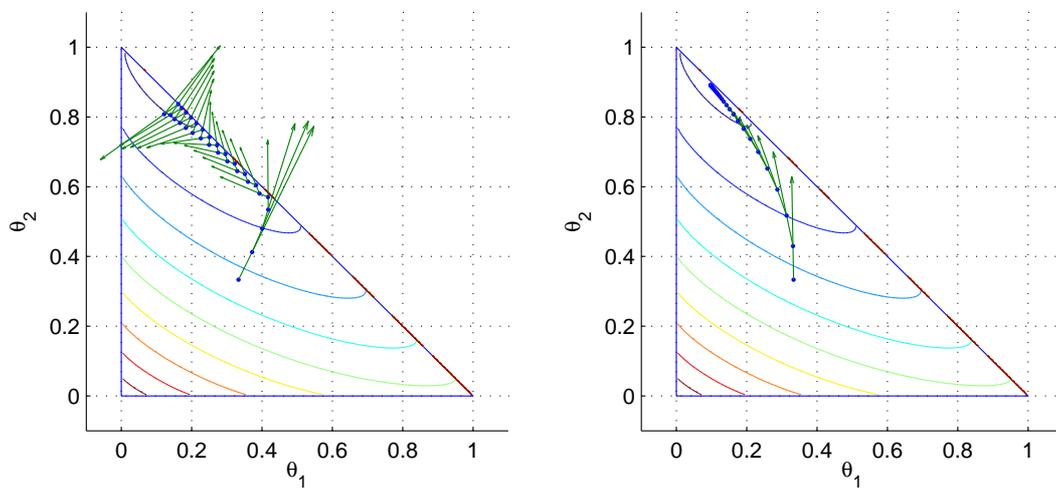


Figura 3.4: Evolução dos parâmetros θ nos métodos do gradiente standard (esq.) e natural (dir.).

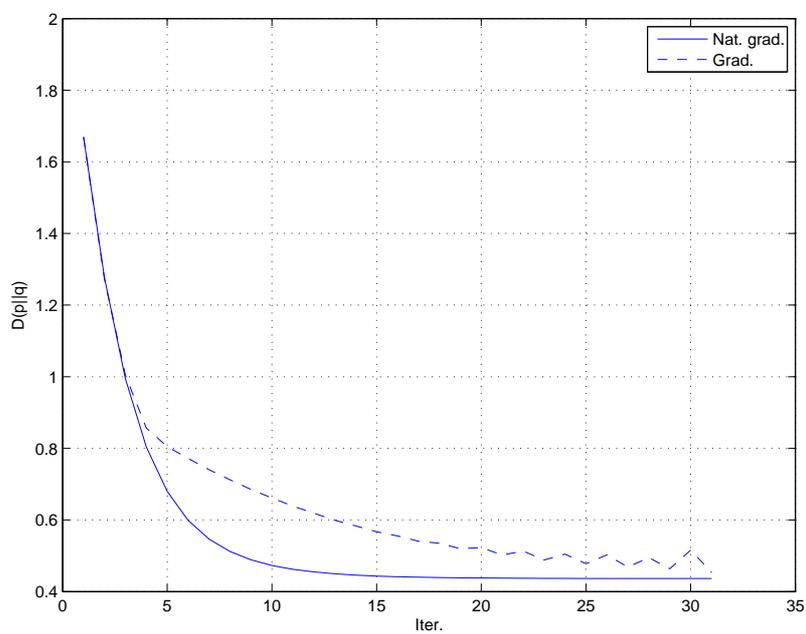


Figura 3.5: Evolução do custo $D(p||q)$ nos métodos do gradiente standard e natural.

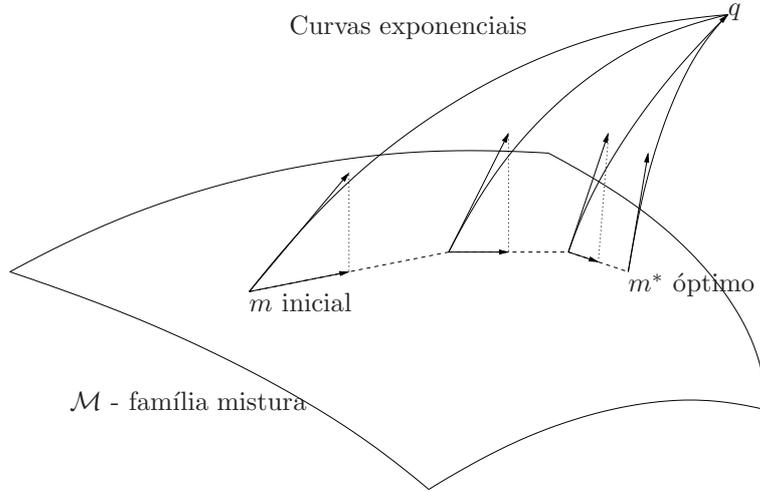


Figura 3.6: Algoritmo iterativo para otimização de $D(m||q)$.

O método do gradiente natural consiste em actualizar os parâmetros θ na direcção de $\tilde{\nabla}D$ ou $-\tilde{\nabla}D$, conforme se trata de um problema de maximização ou minimização. Em seguida mostra-se que este procedimento é equivalente a um procedimento geométrico baseado no teorema de Pitágoras.

Sabe-se que a distribuição $m^* \in \mathcal{M}$ que minimiza a divergência $D(m||q)$ é tal que a curva exponencial que une m^* a q intersecta ortogonalmente \mathcal{M} . Tal não se verifica para as restantes distribuições $m \in \mathcal{M}$.

Considerando o vector tangente a uma curva exponencial, tomado no respectivo ponto de partida $m \in \mathcal{M}$, é possível efectuar a sua projecção no espaço tangente usando a métrica de Fisher. Esta projecção não se anula a não ser para a distribuição óptima m^* , para a qual a curva exponencial é ortogonal a \mathcal{M} .

Estes factos sugerem que a pesquisa da distribuição óptima m^* pode ser efectuada seguindo a direcção da “sombra” da curva exponencial, tomada em cada passo. A figura 3.6 ilustra este procedimento.

A optimização resume-se no seguinte algoritmo:

1. Especificar uma estimativa inicial $m \in \mathcal{M}$.
2. Definir a curva exponencial que une m a q :

$$r(x|a) \stackrel{\text{def}}{=} \frac{m^{1-a} q^a}{\gamma(a)}, \quad (3.60)$$

onde $\gamma(a)$ é um factor de normalização.

3. Calcular o vector tangente à curva r no ponto m :

$$\left. \frac{\partial \log r}{\partial a} \right|_{a=0}. \quad (3.61)$$

4. Efectuar a projecção do vector tangente à curva para o espaço tangente $T_m \mathcal{M}$.
5. Actualizar os parâmetros θ na direcção do vector tangente.

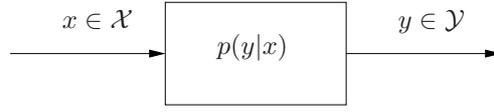


Figura 3.7: Canal discreto sem memória.

6. Repetir os passos 2 a 5 até que a projecção se anule.

Efectuando os passos descritos no algoritmo anterior, obtém-se um vector tangente com componentes

$$\begin{aligned}
 b^j &= \sum_{i=1}^n g^{ij} E_p \left[\frac{\partial \log m}{\partial \theta^i} \frac{\partial \log r}{\partial a} \Big|_{a=0} \right] \\
 &= \sum_{i=1}^n g^{ij} \int_{\mathcal{X}} m \frac{\partial \log m}{\partial \theta^i} \log \frac{q}{m} dx \\
 &= \sum_{i=1}^n g^{ij} \int_{\mathcal{X}} (p_i - p_0) \log \frac{q}{m} dx \\
 &= -a^j,
 \end{aligned} \tag{3.62}$$

ou seja, um vector com direcção contrária à do gradiente natural (3.59). Assim, a aplicação deste procedimento resulta precisamente no algoritmo do gradiente natural da secção anterior.

Este resultado fornece uma interpretação geométrica compatível com o método do gradiente natural para a optimização da divergência de K-L em famílias mistura. A sua validade está restrita às situações para as quais a interpretação Pitagoriana da divergência K-L é válida, conforme observado na secção 2.5.1.

3.8 Capacidade de um canal discreto sem memória

Um canal discreto sem memória (figura 3.7) é caracterizado pela tripla $(\mathcal{X}, \mathcal{Y}, p(y|x))$, onde

- $\mathcal{X} \stackrel{\text{def}}{=} \{0, 1, 2, \dots, n\}$ é um alfabeto finito de símbolos de entrada;
- $\mathcal{Y} \stackrel{\text{def}}{=} \{0, 1, 2, \dots, m\}$ é um alfabeto finito de símbolos de saída;
- $p(y|x)$ são as probabilidades de transição que caracterizam, para cada $x \in \mathcal{X}$ enviado, a probabilidade de ser recebido cada um dos símbolos $y \in \mathcal{Y}$.

Intuitivamente, a capacidade de um canal é um limite assintótico na quantidade máxima de informação que é possível transmitir de forma fiável por esse canal [30, 12]. Formalmente, a capacidade do canal é definida por

$$C \stackrel{\text{def}}{=} \sup_{p(x)} I(X; Y), \tag{3.63}$$

onde $I(X; Y)$ é a informação mútua entre as variáveis aleatórias X e Y . A informação mútua pode ser escrita como a divergência de Kullback-Leibler entre a distribuição conjunta $p(x, y)$ e o produto das marginais $p(x)p(y)$:

$$I(X; Y) \stackrel{\text{def}}{=} \sum_{x \in \mathcal{X}, y \in \mathcal{Y}} p(x, y) \log \frac{p(x, y)}{p(x)p(y)}. \tag{3.64}$$

Desta forma, se X e Y são estatisticamente independentes, então $p(x, y) = p(x)p(y)$ e a sua informação mútua é nula. Caso contrário, a informação mútua é um número positivo que indica o grau de dependência entre as variáveis. Se ambas as variáveis estão relacionadas de forma determinística pela fórmula $y = f(x)$, então a informação mútua atinge o seu valor máximo $I(X; Y) = H(Y) \leq H(X)$, com igualdade se f é injectiva.

No caso de um canal discreto sem memória, a distribuição conjunta é obtida pelo produto

$$p(x, y) = p(y|x)p(x) \quad (3.65)$$

onde $p(y|x)$ é o modelo do canal. Assim, a informação mútua (3.64) é dependente da distribuição $p(x)$ dos símbolos à entrada do canal. Uma escolha apropriada de $p(x)$ permite maximizar a sua expressão.

Entre os métodos mais conhecidos para a determinação da capacidade do canal encontram-se os métodos iterativos de Blahut e Arimoto [10, 7], consistindo cada iteração do método em dois passos de optimização diferentes. Mais recentemente, estes métodos foram interpretados recorrendo a ferramentas de geometria da informação. Em [25], um dos passos do método de Blahut-Arimoto é substituído pelo método do gradiente natural numa família exponencial, obtendo-se uma regra de actualização multiplicativa para as probabilidades. Apresenta-se em seguida uma abordagem alternativa, que consiste na aplicação do método do gradiente natural directamente sobre a função a otimizar $I(X; Y)$.

3.8.1 Diferencial da informação mútua

Parametrizando a distribuição de probabilidade $p(x|\theta)$ tal como nas secções anteriores, obtêm-se as derivadas parciais

$$\frac{\partial I}{\partial \theta^i} = \sum_{x=1}^n \frac{\partial p(x|\theta)}{\partial \theta^i} D(p(y|x)||p(y|\theta)), \quad (3.66)$$

onde a divergência K-L é função de x e θ , e $p(y|\theta)$ depende de θ pela expressão

$$p(y|\theta) = \sum_{x=1}^n p(y|x)p(x|\theta). \quad (3.67)$$

Definindo as matrizes

$$\mathbf{p}_X \stackrel{\text{def}}{=} \begin{bmatrix} 1 - \sum_{i=1}^n \theta^i \\ \theta^1 \\ \vdots \\ \theta^n \end{bmatrix}, \quad \mathbf{P}_{Y|X} \stackrel{\text{def}}{=} \begin{bmatrix} p(0|0) & p(0|1) & \cdots & p(0|n) \\ p(1|0) & p(1|1) & \cdots & p(1|n) \\ \vdots & \vdots & \ddots & \vdots \\ p(m|0) & p(m|1) & \cdots & p(m|n) \end{bmatrix}, \quad (3.68)$$

então

$$\mathbf{p}_Y = \mathbf{P}_{Y|X} \cdot \mathbf{p}_X, \quad (3.69)$$

e as derivadas parciais $\partial I / \partial \theta^i$ são as componentes da matriz linha $1 \times n$ obtida pelo produto

$$\frac{\partial I}{\partial \theta} = \underbrace{[1 \quad \cdots \quad 1]}_{1 \times (m+1)} \underbrace{\left[\mathbf{P}_{Y|X} \circ (\log \mathbf{P}_{Y|X} - \log \mathbf{p}_Y \cdot [1 \quad \cdots \quad 1]) \right]}_{(m+1) \times (n+1)} \underbrace{\begin{bmatrix} -1 & \cdots & -1 \\ 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{bmatrix}}_{(n+1) \times n}. \quad (3.70)$$

3.8.2 Método do gradiente natural da informação mútua

Tal como anteriormente, o gradiente natural de $I(X; Y)$ é calculado pela fórmula

$$\tilde{\nabla}I = \theta \circ \nabla I - \theta(\theta \cdot \nabla I). \quad (3.71)$$

A aplicação directa do método do gradiente natural permite obter a distribuição óptima $p(x|\theta)$ que maximiza $I(X; Y)$. Como exemplo, foram considerados dois canais propostos em [23]. O primeiro canal, com $m = n = 4$ (alfabetos de dimensão 5), tem probabilidades de transição dadas por

$$\mathbf{P}_{Y|X} = \begin{bmatrix} 0.23 & 0.02 & 0.44 & 0.18 & 0.41 \\ 0.25 & 0.44 & 0.17 & 0.23 & 0.48 \\ 0.13 & 0.31 & 0.18 & 0.33 & 0.02 \\ 0.34 & 0.12 & 0.03 & 0.18 & 0.03 \\ 0.05 & 0.11 & 0.18 & 0.08 & 0.06 \end{bmatrix}. \quad (3.72)$$

O segundo canal, com $n = 9$ e $m = 4$ (10 símbolos de entrada e 5 símbolos de saída), tem probabilidades de transição dadas por

$$\mathbf{P}_{Y|X} = \begin{bmatrix} 0.16 & 0.05 & 0.02 & 0.03 & 0.07 & 0.23 & 0.05 & 0.05 & 0.11 & 0.04 \\ 0.33 & 0.14 & 0.24 & 0.23 & 0.18 & 0.04 & 0.19 & 0.13 & 0.21 & 0.06 \\ 0.03 & 0.37 & 0.15 & 0.17 & 0.39 & 0.17 & 0.22 & 0.20 & 0.04 & 0.26 \\ 0.17 & 0.32 & 0.33 & 0.50 & 0.07 & 0.08 & 0.42 & 0.36 & 0.40 & 0.40 \\ 0.31 & 0.12 & 0.26 & 0.07 & 0.29 & 0.48 & 0.12 & 0.26 & 0.24 & 0.24 \end{bmatrix}. \quad (3.73)$$

Para o primeiro canal, as figuras 3.8 e 3.9 mostram a evolução da estimativa da capacidade do canal C , a norma $\|\tilde{\nabla}I\|_g$ calculada como em (3.44), e os parâmetros θ^i que caracterizam $p(x|\theta)$. A optimização foi efectuada com a versão discreta do método do gradiente natural usando um passo fixo $\eta = 1$, tendo sido obtida uma capacidade $C \approx 0.31124213$ bits/s. Para o segundo canal foi usado o mesmo procedimento tendo sido obtida uma capacidade $C \approx 0.36217799$ bits/s. As figuras 3.10 e 3.11 mostram a evolução tomada pelo algoritmo neste caso.

Em ambos os canais, as distribuições óptimas $p(x)$ têm probabilidades nulas ou próximas de zero em um ou mais símbolos x . Este facto não introduz problemas adicionais no algoritmo, uma vez que o gradiente natural garante intrinsecamente as restrições nas probabilidades.

O número de iterações necessárias para a obtenção da capacidade do canal (com a mesma precisão numérica) é, nos exemplos estudados muito próximo ao obtido com os métodos de Arimoto e Blahut em [23].

O algoritmo, aplicado à informação mútua, não goza da propriedade *quasi-Newton*, como acontece com a divergência de Kullback-Leibler. Tal deve-se ao facto de, na função

$$I(X; Y) = D(p(x, y) \| p(x)p(y)), \quad (3.74)$$

ambos os argumentos da divergência K-L dependem dos parâmetros a otimizar e, consequentemente, a matriz Hessiana não coincide com a matriz de informação de Fisher.

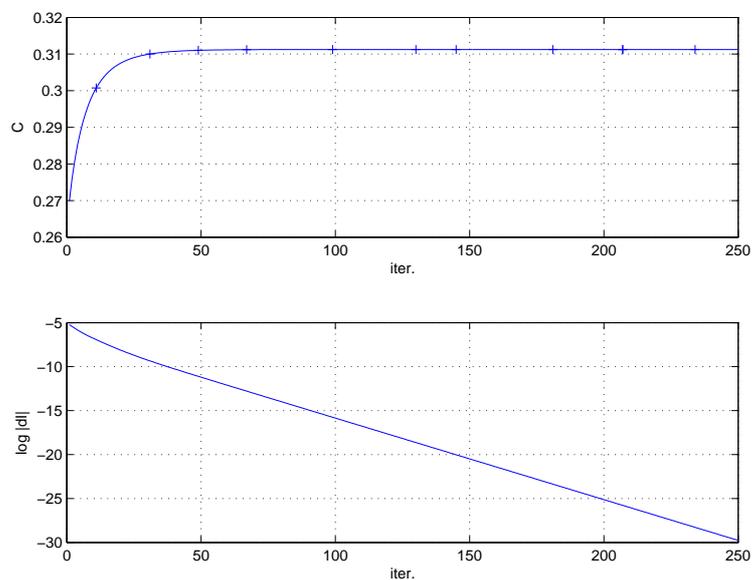


Figura 3.8: Evolução da estimativa da capacidade do canal C e do diferencial $dI(\tilde{\nabla}I)$ em escala logarítmica. As marcas no gráfico de cima sinalizam os momentos em que foram obtidos mais dígitos significativos na representação decimal da capacidade do canal.

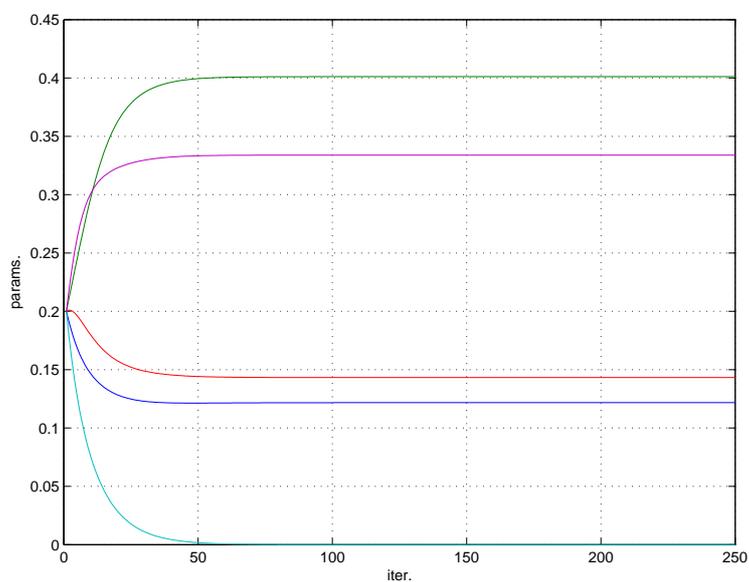


Figura 3.9: Evolução dos parâmetros θ^i . Os seus valores finais determinam a distribuição $p(x)$ que maximiza $I(X; Y)$.

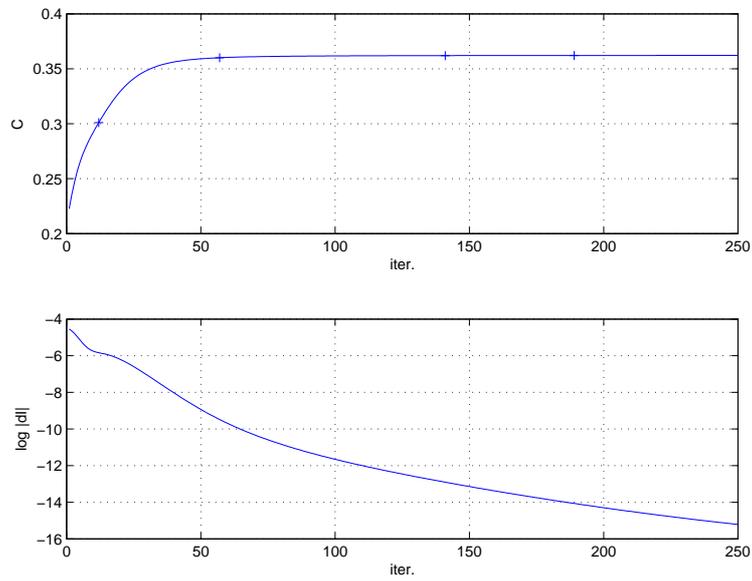


Figura 3.10: Evolução da estimativa da capacidade do canal C e do diferencial $dI(\tilde{\nabla}I)$ em escala logarítmica. As marcas no gráfico de cima sinalizam os momentos em que foram obtidos mais dígitos significativos na representação decimal da capacidade do canal.

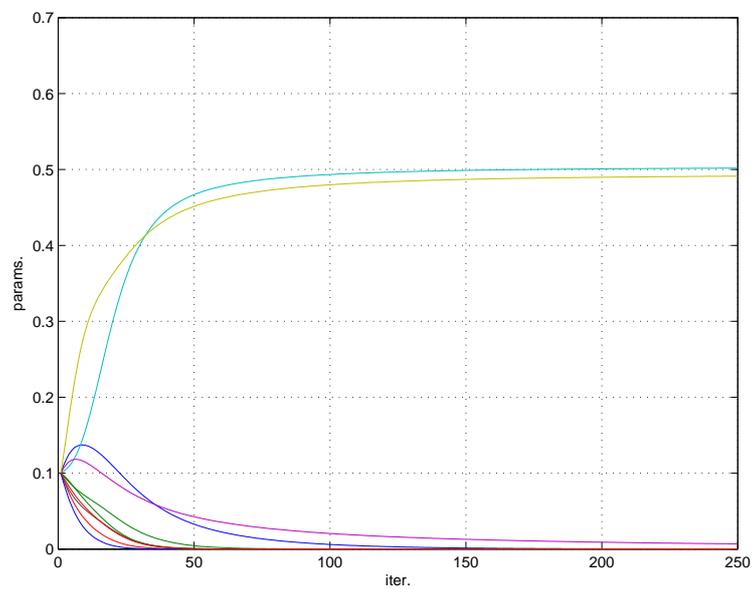


Figura 3.11: Evolução dos parâmetros θ^i . Os seus valores finais determinam a distribuição $p(x)$ que maximiza $I(X; Y)$.

Capítulo 4

Controlo probabilístico

4.1 Introdução

Tradicionalmente, os modelos de sistemas de controlo em cadeia fechada subdividem o sistema global em dois blocos interligados. Estes correspondem ao modelo do sistema a controlar e ao respectivo controlador. Cada bloco transforma os sinais presentes nas entradas em sinais que são colocados nas saídas. A transformação entrada-saída pode eventualmente depender de um estado guardado numa memória interna de cada bloco, e que é actualizado em cada instante.

Considerando um sistema em tempo discreto, as actualizações do estado e da saída podem ser descritas por duas equações como as seguintes,

$$x(t+1) = f(x(t), u(t), t) \quad (4.1)$$

$$y(t) = g(x(t), u(t), t) \quad (4.2)$$

designadas por equação de estado e saída, respectivamente, e onde f e g caracterizam o comportamento dinâmica do sistema.

Um sistema deste tipo diz-se determinístico uma vez que, dados t , $x(t)$ e $u(t)$, as variáveis $x(t+1)$ e $y(t)$ são determinadas de forma exacta. Havendo incerteza nestas variáveis, o sistema diz-se estocástico e as variáveis $x(t+1)$ e $y(t)$ são variáveis aleatórias descritas pelas respectivas distribuições de probabilidade. Num sistema estocástico toda a evolução futura do estado e da saída é descrita por sinais aleatórios.

Um exemplo de um sistema estocástico é um sistema linear com perturbações no estado e ruído nas observações,

$$x(t+1) = Ax(t) + Bu(t) + w(t) \quad (4.3)$$

$$y(t) = Cx(t) + Du(t) + v(t), \quad (4.4)$$

em que (A, B, C, D) são matrizes de dimensões apropriadas e $w(t)$ e $v(t)$ são sucessões de v.a. Gaussianas independentes. Neste sistema, os sinais $x(t)$ e $y(t)$ são aleatórios mesmo quando o sinal $u(t)$ e o estado inicial $x(t_0)$ são determinísticos.

A formulação anterior é frequente na modulação de processos devendo grande parte da sua popularidade à existência de ferramentas teóricas potentes, tais como o filtro Kalman e o controlo LQG.

Neste capítulo, é seguida uma abordagem em que todas as variáveis – entrada, saída e estado – são variáveis aleatórias descritas pelas respectivas distribuições de probabilidade.

Para permitir o maior grau de generalidade possível, não é fixada uma classe particular de sistemas (como os lineares Gaussianos).

Os modelos são descritos por distribuições de probabilidade condicionada genéricas. Como exemplo, a equação de estado do sistema linear Gaussiano (4.3) pode ser escrito nesta formulação probabilística como

$$p(x_{t+1}|x_t, u_t) = \frac{1}{\sqrt{(2\pi)^n |\Sigma|}} \exp\left(-\frac{1}{2}(x_{t+1} - Ax_t - Bu_t)^T \Sigma^{-1}(x_{t+1} - Ax_t - Bu_t)\right), \quad (4.5)$$

onde $n = \dim(x_t)$ e $\Sigma = E[ww^T]$ é a matriz de covariância da perturbação w .

As secções seguintes desenvolvem uma abordagem puramente probabilística para o problema do controlo. A secção 4.2 introduz os modelos probabilísticos e mostra de que forma são determinadas as trajectórias probabilísticas. São apresentados exemplos ilustrativos para variáveis aleatórias contínuas e discretas. A secção 4.3 formaliza o problema de controlo probabilístico, apresentando um funcional de custo que mede o desvio da trajectória probabilística relativamente a uma trajectória ideal. A secção 4.4 apresenta a solução para o problema proposto para os casos contínuos e discretos. É proposta a extensão para horizonte temporal infinito. São consideradas as questões numéricas de implementação do caso discreto para problemas com horizontes temporais elevados. A secção 4.5 formaliza e obtém a solução para o controlo probabilístico em famílias paramétricas de distribuições de probabilidade. A secção 4.6 consiste na aplicação da solução encontrada em 4.5 para o controlo linear Gaussiano com matrizes de covariância arbitrárias. A secção 4.7 dá uma interpretação geométrica da solução paramétrica, à luz da geometria da informação. A secção 4.8 considera famílias mistura de controladores em geral, enquanto a secção 4.9 cinge-se ao caso particular de famílias discretas, e mostra como se pode aplicar o método do gradiente natural a esta família. Finalmente, a secção 4.10 apresenta dois exemplos com graus de complexidade diferentes. São considerados dois sistemas: uma cadeia de Markov controlável, de baixa dimensão; e um sistema de dimensão elevada, obtido por conversão A/D de um sistema linear.

4.2 Modelos probabilísticos

Um modelo probabilístico de um sistema dinâmico é um modelo em que a dinâmica é descrita por uma distribuição de probabilidade condicional. Como exemplo, considere-se o sistema linear em tempo discreto

$$x_{t+1} = Ax_t + w_t, \quad (4.6)$$

em que $x_t \in \mathbb{R}^n$, $w_t \in \mathbb{R}^n$ é uma sucessão aleatória iid. Esta equação pode ser modelada probabilisticamente pela distribuição de probabilidade condicional $p(x_{t+1}|x_t)$, onde a distribuição de probabilidade caracteriza a incerteza em x_{t+1} devida ao termo aleatório w_t . Os seguintes exemplos ilustram dois casos particulares:

Processo estocástico (4.6) com ruído gaussiano multivariado $w_t \sim \mathcal{N}(\mathbf{0}, \Sigma)$ é modelado probabilisticamente por

$$p(x_{t+1}|x_t) = \frac{1}{\sqrt{(2\pi)^n |\Sigma|}} \exp\left(-\frac{1}{2}(x_{t+1} - Ax_t)^T \Sigma^{-1}(x_{t+1} - Ax_t)\right). \quad (4.7)$$

Processo determinístico correspondente a (4.6), com ausência de perturbação, é modelado probabilisticamente usando a distribuição delta de Dirac

$$p(x_{t+1}|x_t) = \delta(x_{t+1} - Ax_t). \quad (4.8)$$

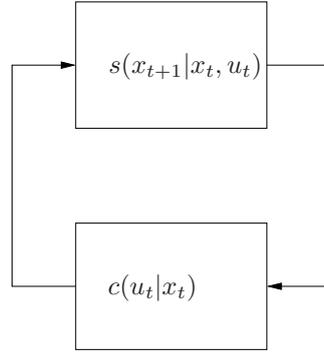


Figura 4.1: Sistema de controle probabilístico.

Os modelos probabilísticos podem ser encadeados em série: se $p(x_t|x_{t-1}, u_t)$ e $p(y_t|y_{t-1}, x_t)$ são dois modelos probabilísticos, então o produto

$$p(y_t, x_t|x_{t-1}, y_{t-1}, u_t) = p(x_t|x_{t-1}, u_t)p(y_t|y_{t-1}, x_t) \quad (4.9)$$

caracteriza a distribuição conjunta das variáveis x_t e y_t . O modelo probabilístico com entrada u e saída y , correspondente ao encadeamento em série, obtém-se calculando a distribuição marginal (equação de Chapman-Kolmogorov)

$$p(y_t|x_{t-1}, y_{t-1}, u_t) = \int_{\mathcal{X}} p(x_t|x_{t-1}, u_t)p(y_t|y_{t-1}, x_t)dx_t. \quad (4.10)$$

Diz-se que um sistema realimentado é um *sistema de controle probabilístico* se tanto o sistema s como o controlador c são modelados probabilisticamente. A figura 4.1 ilustra um sistema de controle deste tipo.

A evolução temporal deste tipo de sistemas é determinada pela regra da cadeia das probabilidades. Se a condição inicial é caracterizada por $p(x_0)$, então temos sucessivamente:

$$\begin{aligned} p(x_0, u_0) &= c(u_0|x_0)p(x_0) \\ p(x_0, x_1, u_0) &= s(x_1|x_0, u_0)c(u_0|x_0)p(x_0) \\ &\vdots \\ p(x_{0:t}, u_{0:t-1}) &= s(x_t|x_{t-1}, u_{t-1})c(u_{t-1}|x_{t-1}) \cdots s(x_1|x_0, u_0)c(u_0|x_0)p(x_0). \end{aligned}$$

Num horizonte temporal arbitrário T , a trajetória do sistema em cadeia fechada é caracterizada pela distribuição de probabilidade conjunta

$$p(x_{0:T}, u_{0:T-1}) = p(x_0) \prod_{t=0}^{T-1} s(x_{t+1}|x_t, u_t)c(u_t|x_t) \quad (4.11)$$

em que $x_{0:T} = (x_0, x_1, \dots, x_T)$ e $u_{0:T-1} = (u_0, u_1, \dots, u_{T-1})$ representam de forma condensada a sequência de sinais x_t e u_t no horizonte temporal de 0 a T .

Até ao momento não foi referido o domínio das variáveis x e u . Estas variáveis aleatórias podem ser discretas ou contínuas. No caso de v.a. contínuas, onde se lê distribuições de probabilidade deve ler-se densidade de probabilidade. Os modelos são, nesse caso, funções densidade de probabilidade condicionada mas as fórmulas têm a mesma apresentação.

Os sinais x_t e u_t podem ainda ser escalares ou vectoriais mantendo-se a validade de todas as expressões. Qualquer que seja a natureza das variáveis, estas serão chamadas variáveis aleatórias.

4.2.1 Exemplos

Controlo determinístico (caso contínuo) Suponha-se que as variáveis aleatórias x e u são contínuas e que o controlo gera um sinal u_t determinístico obtido por $u_t = f(x_t)$. Neste caso, a descrição probabilística do controlador é dada por uma função delta de Dirac

$$c(u_t|x_t) = \delta(u_t - f(x_t)). \quad (4.12)$$

A trajectória do sistema realimentado é, aplicando (4.11),

$$p(x_{0:T}, u_{0:T-1}) = p(x_0) \prod_{t=0}^{T-1} s(x_{t+1}|x_t, u_t) \delta(u_t - f(x_t)). \quad (4.13)$$

Uma vez que u é determinístico, a representação probabilística da trajectória pode ser feita apenas na variável x , sem que haja perda de informação¹. Assim, obtém-se a distribuição marginal

$$\begin{aligned} p(x_{0:T}) &= \int p(x_{0:T}, u_{0:T-1}) du_{0:T-1} \\ &= \int p(x_0) \prod_{t=0}^{T-1} s(x_{t+1}|x_t, u_t) \delta(u_t - f(x_t)) du_{0:T-1} \\ &= p(x_0) \prod_{t=0}^{T-1} \int s(x_{t+1}|x_t, u_t) \delta(u_t - f(x_t)) du_t \\ &= p(x_0) \prod_{t=0}^{T-1} s(x_{t+1}|x_t, f(x_t)). \end{aligned} \quad (4.14)$$

Definindo

$$s^{\text{cl}}(x_{t+1}|x_t) \stackrel{\text{def}}{=} s(x_{t+1}|x_t, f(x_t)), \quad (4.15)$$

então a trajectória pode ser caracterizada probabilisticamente usando apenas o processo de Markov descrito por s^{cl} .

Controlo determinístico (caso discreto) Considere-se que as variáveis aleatórias x e u pertencem a alfabetos finitos \mathcal{X} e \mathcal{U} , respectivamente. Os símbolos, elementos dos conjuntos \mathcal{X} e \mathcal{U} , podem ser indexados pelos inteiros $\{0, 1, 2, \dots, n\}$ e $\{0, 1, 2, \dots, m\}$, onde $n = \#\mathcal{X} - 1$ e $m = \#\mathcal{U} - 1$, de maneira que os índices possam ser usados como código dos símbolos. Suponhamos que a lei de controlo determinístico é $u_t = f(x_t)$. A sua descrição probabilística correspondente é

$$c(u_t|x_t) = \delta_{u_t}^{f(x_t)}, \quad (4.16)$$

em que δ_i^j é a função delta de Kronecker², e a trajectória do sistema realimentado é

$$p(x_{0:T}, u_{0:T-1}) = p(x_0) \prod_{t=0}^{T-1} s(x_{t+1}|x_t, u_t) \delta_{u_t}^{f(x_t)}. \quad (4.17)$$

¹A informação de uma variável aleatória X é, pela teoria da informação de Shannon, medida pela sua entropia $H(X)$. Se X é uma v.a. e $Y = f(X)$ então $H(X, Y) = H(X)$.

²A função delta de Kronecker δ_i^j é 1 (um) se $i = j$ e zero caso contrário.

Tal como no caso contínuo, é possível escrever a distribuição de probabilidade que caracteriza a trajectória do sistema realimentado, eliminando a componente determinística. Obtém-se

$$\begin{aligned}
p(x_{0:T}) &= \sum_{u_{0:T-1}} p(x_{0:T}, u_{0:T-1}) \\
&= \sum_{u_{0:T-1}} p(x_0) \prod_{t=0}^{T-1} s(x_{t+1}|x_t, u_t) \delta_{u_t}^{f(x_t)} \\
&= p(x_0) \prod_{t=0}^{T-1} \sum_{u_t} s(x_{t+1}|x_t, u_t) \delta_{u_t}^{f(x_t)} \\
&= p(x_0) \prod_{t=0}^{T-1} s(x_{t+1}|x_t, f(x_t)).
\end{aligned} \tag{4.18}$$

Definindo

$$s^{\text{cl}}(x_{t+1}|x_t) \stackrel{\text{def}}{=} s(x_{t+1}|x_t, f(x_t)), \tag{4.19}$$

a trajectória de x é gerada pela cadeia de Markov $s^{\text{cl}}(x_{t+1}|x_t)$.

Sistemas variantes no tempo Tanto o sistema s como o controlador c pode ser variante no tempo, isto é, as probabilidades de transição que caracterizam probabilisticamente o sistema ou controlador podem depender do instante de tempo. Assim, a trajectória dos sinais x e u é dada por

$$p(x_{0:T}, u_{0:T-1} | \tau_{0:T-1}) = p(x_0) \prod_{t=0}^{T-1} s(x_{t+1}|x_t, u_t, \tau_t) c(u_t|x_t, \tau_t), \tag{4.20}$$

em que a sequência $\tau_{0:T-1}$ selecciona o horizonte temporal no qual se efectua a evolução do sistema.

Sistema de controlo paramétrico Um sistema de controlo probabilístico em que o controlador pertence a uma classe paramétrica de modelos $c(u_t|x_t, \theta_t)$ com parametro $\theta_t \in \Theta$, gera trajectórias do estado parametrizadas pela sequência $\theta_{0:T-1} = (\theta_0, \theta_1, \dots, \theta_{T-1})$.

$$\begin{aligned}
p(x_{0:T} | \theta_{0:T-1}) &= \int_{\mathcal{U}^T} p(x_{0:T}, u_{0:T-1} | \theta_{0:T-1}) du_{0:T-1} \\
&= \int_{\mathcal{U}^T} p(x_0) \prod_{t=0}^{T-1} s(x_{t+1}|x_t, u_t) c(u_t|x_t, \theta_t) du_{0:T-1} \\
&= p(x_0) \prod_{t=0}^{T-1} \underbrace{\int_{\mathcal{U}} s(x_{t+1}|x_t, u_t) c(u_t|x_t, \theta_t) du_{0:T-1}}_{p(x_{t+1}|x_t, \theta_t)}
\end{aligned} \tag{4.21}$$

4.3 Formulação do problema de controlo

Num sistema de controlo probabilístico, como o apresentado na secção anterior, o controlador deve ser seleccionado de forma a que, em malha fechada, as trajectórias das variáveis x_t e u_t estejam próximas de uma trajectória desejada.

Este objectivo é análogo à minimização do funcional de custo quadrático para sistemas determinísticos

$$J = \frac{1}{2} \sum_{t=0}^T \|x_t\|_Q^2 + \|u_t\|_R^2, \quad (4.22)$$

onde as trajectórias desejadas são $x_t = 0$ e $u_t = 0$ num horizonte temporal $t = 0, \dots, T$. Neste caso, são penalizadas todas as trajectórias que se afastem da trajectória ideal no sentido Euclidiano e com métricas definidas pelas matrizes de penalização Q e R .

No caso probabilístico, a utilização do funcional de custo (4.22) resulta numa variável aleatória que não pode ser minimizada directamente. É prática comum, nestes casos, efectuar a optimização sobre o seu valor esperado. Neste trabalho é seguido um caminho diferente, introduzido por Kárný [21], e que pode ser interpretado do ponto de vista da teoria da informação como se descreve em seguida.

Consideram-se dois sistemas “idealizados” $S(x_{t+1}|x_t, u_t)$ e $C(u_t|x_t)$ em cadeia fechada. A informação gerada por estes sistemas num horizonte temporal T é dada pela entropia conjunta

$$h(x_{0:T}, u_{0:T-1}) = \int q(x_{0:T}, u_{0:T-1}) \log \frac{1}{q(x_{0:T}, u_{0:T-1})} dx_{0:T} du_{0:T-1}, \quad (4.23)$$

onde

$$q(x_{0:T}, u_{0:T-1}) \stackrel{\text{def}}{=} q(x_0) \prod_{t=0}^{T-1} S(x_{t+1}|x_t, u_t) C(u_t|x_t). \quad (4.24)$$

Os sinais $x_{0:T}$ e $u_{0:T-1}$ gerados podem ser armazenados, em média, num espaço não inferior ao valor da entropia [12]. As propriedades estatísticas são usadas para otimizar o espaço ocupado, por exemplo no desenho de um código de dicionário. No caso dos modelos se desviarem dos originais para $s(x_{t+1}|x_t, u_t)$ e $c(u_t|x_t)$, gerando uma trajectória em malha fechada com distribuição $p(x_{0:T}, u_{0:T-1})$, o código anterior irá produzir, em média, palavras de código subóptimas, uma vez que o código original não está adaptado às novas propriedades estatísticas dos sinais. Mais precisamente, obtém-se uma penalização no tamanho dos dados igual a

$$\int p(x_{0:T}, u_{0:T-1}) \log \frac{1}{q(x_{0:T}, u_{0:T-1})} dx_{0:T} du_{0:T-1} - h(p) = D(p||q), \quad (4.25)$$

onde $D(p||q)$ é a divergência de Kullback-Leibler. A figura 4.2 ilustra de forma intuitiva este resultado. Esta penalização pode ser usada como funcional de custo:

$$J_{0:T} \stackrel{\text{def}}{=} D(p(x_{0:T}, u_{0:T-1})||q(x_{0:T}, u_{0:T-1})). \quad (4.26)$$

O problema de controlo consiste em determinar a distribuição $c(u|x)$ que minimiza este custo.

A interpretação anterior, que justifica a divergência de Kullback-Leibler partindo de um problema de compressão de dados, é apenas uma das muitas possíveis. Na realidade poderiam ser consideradas outras medidas de distância ou divergência entre as duas distribuições de probabilidade real e ideal p e q . A divergência de Kullback-Leibler tem no entanto uma propriedade que, de um ponto de vista pragmático, a justifica sobre as outras distâncias e divergências. Esta propriedade é a regra da cadeia das probabilidades, que tem uma expressão análoga na divergência de K-L, onde a divergência total é escrita como a soma de várias divergências de K-L parciais. Esta propriedade é explorada nas secções seguintes para obter um algoritmo recursivo (programação dinâmica). As restantes divergências, não possuindo esta propriedade, têm uma aplicação prática mais complexa.

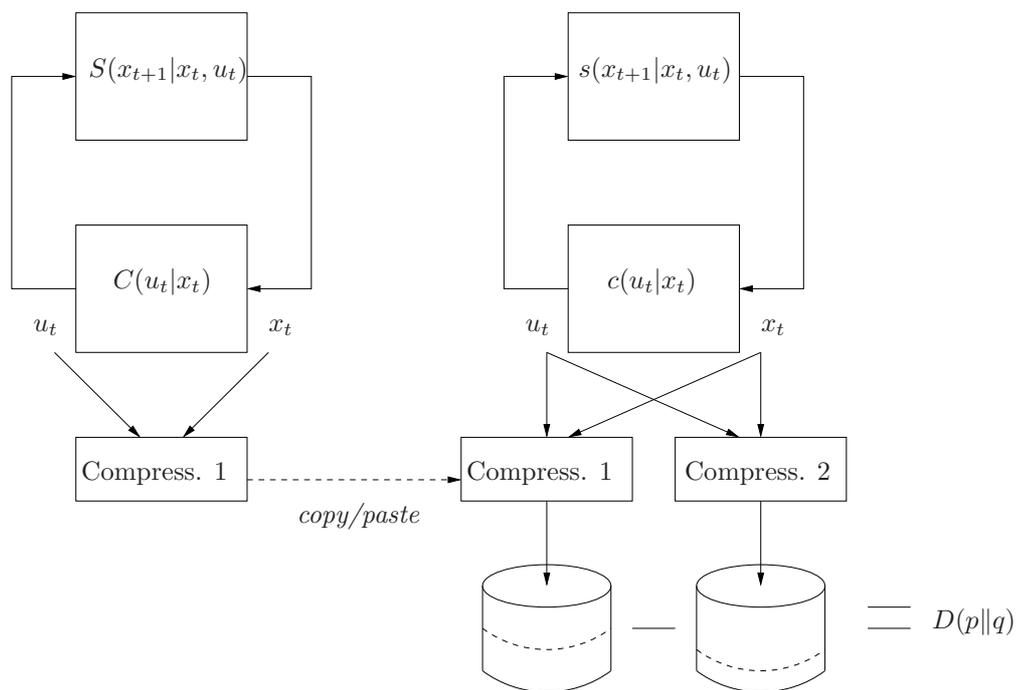


Figura 4.2: Comparação entre a quantidade de dados gerada por um compressor adaptado e outro não adaptado à cadeia de controlo.

4.4 Solução óptima do controlo probabilístico

A formulação apresentada na secção anterior permite considerar problemas com variáveis discretas, assim como problemas com variáveis contínuas. No caso de variáveis aleatórias discretas, o sistema $s(x_{t+1}|x_t, u_t)$ corresponde a uma cadeia de Markov controlável pela variável u_t , enquanto o controlo $c(u_t|x_t)$ corresponde a uma matriz de transição. O objectivo é, neste caso, determinar esta matriz de transição. No caso de variáveis aleatórias contínuas, as distribuições $s(x_{t+1}|x_t, u_t)$ e $c(u_t|x_t)$ são funções densidade de probabilidade condicionadas que modelam dois processos de Markov.

Considera-se que, em cada instante de tempo, as variáveis aleatórias x_t e u_t têm distribuições de probabilidade ideais $S(x_{t+1}|x_t, u_t, t)$ e $C(u_t|x_t, t)$, respectivamente. Com estas distribuições, e partindo da distribuição inicial $p(x_0)$, a trajectória ideal é

$$q(x_{0:T}, u_{0:T-1}) = p(x_0) \prod_{t=0}^{T-1} S(x_{t+1}|x_t, u_t, t) C(u_t|x_t, t). \quad (4.27)$$

Esta é a trajectória ideal considerada nas secções 4.4.1 e 4.4.2 seguintes, e que apresentam as soluções óptimas devidas a Kárný [21, 27] para os casos discreto e contínuo.

4.4.1 Caso discreto

Sejam x_t e u_t variáveis aleatórias discretas, cuja dinâmica é determinada pelos sistemas probabilísticos $s(x_{t+1}|x_t, u_t)$ e $c(u_t, x_t)$, e seja $q(x_{0:T}, u_{0:T-1})$ uma trajectória ideal escrita como em (4.27). Pretende-se minimizar o funcional de custo $J = D(p(x_{0:T}, u_{0:T-1})||q(x_{0:T}, u_{0:T-1}))$ seleccionando a sucessão de controladores óptimos $c(u_t|x_t, t)$, com $t = 0, \dots, T-1$.

A solução óptima, encontrada em [27], consiste em definir

$$\gamma(x_T, T) = 1, \quad (4.28)$$

e iterar recursivamente para trás no tempo o controlador

$$c(u_t|x_t, t) = \frac{1}{\gamma(x_t, t)} C(u_t|x_t, t) \exp(-a(x_t, u_t, t) - b(x_t, u_t, t)), \quad (4.29)$$

em que os termos a , b e γ são dados por

$$\begin{aligned} a(x_t, u_t, t) &= D(s(x_{t+1}|x_t, u_t, t)||S(x_{t+1}|x_t, u_t, t)), \\ b(x_t, u_t, t) &= E_{s(x_{t+1}|x_t, u_t, t)}[-\log \gamma(x_{t+1}, t+1)], \\ \gamma(x_t, t) &= \sum_{u_t \in \mathcal{U}} C(u_t|x_t, t) \exp(-a(x_t, u_t, t) - b(x_t, u_t, t)). \end{aligned} \quad (4.30)$$

Estes termos, calculados em cada instante da recorrência, têm a seguinte interpretação:

- O termo a mede a divergência entre as distribuições de probabilidade reais e ideais de x_{t+1} , em função de u_t e do estado actual x_t . Note-se que esta divergência afecta apenas o instante $t+1$ da variável x , não tendo em consideração todo o horizonte temporal;
- O termo b mede o *cost to go* esperado no restante horizonte temporal. Na primeira iteração, este termo é nulo;
- O termo γ é um factor de normalização que garante $\sum_{u_t} c(u_t|x_t, t) = 1$.

4.4.2 Caso contínuo

A solução do caso contínuo, obtida em [21], é semelhante à solução do caso discreto. Esta consiste em definir

$$\gamma(x_T, T) = 1, \quad (4.31)$$

e, em seguida, iterar recursivamente para trás no tempo o controlador

$$c(u_t|x_t, t) = \frac{1}{\gamma(x_t, t)} C(u_t|x_t, t) \exp(-a(x_t, u_t, t) - b(x_t, u_t, t)), \quad (4.32)$$

em que os termos a , b e γ são dados por

$$\begin{aligned} a(x_t, u_t, t) &= D(s(x_{t+1}|x_t, u_t, t) \| S(x_{t+1}|x_t, u_t, t)), \\ b(x_t, u_t, t) &= E_{s(x_{t+1}|x_t, u_t, t)}[-\log \gamma(x_{t+1}, t+1)], \\ \gamma(x_t, t) &= \int_{\mathcal{U}} C(u_t|x_t, t) \exp(-a(x_t, u_t, t) - b(x_t, u_t, t)) du_t. \end{aligned} \quad (4.33)$$

A interpretação dos termos é a mesma do caso discreto.

Ao contrário do caso discreto (finito), a solução no caso contínuo está dependente da possibilidade de uma solução analítica para os três termos em (4.33), termos esses que envolvem, nos três casos, cálculos integrais. Em geral, estes não têm solução analítica, pelo que a fórmula (4.32), aparentemente explícita, não pode na realidade ser implementada a não ser em casos muito particulares. Um desses casos é quando todos os modelos envolvidos são linear-gaussianos. Estabelece-se, nesse caso, uma relação entre o controlo probabilístico e o controlo linear quadrático clássico (controlo LQ).

Relação com o controlo linear quadrático

A relação entre a formulação probabilística e o controlo LQ, explorada em [21], mostra que as matrizes de covariância dos modelos probabilísticos lineares gaussianos, desempenham um papel equivalente às matrizes de penalização do estado e do controlo na formulação clássica. A formulação é a seguinte:

Considere-se o sistema linear

$$x_{t+1} = Ax_t + Bu_t + w_t, \quad w_t \sim \mathcal{N}(\mathbf{0}, Q), \quad (4.34)$$

onde w_t é uma perturbação no estado, gaussiana multivariada.

Alternativamente, a equação de estado (4.34) pode ser escrita na forma probabilística pela densidade de probabilidade condicionada

$$s(x_{t+1}|x_t, u_t) = \mathcal{N}_{x_{t+1}}(Ax_t + Bu_t, Q). \quad (4.35)$$

Suponha-se que as densidades ideais do estado e do controlo são distribuições gaussianas de média nula e covariância Q e R ,

$$S(x_{t+1}) = \mathcal{N}_{x_{t+1}}(\mathbf{0}, Q), \quad (4.36)$$

$$C(u_t) = \mathcal{N}_{u_t}(\mathbf{0}, R). \quad (4.37)$$

A aplicação das fórmulas (4.32) e (4.33) permite obter o controlo óptimo $c(u_t|x_t)$ para estes modelos. Obtém-se o controlo óptimo dado pela função densidade de probabilidade gaussiana

$$c(u_t|x_t) = \mathcal{N}_{u_t}(-L_t x_t, R_t), \quad (4.38)$$

em que os ganhos L_t e a matriz de covariância R_t resultam da recorrência para trás no tempo

$$\Sigma_T = \mathbf{0}, \quad (4.39)$$

$$\Sigma_t = A^T [\Sigma_{t+1} - \Sigma_{t+1} B (R^{-1} + B^T \Sigma_{t+1} B)^{-1} B^T \Sigma_{t+1}] A + Q^{-1}, \quad (4.40)$$

$$R_t = (R^{-1} + B^T \Sigma_t B)^{-1}, \quad (4.41)$$

$$L_t = R_t B^T \Sigma_t A. \quad (4.42)$$

Estes ganhos são precisamente os obtidos na solução clássica do regulador linear quadrático (ver *e.g.* [22]). A dedução destes resultados pode ser consultada em [21]. A tabela 4.1

| Abordagem: | Probabilística | Clássica |
|------------------|---|---|
| Modelo: | $s(x_{t+1} x_t, u_t) = \mathcal{N}(Ax_t + Bu_t, Q)$ | $x_{t+1} = Ax_t + Bu_t$ |
| Custo: | $D(p(x_{0:T}, u_{0:T-1}) q(x_{0:T}, u_{0:T-1}))$ | $\sum_{t=0}^{T-1} (x_t^T Q^{-1} x_t + u_t^T R^{-1} u_t)$ |
| Lei de controlo: | $c(u_t x_t, t) = \mathcal{N}_{u_t}(-L_t x_t, R_t)$ | $u_t = -L_t x_t$ |
| Recorrência: | | $\Sigma_T = \mathbf{0}$ |
| | | $\Sigma_t = A^T [\Sigma_{t+1} - \Sigma_{t+1} B (R^{-1} + B^T \Sigma_{t+1} B)^{-1} B^T \Sigma_{t+1}] A + Q^{-1}$ |
| | | $R_t = (R^{-1} + B^T \Sigma_t B)^{-1}$ |
| | | $L_t = R_t B^T \Sigma_t A$ |

Tabela 4.1: Comparação das abordagens: probabilística linear-gaussiana *vs* clássica com controlo linear e custo quadrático.

mostra ambas as formulações e as respectivas soluções. Note-se que, embora os ganhos coincidam, uma formulação é estocástica enquanto a outra é determinística. A coincidência dos ganhos permite efectuar uma interpretação interessante entre as duas abordagens, nomeadamente a correspondência directa entre as matrizes de covariância ideais Q e R do controlo probabilístico e as matrizes de penalização Q^{-1} e R^{-1} do regulador linear quadrático.

4.4.3 Extensão para horizonte temporal infinito

As soluções apresentadas anteriormente são válidas para horizontes temporais T finitos. Se, no funcional de custo anterior, se estender o horizonte para infinito $T \rightarrow \infty$, surge uma dificuldade que resulta do custo J_∞ em geral divergir. Nesta secção, o funcional de custo J_t é modificado de forma a garantir convergência para uma certa classe de problemas.

Considere-se a sucessão J_t que consiste na sequência dos custos óptimos, obtidos no problema de optimização em horizontes temporais sucessivamente maiores. O custo adicional incorrido pelo aumento do horizonte em uma unidade de tempo é dado por

$$\begin{aligned} J_{t+1} - J_t &= D(p_{0:t+1} || q_{0:t+1}) - D(p_{0:t} || q_{0:t}) \\ &= D(s_{t+1} c_t p_{0:t} || S_{t+1} C_t q_{0:t}) - D(p_{0:t} || q_{0:t}) \\ &= E_{p_{0:t}} \left[E_{c_t} \left[D(s_{t+1} || S_{t+1}) \right] + D(c_t || C_t) \right] \\ &= E_{p(x_t)} \left[E_{c_t} \left[D(s_{t+1} || S_{t+1}) \right] + D(c_t || C_t) \right]. \end{aligned} \quad (4.43)$$

Assim, é suficiente que exista alguma subsucessão para a qual $D(s_{t+1}||S_{t+1})$ ou $D(c_t||C_t)$ sejam limitadas inferiormente por uma constante positiva, o que acontece normalmente, para que a sucessão J_t divirja para infinito.

Admitindo que (4.43) é majorada por uma constante b e minorada por a , então o funcional modificado

$$\tilde{J}_t \stackrel{\text{def}}{=} \frac{1}{t} J_t \quad (4.44)$$

forma uma sucessão limitada uma vez que

$$\begin{aligned} \frac{1}{t} J_t &= \frac{1}{t} \sum_{\tau=0}^{t-1} J_{\tau+1} - J_\tau \\ &\leq \frac{1}{t} \sum_{\tau=0}^{t-1} b \\ &= b, \end{aligned} \quad (4.45)$$

e da mesma forma

$$\begin{aligned} \frac{1}{t} J_t &= \frac{1}{t} \sum_{\tau=0}^{t-1} J_{\tau+1} - J_\tau \\ &\geq \frac{1}{t} \sum_{\tau=0}^{t-1} a \\ &= a, \end{aligned} \quad (4.46)$$

mas nada se pode dizer acerca da convergência de \tilde{J}_t em geral. Existem, no entanto, situações particulares para as quais se garante convergência.

Proposição 8. *Se todos os modelos $s(x_{t+1}|x_t, u_t)$, $c(u_t|x_t)$, $S(x_{t+1})$ e $C(u_t)$ são invariantes no tempo e a cadeia de Markov induzida*

$$m(x_{t+1}|x_t) \stackrel{\text{def}}{=} \sum_{u_t} s(x_{t+1}|x_t, u_t) c(u_t|x_t) \quad (4.47)$$

é irreduzível³ e aperiódica⁴, então $\tilde{J}_t \rightarrow \tilde{J}_\infty$ quando $t \rightarrow \infty$.

Demonstração. Se $m(x_{t+1}|x_t)$ é irreduzível e aperiódica, então existe uma única distribuição estacionária μ e a sucessão de distribuições $p(x_t)$ converge para μ (ver [12, 15]).

Se os sistemas s , c , S e C são invariantes no tempo, então a função

$$\alpha(x_t) \stackrel{\text{def}}{=} E_{c_t} [D(s_{t+1}||S_{t+1}) + D(c_t||C_t)] \quad (4.48)$$

é constante no tempo e consequentemente $J_{t+1} - J_t = E_{p(x_t)}[\alpha(x_t)] \rightarrow E_\mu[\alpha(x_t)]$ quando $t \rightarrow \infty$. Finalmente, definindo $\tilde{J}_\infty \stackrel{\text{def}}{=} E_\mu[\alpha(x_t)]$, então

$$\tilde{J}_t = \frac{1}{t} \sum_{\tau=0}^{t-1} (J_{\tau+1} - J_\tau) \rightarrow \tilde{J}_\infty, \quad (4.49)$$

³Uma cadeia de Markov é irreduzível se para cada par de estados i, j a probabilidade de transitar entre esses estados é não nula para algum intervalo de tempo suficientemente grande n .

⁴Um estado i é periódico com período τ se o regresso a esse estado está condicionado a ocorrer em intervalos de tempo múltiplos de τ . Se o período é $\tau = 1$ para todos os estados, então a cadeia de Markov é aperiódica.

pelo lema da soma de Cesáro⁵. Assim, o funcional modificado converge: $\tilde{J}_t \rightarrow \tilde{J}_\infty$. \square

Note-se que a proposição anterior parte do princípio que c_t é invariante no tempo, o que não acontece para horizontes finitos. No entanto, para horizontes temporais sucessivamente maiores, o controlo c_t converge eventualmente para uma versão estacionária c_∞ . A análise detalhada da (eventual) convergência é ainda um problema em aberto.

Relação com o ritmo de entropia

O ritmo de entropia de um processo estocástico Y_t é definido por

$$H(\mathcal{Y}) = \lim_{t \rightarrow \infty} \frac{1}{t} H(Y_1, Y_2, \dots, Y_t) \quad (4.50)$$

Prova-se que, para um processo estocástico estacionário, o ritmo de entropia é igual à entropia condicional (ver [12] págs. 63–66)

$$H(\mathcal{Y}) = \lim_{t \rightarrow \infty} H(Y_t | Y_{t-1}, Y_{t-2}, \dots, Y_1). \quad (4.51)$$

Se Y é um processo de Markov estacionário, a equação (4.51) simplifica-se para

$$H(\mathcal{Y}) = H(Y_t | Y_{t-1}). \quad (4.52)$$

O funcional de custo usado na formulação do controlo probabilístico é semelhante à definição de ritmo de entropia, em que a entropia H é substituída pela entropia relativa D . Mostra-se em seguida que a igualdade (4.52) tem uma expressão análoga no controlo probabilístico.

Usando a separação temporal da entropia relativa $D(p_{0:t} \| q_{0:t})$ pode escrever-se

$$\lim_{t \rightarrow \infty} \frac{D(p_{0:t} \| q_{0:t})}{t} = \lim_{t \rightarrow \infty} \frac{1}{t} \sum_{\tau=1}^t E_{p(x_\tau)} \left[D(p(x_{\tau+1}, u_\tau | x_\tau) \| q(x_{\tau+1}, u_\tau | x_\tau)) \right], \quad (4.53)$$

que em regime estacionário resulta em

$$\lim_{t \rightarrow \infty} \frac{D(p_{0:t} \| q_{0:t})}{t} = E_{p(x_t)} \left[D(p(x_{t+1}, u_t | x_t) \| q(x_{t+1}, u_t | x_t)) \right]. \quad (4.54)$$

Isto é, num sistema de controlo probabilístico em regime estacionário, o ritmo de entropia relativa é igual à entropia relativa condicional (4.54), onde

$$p(x_{t+1}, u_t | x_t) = s(x_{t+1} | x_t, u_t) c_\infty(u_t | x_t) \quad (4.55)$$

e

$$q(x_{t+1}, u_t | x_t) = S(x_{t+1} | x_t, u_t) C(u_t | x_t). \quad (4.56)$$

4.4.4 Considerações de implementação da solução discreta

A utilização directa da solução óptima (4.29)-(4.30) em vírgula flutuante poderá não ter os resultados esperados, devido à forma como o *standard IEEE754*, implementado na generalidade do *hardware* actualmente disponível, define o formato e as operações de vírgula flutuante suportadas.

Admitindo que o algoritmo é implementado em precisão dupla numa máquina que segue este *standard*, surgem problemas devidos às situações seguintes:

⁵A soma de Cesáro afirma que se a sucessão $a_n \rightarrow a$ então $b_n = \frac{1}{n} \sum_{m=1}^n a_m \rightarrow a$.

1. Havendo probabilidades nulas no modelo $s(x_{t+1}|x_t, u_t)$, o cálculo da divergência $D(s||S)$ resulta em NaN (Not-A-Number). Este problema resulta do produto

$$0 \cdot \log 0 = 0 \cdot (-\infty) = \text{NaN}, \quad (4.57)$$

efectuado durante o cálculo, propagar-se até ao resultado final. Uma solução consiste em substituir imediatamente todas as ocorrências de NaN por zeros, e que corresponde à usual definição por continuidade:

$$0 \cdot \log 0 \stackrel{\text{def}}{=} 0. \quad (4.58)$$

2. O termo $E[-\log \gamma]$ cresce em amplitude à medida que as recorrência avança (para trás no tempo). Este termo está relacionado com o *cost to go*, pelo que, em horizontes temporais longos, tende a assumir valores elevados. A sua presença no argumento da função exponencial tem como efeito que, ao fim de algumas recorrências, a exponencial sofre *underflow*, obtendo-se em seguida uma divisão por zero devida à normalização. Para solucionar este problema, é efectuada uma normalização do argumento da exponencial antes da sua avaliação, evitando-se o *underflow*. A normalização é cancelada no final, recuperando-se assim o resultado correcto. A tabela 4.2 apresenta, na coluna da direita, a sequência de operações usadas para o efeito.

| Solução original | Solução modificada |
|---|--|
| $-\log \gamma_t(x_t) = \sum_{u_t} C_t \exp \left(-D(s_{t+1} S_{t+1}) + \right. \\ \left. -E_{s_{t+1}}[-\log \gamma_{t+1}(x_{t+1})] \right)$ | $\log \hat{c}_t(u_t x_t) = \log C_t - D(s_{t+1} S_{t+1}) + \\ -E_{s_{t+1}}[-\log \gamma_{t+1}(x_{t+1})]$ $\mu_t(x_t) = \max_{u_t} \log \hat{c}_t(u_t x_t)$ $\bar{c}_t(u_t x_t) = \exp(\log \hat{c}_t - \mu_t)$ $\bar{\gamma}_t(x_t) = \sum_{u_t} \bar{c}_t(u_t x_t)$ $-\log \gamma_t(x_t) = -\log \bar{\gamma}_t(x_t) - \mu_t(x_t)$ |
| $c_t(u_t x_t) = \frac{1}{\gamma_t(x_t)} C_t \exp \left(-D(s_{t+1} S_{t+1}) + \right. \\ \left. -E_{s_{t+1}}[-\log \gamma_{t+1}(x_{t+1})] \right)$ | $c_t(u_t x_t) = \bar{c}_t(u_t x_t) / \bar{\gamma}_t(x_t)$ |

Tabela 4.2: Versões original e modificada da solução óptima discreta.

4.5 Solução paramétrica do controlo probabilístico

As soluções explícitas apresentadas nas secções 4.4.1 e 4.4.2, embora elegantes do ponto de vista teórico, podem apresentar dificuldades práticas na sua implementação:

1. A versão contínua não tem, em geral, solução analítica pelo que a sua aplicabilidade se resume a casos particulares como o LQ-Gaussiano;

2. A versão discreta, por outro lado, tem solução computável desde que se tenham em consideração as questões numéricas que se podem levantar na sua implementação (ver secção 4.4.4). A solução explícita não permite que sejam incorporadas restrições na família de controladores probabilísticos (*e.g.*, limitar a gama de valores que certas probabilidades podem tomar).

Nesta secção é proposta uma solução para estes problemas, que garante à partida ser fazível, e que permite a incorporação de restrições adicionais.

Considera-se que os controladores candidatos pertencem a uma família paramétrica de controladores \mathcal{K} , fixa, e parametrizada por n parâmetros $(\theta_t^1, \theta_t^2, \dots, \theta_t^n)$ em cada instante de tempo t . A restrição do controlo a uma família paramétrica permite que sejam apenas consideradas leis de controlo $c(u|x, \theta)$ para os quais a sua implementação em tempo real é viável. A carga computacional é movida para a resolução da optimização, que é geralmente efectuada *offline*.

A minimização do funcional de custo (paramétrico)

$$J_{0:T} = D(p(x_{0:T}, u_{0:T-1} | \theta_{0:T-1}) \| q(x_{0:T}, u_{0:T-1})) \quad (4.59)$$

é agora efectuada relativamente aos nT parâmetros θ_t^i ao longo de todo o horizonte de tempo. O mínimo de J satisfaz a condição de estacionariedade

$$dJ\left(\frac{\partial}{\partial \theta_t^i}\right) = 0, \quad i = 1, \dots, n, \quad t = 0, \dots, T-1 \quad (4.60)$$

onde dJ é o diferencial de J definido por

$$dJ\left(\frac{\partial}{\partial \theta_t^i}\right) \stackrel{\text{def}}{=} \frac{\partial J}{\partial \theta_t^i}. \quad (4.61)$$

A solução de (4.60) é encontrada resolvendo o sistema de nT equações

$$\frac{\partial}{\partial \theta_t^i} \int p(x_{0:T}, u_{0:T-1} | \theta) \log \frac{p(x_{0:T}, u_{0:T-1} | \theta)}{q(x_{0:T}, u_{0:T-1})} dx_{0:T} du_{0:T-1} = 0, \quad (4.62)$$

para $i = 1, \dots, n$ e $t = 0, \dots, T-1$.

Admitindo que as distribuições das trajectórias podem ser decompostas pela regra da cadeia das probabilidades, onde intervêm os modelos s e c , e os modelos ideais S e C , de acordo com

$$p(x_{1:T}, u_{0:T-1} | x_0, \theta_{0:T-1}) = \prod_{t=0}^{T-1} s(x_{t+1} | x_t, u_t, t) c(u_t | x_t, \theta_t) \quad (4.63)$$

e

$$q(x_{1:T}, u_{0:T-1} | x_0) = \prod_{t=0}^{T-1} S(x_{t+1} | x_t, u_t, t) C(u_t | x_t, t), \quad (4.64)$$

então o funcional de custo J pode ser rescrito como a soma de dois termos

$$J_{0:T} = J_{t+1:T} + J_{0,t}, \quad 0 \leq t \leq T \quad (4.65)$$

em que $J_{0,t}$ representa a parte do custo devida ao intervalo de tempo de 0 a t e $J_{t+1:T}$ a parte devida ao restante horizonte temporal. O seguinte lema será útil para colocar o funcional J nesta forma.

Lema 9 (Regra da cadeia). *A divergência de Kullback-Leibler entre duas funções $p(z_1, z_2)$ e $q(z_1, z_2)$ satisfaz as igualdades*

$$\begin{aligned} D(p(z_1, z_2) \| q(z_1, z_2)) &= E_{p(z_1)} \left[D(p(z_2 | z_1) \| q(z_2 | z_1)) \right] + \\ &\quad + D(p(z_1) \| q(z_1)) \end{aligned} \quad (4.66)$$

$$\begin{aligned} &= E_{p(z_2)} \left[D(p(z_1 | z_2) \| q(z_1 | z_2)) \right] + \\ &\quad + D(p(z_2) \| q(z_2)). \end{aligned} \quad (4.67)$$

Demonstração. É um resultado standard da teoria da informação (ver [12]). \square

Defina-se $z_1 \stackrel{\text{def}}{=} (x_{0:t}, u_{0:t-1})$ e $z_2 \stackrel{\text{def}}{=} (x_{t+1:T}, u_{t:T-1})$. Então, aplicando o lema 9 obtém-se

$$\begin{aligned} J_{0:T} &= E_{p(x_{0:t}, u_{0:t-1})} \left[D(p(x_{t+1:T}, u_{t:T-1} | x_{0:t}, u_{0:t-1}, \theta_{0:T-1}) \| q(x_{t+1:T}, u_{t:T-1} | x_{0:t}, u_{0:t-1})) \right] \\ &\quad + D(p(x_{0:t}, u_{0:t-1} | \theta_{0:t-1}) \| q(x_{0:t}, u_{0:t-1})) \\ &= E_{p(x_t)} \left[\underbrace{D(p(x_{t+1:T}, u_{t:T-1} | x_t, \theta_{t:T-1}) \| q(x_{t+1:T}, u_{t:T-1} | x_t))}_{=J_{t+1:T}} \right] \\ &\quad + \underbrace{D(p(x_{0:t}, u_{0:t-1} | \theta_{0:t-1}) \| q(x_{0:t}, u_{0:t-1}))}_{=J_{0,t}}. \end{aligned} \quad (4.68)$$

Esta propriedade de separação temporal permite a aplicação de programação dinâmica, obtendo-se um algoritmo recursivo, como se mostra em seguida.

Derivando (4.68) em ordem a θ_t^i obtém-se

$$\frac{\partial J_{0:T}}{\partial \theta_t^i} = \frac{\partial}{\partial \theta_t^i} E_{p(x_t)} \left[D(p(x_{t+1:T}, u_{t:T-1} | x_t, \theta_{t:T-1}) \| q(x_{t+1:T}, u_{t:T-1} | x_t)) \right], \quad (4.69)$$

uma vez que o termo $J_{0:t}$ de (4.68) não depende, directa ou indirectamente, de θ_t^i . Como consequência, a optimização relativamente a θ_t^i é independente dos parâmetros dos controladores usados em instantes de tempo inferiores a t .

Iniciando a optimização no instante $t = T-1$, é possível determinar os parâmetros óptimos θ_{T-1} independentemente de todos os outros instantes de tempo. Em seguida, efectua-se a optimização em $T-2$. A optimização de θ_{T-2} depende dos parâmetros óptimos θ_{T-1} calculados anteriormente. Procedendo-se recursivamente para trás no tempo, obtém-se todos os parâmetros do horizonte temporal de optimização. O teorema seguinte sumariza o algoritmo recursivo que se obtém com este procedimento.

Teorema 10 (Controlo probabilístico paramétrico). *Considere-se um sistema probabilístico $s_{t+1} = s(x_{t+1} | x_t, u_t, t)$, possivelmente variante no tempo, um controlador $c_t = c(u_t | x_t, \theta_t)$, parametrizado em cada instante por um vector de parâmetros θ_t , e uma trajetória ideal probabilística $q = q(x_{1:T}, u_{0:T-1} | x_0)$ gerada pelos sistemas ideais $S_{t+1} = S(x_{t+1} | x_t, u_t, t)$ e $C_t = C(u_t | x_t, t)$ num horizonte temporal de dimensão T . Seja $p = p(x_{1:T}, u_{0:T-1} | x_0)$ a trajetória probabilística gerada pelos sistemas s e c no mesmo horizonte temporal. Um ponto de estacionariedade da função*

$$J_{0:T} = E_{p(x_0)} [D(p \| q)] \quad (4.70)$$

é dado pelo seguinte algoritmo recursivo:

1. Inicialização: $t = T - 1$ e $\gamma_T(x_T) \stackrel{\text{def}}{=} 1$.
2. Calcular os parâmetros θ_t^i , $i = 1, \dots, n$, que são solução do sistema de n equações

$$E_{c_t} \left[\left(E_{s_{t+1}}[-\log \gamma_{t+1}] + D(s_{t+1} \| S_{t+1}) + \log \frac{c_t}{C_t} \right) \frac{\partial \log c(u_t | x_t, \theta_t)}{\partial \theta_t^i} \right] = 0. \quad (4.71)$$

3. Calcular:

$$-\log \gamma_t(x_t) \stackrel{\text{def}}{=} E_{c_t} \left[E_{s_{t+1}}[-\log \gamma_{t+1}] + D(s_{t+1} \| S_{t+1}) + \log \frac{c_t}{C_t} \right]. \quad (4.72)$$

4. Decrementar t e repetir desde 2 enquanto $t \geq 0$.
5. A sequência de parâmetros θ_t obtida nos passos anteriores define o controlo óptimo. O custo total é:

$$J_{0:T} = E_{p(x_0)}[-\log \gamma_0(x_0)]. \quad (4.73)$$

Demonstração. Ver apêndice C.1 □

A condição de optimalidade (4.71) pode ser encarada, num certo sentido, como uma generalização das soluções óptimas explícitas de Kárný. De facto, pode verificar-se que a expressão do controlador probabilístico $c(u_t | x_t)$ de Kárný, obtida em (4.29) e (4.32), satisfaz a equação (4.71) caso a família paramétrica de controladores seja suficientemente abrangente e a contenha. Além desta, é possível obter soluções sub-óptimas, pertencentes a classes paramétricas de controladores, que também satisfazem a equação. Nesse caso, a solução sub-óptima terá um custo superior, mas será o menor custo para a família paramétrica considerada.

4.6 Formulação probabilística do controlo linear Gaussiano

Considere-se o sistema linear

$$x_{t+1} = Ax_t + Bu_t + w_t \quad (4.74)$$

onde $x_t \in \mathbb{R}^n$ é o estado, $u_t \in \mathbb{R}^m$ é a entrada, e $w_t \sim \mathcal{N}(\mathbf{0}, Q_t)$ é uma perturbação Gaussiana de média nula e matriz de covariância Q_t , possivelmente variante no tempo. Suponha-se que se pretende projectar um regulador linear

$$u_t = K_t x_t + v_t, \quad (4.75)$$

onde K_t é uma matriz de ganhos a determinar e $v_t \sim \mathcal{N}(\mathbf{0}, R_t)$ uma perturbação à entrada de média nula a matriz de covariância R_t , também possivelmente variante no tempo. Os modelos (4.74) e (4.75) escrevem-se probabilisticamente da seguinte maneira:

$$x_{t+1} \sim \mathcal{N}(Ax_t + Bu_t, Q_t), \quad (4.76)$$

$$u_t \sim \mathcal{N}(K_t x_t, R_t). \quad (4.77)$$

Definindo uma trajectória probabilística ideal $q(x_{0:T}, u_{0:T-1})$, num horizonte finito T , a partir dos modelos ideais

$$S(x_{t+1}) = \mathcal{N}(\mathbf{0}, \bar{Q}_t) \quad (4.78)$$

e

$$C(u_t) = \mathcal{N}(\mathbf{0}, \bar{R}_t), \quad (4.79)$$

é possível aplicar o algoritmo de controlo probabilístico paramétrico desenvolvido na secção anterior.

Note-se que este problema é diferente do obtido por Kárný em [21] (ver secção 4.4.2). Enquanto que na formulação anterior se considera a mesma matriz de covariância para o modelo real s e ideal S , na formulação aqui considerada as matrizes de covariância Q_t e \bar{Q}_t são arbitrárias. Também, para o controlador, a solução explícita de Kárný impõe uma covariância R_t variante no tempo, obtida no processo de optimização. Aqui considera-se uma matriz de covariância R_t arbitrária, eventualmente constante. Em particular pode considerar-se uma covariância arbitrariamente pequena, aproximando-se assim a implementação de um controlador determinístico, se desejado.

A aplicação do algoritmo do teorema 10 na formulação linear Gaussiana tem como solução os resultados apresentados na tabela 4.3. Nestas expressões são apresentadas matrizes de covariância constantes, embora os resultados sejam igualmente válidos para o caso em que todas as matrizes são variantes no tempo. Os cálculos detalhados apresentam-se em apêndice (ver C.2).

| | |
|----------------|--|
| Inicialização: | $\omega_T = 0, \quad \bar{\Sigma}_T^{-1} = \bar{Q}^{-1}, \quad (\Sigma_T^{-1} = \mathbf{0})$ |
| Ganhos: | $K_t = -(B^T \bar{\Sigma}_{t+1}^{-1} B + \bar{R}^{-1})^{-1} B^T \bar{\Sigma}_{t+1}^{-1} A$ $\bar{\Sigma}_t^{-1} = (A + BK_t)^T \bar{\Sigma}_{t+1}^{-1} (A + BK_t) + K_t^T \bar{R}^{-1} K_t + \bar{Q}^{-1}$ |
| Custo: | $\omega_t = \frac{1}{2} \text{Tr}((B^T \bar{\Sigma}_{t+1}^{-1} B + \bar{R}^{-1})R) + \frac{1}{2} \text{Tr}(\bar{\Sigma}_{t+1}^{-1} Q) + \frac{1}{2} \log \frac{ \bar{Q} \cdot \bar{R} }{ Q \cdot R } - \frac{N+M}{2} + \omega_{t+1}$ $J_{0:T} = \frac{1}{2} E_{p(x_0)} [x_0^T \Sigma_0^{-1} x_0] + \omega_0$ |

Tabela 4.3: Sumário do controlo probabilístico linear gaussiano.

As equações da tabela 4.3 que determinam os ganhos do controlador coincidem precisamente com a solução do problema do regulador linear quadrático (ver [22], p. 52–57), onde as matrizes de penalização são dadas pelas inversas das matrizes de covariância ideais \bar{Q}^{-1} e \bar{R}^{-1} . As equações são, respectivamente, os ganhos de Kalman e a versão estabilizada de Joseph da equação de Riccati.

Realça-se o facto de os ganhos óptimos K_t dependerem apenas das matrizes A e B da dinâmica e das matrizes de covariância \bar{Q} e \bar{R} dos modelos ideais. Isto é, os ganhos do controlador não dependem das perturbações à entrada e saída do processo. No entanto, o custo total no horizonte temporal considerado depende das covariâncias Q e R das perturbações, para além dos parâmetros anteriores.

O termo escalar ω_t do funcional de custo pode ainda ser escrito como

$$\omega_t = \frac{1}{2} \text{Tr} \left(\begin{bmatrix} \bar{\Sigma}_{t+1} & \mathbf{0} \\ \mathbf{0} & \hat{R}_t \end{bmatrix}^{-1} \begin{bmatrix} Q & \mathbf{0} \\ \mathbf{0} & R \end{bmatrix} \right) + \frac{1}{2} \log \frac{|\bar{\Sigma}_{t+1}| \cdot |\hat{R}_t|}{|Q| \cdot |R|} - \frac{N+M}{2} +$$

$$+ \frac{1}{2} \log ((2\pi e)^{n+m} |\bar{Q}| |\bar{R}|) - \frac{1}{2} \log ((2\pi e)^{n+m} |\bar{\Sigma}_{t+1}| |\hat{R}_t|) + \omega_{t+1}, \quad (4.80)$$

onde $\hat{R}_t \stackrel{\text{def}}{=} (B^T \bar{\Sigma}_{t+1}^{-1} B + \bar{R}^{-1})^{-1}$. Na equação (4.80) podem ser reconhecidas expressões correspondentes a uma divergência K-L e duas entropias diferenciais sobre distribuições Gaussianas (ver apêndice C.3.3 e C.3.4). Definindo

$$\hat{p}_t \left(\begin{bmatrix} x_{t+1} \\ u_t \end{bmatrix} \right) \stackrel{\text{def}}{=} \mathcal{N} \left(\begin{bmatrix} Ax_t + Bu_t \\ K_t x_t \end{bmatrix}, \begin{bmatrix} \bar{\Sigma}_{t+1} & \mathbf{0} \\ \mathbf{0} & \hat{R}_t \end{bmatrix} \right) \quad (4.81)$$

e considerando a distribuição gerada pelos modelos reais

$$p_t \left(\begin{bmatrix} x_{t+1} \\ u_t \end{bmatrix} \right) \stackrel{\text{def}}{=} \mathcal{N} \left(\begin{bmatrix} Ax_t + Bu_t \\ K_t x_t \end{bmatrix}, \begin{bmatrix} Q & \mathbf{0} \\ \mathbf{0} & R \end{bmatrix} \right), \quad (4.82)$$

então a equação (4.80) toma a forma da soma das entropias

$$\omega_t = D(p_t \| \hat{p}_t) + h(q_t) - h(\hat{p}_t) + \omega_{t+1}. \quad (4.83)$$

A equação anterior permite uma interpretação envolvendo três distribuições obtidas numa circulação da cadeia de controlo:

- A distribuição ideal q_t , gerada pelas distribuições ideais $S(x_{t+1})$ e $C(u_t)$;
- A distribuição “óptima” \hat{p}_t , gerada pelo modelo \hat{s}_t com a matriz de covariância modificada para $\bar{\Sigma}_{t+1}$, e com o controlo óptimo \hat{c}_t obtido pela solução de Kárný;
- A distribuição real p_t , gerada pelo modelo real $\mathcal{N}(Ax_t + Bu_t, Q)$ e pelo controlador projectado $\mathcal{N}(K_t x_t, R)$.

A optimização do funcional de custo compreende a aproximação da distribuição real p_t à fazível \hat{p}_t , e desta última à ideal q_t .

4.7 Interpretação geométrica

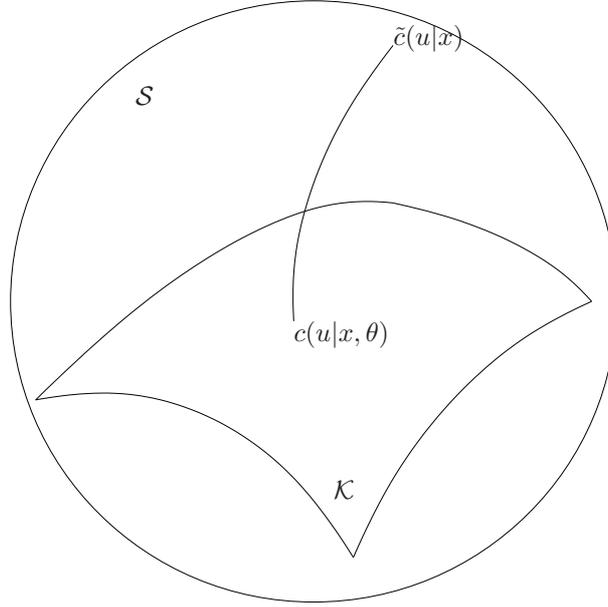
A solução paramétrica sub-óptima apresentada no teorema 10 pode ser interpretada geometricamente usando o formalismo da geometria da informação.

Considerando a família de controladores \mathcal{K} como uma variedade com pontos $c(u_t | x_t, \theta_t)$, sistema de coordenadas θ e vectores tangentes $\partial \log c / \partial \theta_t^i$, então a equação (4.71) pode ser vista como uma condição de ortogonalidade. Na equação

$$E_{c_t} \left[\left(E_{s_{t+1}} [-\log \gamma_{t+1}] + D(s_{t+1} \| S_{t+1}) + \log \frac{c_t}{C_t} \right) \frac{\partial \log c(u_t | x_t, \theta_t)}{\partial \theta_t^i} \right] = 0, \quad (4.84)$$

o factor da esquerda pode ser escrito como uma combinação linear de vectores tangentes à variedade \mathcal{K} mais um termo ortogonal:

$$E_{s_{t+1}} [-\log \gamma_{t+1}] + D(s_{t+1} \| S_{t+1}) + \log \frac{c_t}{C_t} = \sum_{j=1}^n a^j \frac{\partial \log c_t}{\partial \theta_t^j} + \pi^\perp. \quad (4.85)$$

Figura 4.3: Projecção ortogonal de \tilde{c} para a variedade \mathcal{K} .

Da substituição de (4.85) em (4.84) resulta que

$$E_{c_t} \left[\sum_{j=1}^n a^j \frac{\partial \log c_t}{\partial \theta_t^j} \frac{\partial \log c_t}{\partial \theta_t^i} \right] + \underbrace{E_{c_t} \left[\pi^\perp \frac{\partial \log c_t}{\partial \theta_t^i} \right]}_{=0 \text{ por def. de } \pi^\perp} = 0 \quad (4.86)$$

$$\Leftrightarrow \sum_{j=1}^n a^j E_{c_t} \left[\frac{\partial \log c_t}{\partial \theta_t^j} \frac{\partial \log c_t}{\partial \theta_t^i} \right] = 0 \quad (4.87)$$

$$\Leftrightarrow \sum_{j=1}^n a^j g_{ij} = 0, \quad (4.88)$$

$$(4.89)$$

de onde se conclui que, sendo g_{ij} definida positiva (e portanto não singular), então $a^j = 0$ para todo j , *i.e.*, a componente tangente de (4.85) anula-se, restando apenas a componente ortogonal π^\perp . Assim, a otimização consiste em determinar o ponto θ que torna o “vector” (4.85) ortogonal ao espaço tangente $T_\theta \mathcal{K}$. A figura 4.3 ilustra esta ideia.

A equação (4.84) pode ser rescrita de modo a coincidir com a condição de optimalidade na minimização da divergência de K-L entre dois controladores c e \tilde{c} . Para o efeito, define-se a distribuição

$$\tilde{c}_t(u_t|x_t) \stackrel{\text{def}}{=} \frac{1}{\tilde{\gamma}_t(x_t)} C_t \exp \left(- E_{s_{t+1}} [-\log \gamma_{t+1}] - D(s_{t+1} \| S_{t+1}) \right). \quad (4.90)$$

Então, a equação (4.84) pode ser rescrita como

$$E_{c_t} \left[\log \frac{c_t}{\tilde{c}_t} \frac{\partial \log c(u_t|x_t, \theta_t)}{\partial \theta_t^i} \right] = 0. \quad (4.91)$$

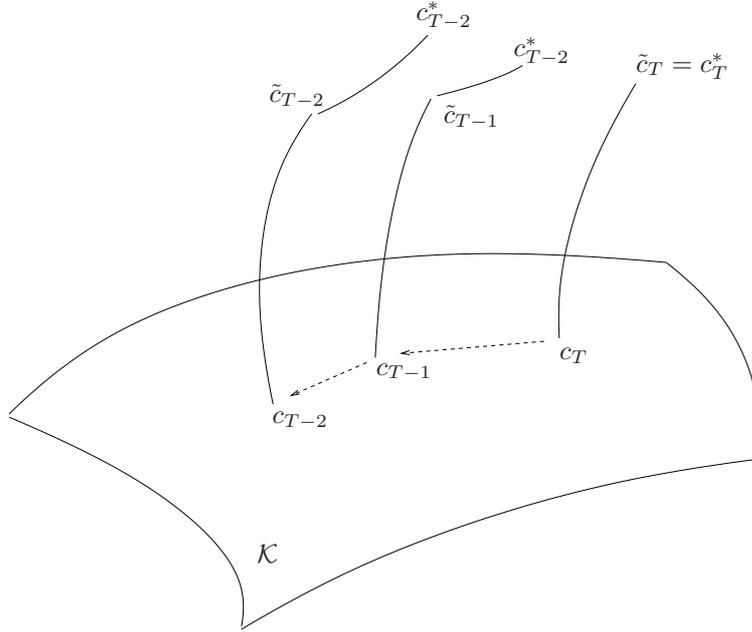


Figura 4.4: A solução paramétrica c_t não é, em geral, obtida pela projecção da solução não paramétrica c_t^* .

Esta equação coincide precisamente com a condição de optimalidade

$$\frac{\partial}{\partial \theta_t^i} D(c_t(u_t|x_t, \theta_t) \parallel \tilde{c}_t(u_t|x_t)) = 0, \quad i = 1, \dots, n, \quad (4.92)$$

correspondente à minimização de $D(c_t \parallel \tilde{c}_t)$. Se a família de controladores \mathcal{K} é uma família mistura, então a perspectiva geométrica apresentada na secção 3.7 é aplicável neste caso. Em particular, o controlador óptimo resulta da projecção ortogonal de \tilde{c} sobre \mathcal{K} ao longo de uma curva exponencial. A aplicação do gradiente natural na optimização resulta num algoritmo excepcionalmente rápido. As secções 4.8 e 4.9 seguintes exploram este caminho.

A equação (4.90) é muito semelhante às equações (4.32) e (4.29), sugerindo que o controlo óptimo paramétrico consiste na projecção ortogonal do controlo óptimo obtido por Kárný. Esta interpretação não é correcta uma vez que o *cost to go*, capturado por γ , é geralmente diferente nos dois casos. Apenas no primeiro passo de optimização, onde é usado $-\log \gamma_T = 0$ em ambos os casos, é válida esta interpretação. Nos restantes instantes da recorrência, a função $-\log \gamma_t(x_t)$ toma formas diferentes nas duas abordagens. A figura 4.4 ilustra intuitivamente esta diferença.

4.8 Optimização de famílias mistura de controladores

Suponha-se que \mathcal{K} é uma família mistura. Para um controlador arbitrário $c \in \mathcal{K}$, a condição de optimalidade não é satisfeita, resultando um conjunto de componentes ε_i

$$E_{c_t} \left[\left(E_{s_{t+1}}[-\log \gamma_{t+1}] + D(s_{t+1} \parallel S_{t+1}) + \log \frac{c_t}{C_t} \right) \frac{\partial \log c(u_t|x_t, \theta_t)}{\partial \theta_t^i} \right] = \varepsilon_i. \quad (4.93)$$

Aplicando o método do gradiente natural para determinar os parâmetros ótimos, resulta no algoritmo

$$\begin{aligned}\dot{\theta}_t^j &= - \sum_i g^{ij}(\theta_t) \frac{\partial D}{\partial \theta_t^i} \\ &= - \sum_i g^{ij}(\theta_t) \varepsilon_i(\theta_t),\end{aligned}\tag{4.94}$$

até que um ponto de estacionariedade $\dot{\theta}_t = \mathbf{0}$ seja encontrado.

A aplicação deste método requer a computação da inversa g^{ij} da métrica g_{ij} , e vários valores esperados com o intuito de determinar as componentes ε_i . Estes cálculos não têm, em geral, solução analítica no caso contínuo, e são difíceis de obter numericamente, sendo necessário recorrer a métodos de integração e amostragem como por exemplo Markov chain Monte Carlo (MCMC). No caso discreto, por outro lado, a solução é computacionalmente muito eficiente como se viu no capítulo 3.

4.9 Optimização de famílias discretas de controladores

Suponha-se que o sistema probabilístico $s(x_{t+1}|x_t, u_t)$ é uma cadeia de Markov com entrada $u_t \in \mathcal{U} \stackrel{\text{def}}{=} \{0, 1, 2, \dots, m\}$ e estado $x_t \in \mathcal{X} \stackrel{\text{def}}{=} \{0, 1, 2, \dots, n\}$, e que o controlador é caracterizado pela distribuição condicional discreta $c(u_t|x_t, \theta_t)$.

As probabilidades de transição do controlador podem ser escritas, em cada instante de tempo, numa matriz de dimensões $(m+1) \times (n+1)$, em que cada elemento (i, j) da matriz contém a probabilidade $c(u_t = i|x_t = j)$:

$$\begin{bmatrix} \theta^{0,0} & \theta^{0,1} & \dots & \theta^{0,n} \\ \theta^{1,0} & \theta^{1,1} & & \theta^{1,n} \\ \vdots & & \ddots & \vdots \\ \theta^{m,0} & \theta^{m,1} & \dots & \theta^{m,n} \end{bmatrix} (t)\tag{4.95}$$

O conjunto de todos os controladores deste tipo são parametrizados por $m \times (n+1)$ parâmetros independentes $\theta^{i,j}$, $i = \{1, \dots, m\}$, $j = \{0, 1, \dots, n\}$. A primeira linha de (4.95), caracterizada pelas probabilidades θ^{0j} , é determinada automaticamente pela restrição $\sum_{i=0}^m \theta^{ij} = 1$, de onde resulta que

$$\theta^{0j} = 1 - \sum_{i=1}^m \theta^{ij}.\tag{4.96}$$

A família de controladores caracterizada desta forma é uma família mistura.

O controlo probabilístico óptimo é encontrado resolvendo a condição de optimalidade que resulta, neste caso, no sistema de equações

$$\sum_{u_t=0}^m \left(\log \frac{c_t}{\tilde{c}_t} \right) \frac{\partial c_t}{\partial \theta_t^{ij}} = 0,\tag{4.97}$$

para todo $i = \{1, \dots, m\}$, $j = \{0, 1, \dots, n\}$, $x \in \mathcal{X}$, e onde

$$\tilde{c}_t = C_t \exp \left(-D(s_{t+1}||S_{t+1}) - E_{s_{t+1}}[-\log \gamma_{t+1}] \right)\tag{4.98}$$

$$-\log \gamma_t(x_t) = \sum_{u_t=0}^m c_t \log \frac{c_t}{\tilde{c}_t}.\tag{4.99}$$

A derivada parcial existente na equação (4.97) resulta, em cada instante de tempo, num objecto indexado pelas quatro variáveis (i, j, u_t, x_t) . A aplicação directa da condição (4.97) resulta na manipulação de objectos de ordem quatro obtendo-se, após o somatório, um objecto de ordem três indexado por (i, j, x_t) .

É possível simplificar este processo observando que a derivada parcial se anula para todo $x_t \neq j$ e consequentemente a condição de optimalidade é automaticamente garantida nestes casos. Esta observação permite reduzir de uma ordem a equação (4.97), impondo $x_t = j$ e resolvendo com esta restrição.

Fixando sucessivamente cada $x \in \mathcal{X}$, é possível escrever a equação na forma matricial, sendo necessário resolver uma equação por cada $x \in \mathcal{X}$. Assim, obtém-se a matriz

$$\begin{bmatrix} \frac{\partial c(u=0|x)}{\partial \theta^1} & \cdots & \frac{\partial c(u=0|x)}{\partial \theta^m} \\ \vdots & & \vdots \\ \frac{\partial c(u=m|x)}{\partial \theta^1} & \cdots & \frac{\partial c(u=m|x)}{\partial \theta^m} \end{bmatrix} = \begin{bmatrix} -1 & -1 & \cdots & -1 \\ 1 & 0 & \cdots & 0 \\ 0 & 1 & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}. \quad (4.100)$$

A condição de optimalidade (4.97) pode então ser escrita como

$$\begin{bmatrix} \log \frac{c(u=0|x)}{\tilde{c}(u=0|x)} & \cdots & \log \frac{c(u=m|x)}{\tilde{c}(u=m|x)} \end{bmatrix} \begin{bmatrix} -1 & -1 & \cdots & -1 \\ 1 & 0 & \cdots & 0 \\ 0 & 1 & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} = \mathbf{0}_{1 \times m} \quad (4.101)$$

para cada $x \in \mathcal{X}$. Como a matriz (4.100) é constante (não depende de x), então, o conjunto das $n + 1$ equações pode ser escrito como uma única equação matricial

$$\begin{bmatrix} \log \frac{c(u=0|x=0)}{\tilde{c}(u=0|x=0)} & \cdots & \log \frac{c(u=m|x=0)}{\tilde{c}(u=m|x=0)} \\ \vdots & & \vdots \\ \log \frac{c(u=0|x=n)}{\tilde{c}(u=0|x=n)} & \cdots & \log \frac{c(u=m|x=n)}{\tilde{c}(u=m|x=n)} \end{bmatrix} \begin{bmatrix} -1 & -1 & \cdots & -1 \\ 1 & 0 & \cdots & 0 \\ 0 & 1 & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} = \mathbf{0}_{(n+1) \times m}. \quad (4.102)$$

Esta equação corresponde a um sistema de $(n + 1) \times m$ equações a $(n + 1) \times m$ incógnitas, e tem uma solução que pode ser obtida forçando a matriz da esquerda a ter linhas constantes. Resulta então que

$$c(u|x) = \frac{1}{\gamma(x)} \tilde{c}(u|x). \quad (4.103)$$

Esta solução coincide com a apresentada na equação (4.29). Em seguida apresenta-se um algoritmo que permite obter a mesma solução de forma iterativa.

4.9.1 Solução pelo método do gradiente natural

Suponha-se que é proposta inicialmente uma distribuição candidata $c(u|x)$, diferente da distribuição óptima. Ao aplicar esta distribuição em (4.102) a igualdade não se verifica, obtendo-se do lado direito uma matriz ε não nula, correspondente ao diferencial dD . Esta matriz ε é usada para calcular o gradiente natural e determinar a direcção a seguir pelo método iterativo.

O algoritmo seguinte apresenta a sequência de cálculos usados na optimização:

1. Inicialização:

- (a) $t = T - 1$, onde T é o tamanho do horizonte temporal;
- (b) $-\log \gamma(x_T) = 0$;
- (c) Especificar uma estimativa inicial da matriz de parâmetros θ para o controlador no instante $t = T$;
- (d) Definir o comprimento do passo η a usar no método do gradiente natural;

2. Calcular:

$$\log \tilde{c}_t \stackrel{\text{def}}{=} \log C_t - D(s_{t+1} \| S_{t+1}) - E_{s_{t+1}}[-\log \gamma_{t+1}(x_{t+1})]. \quad (4.104)$$

3. Repetir os seguintes cálculos:

(a) Cálculo de ε^T :

$$\begin{aligned} \varepsilon^T = & - \begin{bmatrix} 1 \\ \vdots \\ \vdots \\ 1 \end{bmatrix} \begin{bmatrix} \log \frac{c(u=0|x=0)}{\tilde{c}(u=0|x=0)} & \cdots & \log \frac{c(u=0|x=n)}{\tilde{c}(u=0|x=n)} \end{bmatrix} + \\ & + \begin{bmatrix} \log \frac{c(u=1|x=0)}{\tilde{c}(u=1|x=0)} & \cdots & \log \frac{c(u=1|x=n)}{\tilde{c}(u=1|x=n)} \\ \vdots & & \vdots \\ \log \frac{c(u=m|x=0)}{\tilde{c}(u=m|x=0)} & \cdots & \log \frac{c(u=m|x=n)}{\tilde{c}(u=m|x=n)} \end{bmatrix}, \end{aligned} \quad (4.105)$$

onde é usada a decomposição de matrizes $\begin{bmatrix} A & B \\ C & D \end{bmatrix}^T = AC + BD$, de forma a eliminar todos os produtos de (4.102).

(b) Cálculo do gradiente natural:

$$\tilde{\nabla} D = \theta \circ \varepsilon^T - \theta \cdot \text{diag}([1 \ \cdots \ 1] (\theta \circ \varepsilon^T)). \quad (4.106)$$

Esta expressão para o gradiente natural é uma variante da obtida no capítulo 3, modificada para trabalhar em espaços de matrizes.

(c) Actualização dos parâmetros θ na direcção do gradiente natural:

$$\theta \leftarrow \theta - \eta \tilde{\nabla} D. \quad (4.107)$$

Esta actualização corresponde a evoluir o controlador c_t ao longo de uma geodésica (relativamente à conexão mistura) disparada na direcção do gradiente natural.

Os três passos 3a, 3b e 3c anteriores são repetidos até que a norma Riemanniana da matriz ε seja essencialmente nula. Para o efeito, estabelece-se um limiar ϵ e verifica-se se a condição

$$(\varepsilon \tilde{\nabla} D)_{ii} < \epsilon, \quad i = 0, 1, \dots, n, \quad (4.108)$$

é satisfeita para os elementos da diagonal. Em caso afirmativo a iteração é interrompida e os parâmetros θ actuais são considerados óptimos.

4. Usando os parâmetros θ óptimos, calcular

$$-\log \gamma_t(x_t) = E_{c_t} \left[\log \frac{c_t}{\tilde{c}_t} \right]. \quad (4.109)$$

5. Decrementar t , e repetir o cálculo do controlador para esse instante de tempo repetindo todos os passos desde 2. O procedimento é repetido para todos os instantes de tempo t do horizonte temporal considerado.

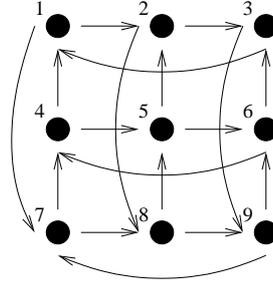


Figura 4.5: Transições permitidas no sistema $s(x_{t+1}|x_t, u_t)$.

4.10 Exemplos

Para testar o método anterior foram criados exemplos com vários graus de complexidade. No primeiro exemplo considera-se um sistema de baixa dimensão. No segundo exemplo considera-se um sistema de dimensão elevada, que prova a viabilidade prática dos algoritmos desenvolvidos.

4.10.1 Problema de baixa complexidade

Considere-se que o sistema $s(x_{t+1}|x_t, u_t)$ tem nove estados x e três entradas u . Para facilitar a visualização, os estados estão organizados numa grelha 3×3 . O sistema permite apenas três tipos de transições entre estados: manter o mesmo estado, rotação para a direita, e rotação para cima. A figura 4.5 mostra as transições permitidas. As probabilidades de transição dependem da entrada externa u . A tabela seguinte mostra de que forma a entrada u influencia as transições: Suponha-se que se pretende levar o sistema de um estado inicial

| u | manter estado | rodar para direita | rodar para cima |
|-----|---------------|--------------------|-----------------|
| 1 | 0.9 | 0.05 | 0.05 |
| 2 | 0.05 | 0.05 | 0.9 |
| 3 | 0.05 | 0.9 | 0.05 |

Tabela 4.4: Probabilidades de transição em função da entrada u .

arbitrário ao estado $x = 3$, usando para o efeito as entradas u sem nenhuma preferência especial. Assim, definiram-se as distribuições ideais

$$S(x) = \frac{1}{1008} [1 \ 1 \ 1000 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1]^T, \quad (4.110)$$

$$C(u) = \left[\frac{1}{3} \ \frac{1}{3} \ \frac{1}{3} \right]^T, \quad (4.111)$$

onde, em $S(x)$, a probabilidade elevada marca o estado desejado.

Aplicando o algoritmo de optimização iterativo num horizonte temporal de comprimento $T = 10$, obtém-se a evolução dos parâmetros como ilustrado na figura 4.6⁶. A optimização

⁶Neste gráfico, a seta do tempo está invertida. Em vez desta, está representada a seta “causal” usada durante a optimização. O índice $T - t$ indicado no eixo horizontal reflecte esta inversão.

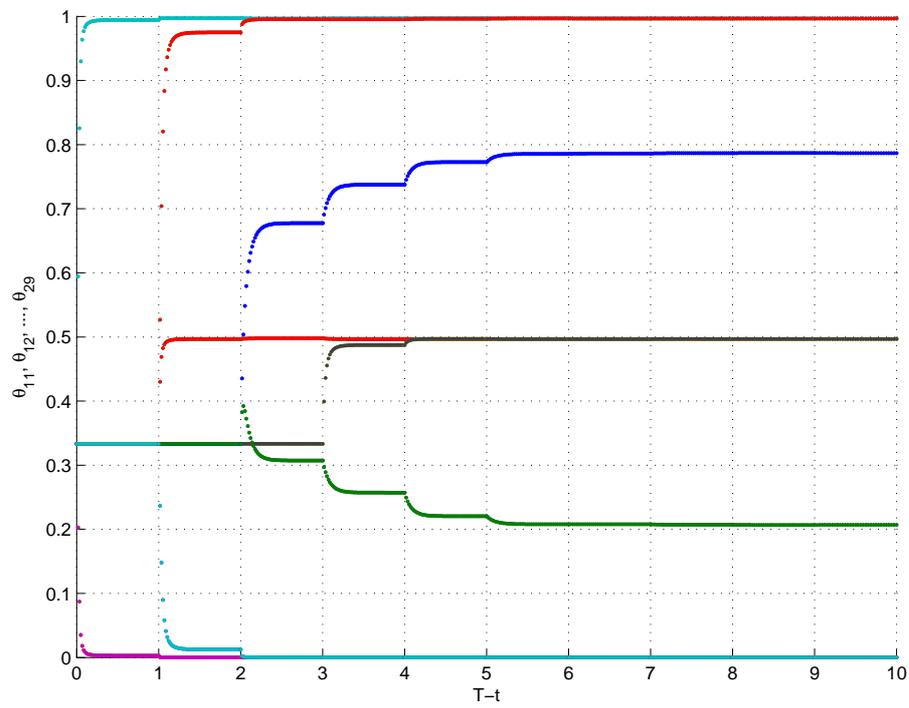


Figura 4.6: Optimizaç o de $c(u|x)$. Os instantes $T - t$ com n meros inteiros correspondem   programaç o din mica, enquanto cada intervalo entre estes corresponde a uma corrida do m todo do gradiente natural.

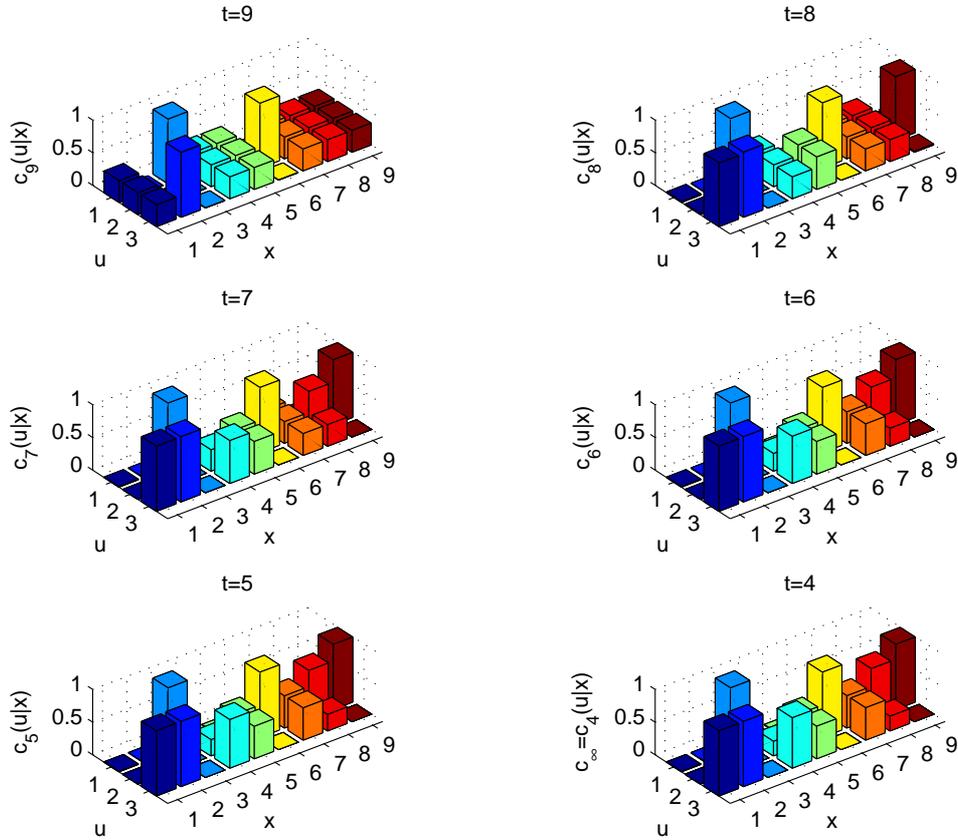


Figura 4.7: Matrizes de transição $c_t(u|x)$, $t = T - 4, \dots, T - 1$, no horizonte temporal de dimensão $T = 10$.

de $c(u|x)$ é iniciada com distribuição uniforme. A otimização para o instante $t = T - 1 = 9$, correspondente à posição $T - t = 1$ no gráfico, devolve um controlo “um-passo-à-frente”. As probabilidades de transição são apresentadas no primeiro gráfico de barras da figura 4.7. Nos quatro instantes de tempo seguintes da otimização, as probabilidades de transição sofrem alterações consideráveis. Ao fim de cinco instantes de tempo obtém-se uma aproximação razoável do controlador $c_\infty(u|x)$. Usando a lei de controlo c_∞ numa simulação estocástica em cadeia fechada, obtiveram-se os histogramas normalizados da figura 4.8. Como se pode observar, o histograma normalizado $p(x)$ do estado x aproxima-se razoavelmente da distribuição ideal $S(x)$. A entrada u apresenta um histograma normalizado diferente da distribuição ideal $C(u)$. Esta diferença é resultado de um compromisso que permite a x ter uma distribuição estacionária perto da especificada por $S(x)$.

4.10.2 Problema de complexidade elevada

No exemplo descrito nesta secção considera-se um sistema linear de primeira ordem, instável em cadeia aberta,

$$x_{t+1} = 1.01x_t + 0.01u_t + w_t, \quad w_t \sim \mathcal{N}(0, 0.1). \quad (4.112)$$

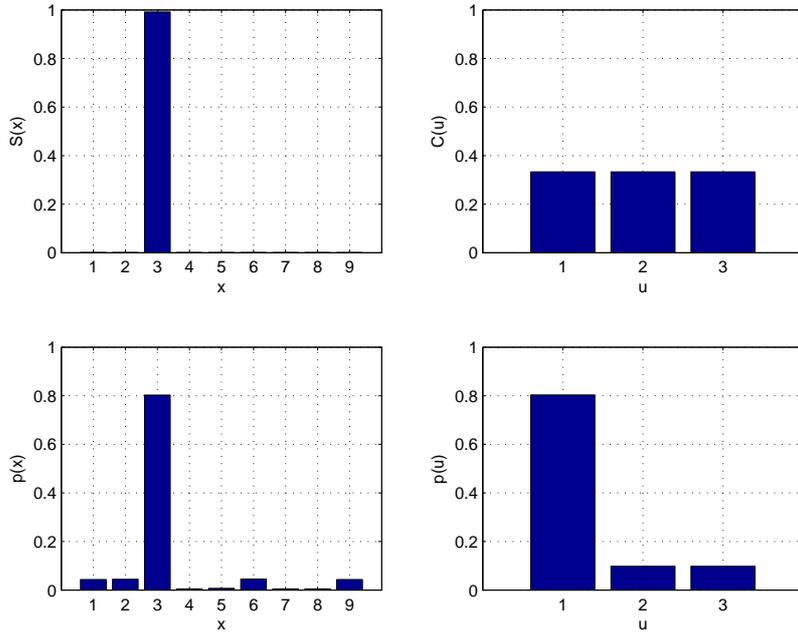


Figura 4.8: Distribuições de probabilidade dos modelos ideais $S(x)$ e $C(u)$, e histograma normalizado das variáveis x e u obtidas em simulação estocástica em regime estacionário.

Na formulação probabilística, o estado x é amostrado com uma precisão de 8 bits no intervalo $[-2, 2]$, e o controlo u no intervalo $[-10, 10]$ com a mesma precisão de 8 bits. Assim, o modelo probabilístico discreto $s(x_{t+1}|x_t, u_t)$ é descrito por $256 \times 256 \times 256 = 16\,777\,216$ probabilidades de transição. Para este modelo, um controlador $c_t(u_t|x_t)$ é descrito por $256 \times 256 = 65\,536$ probabilidades de transição, que podem ser parametrizadas por $255 \times 256 = 65\,280$ parâmetros θ^i independentes.

Pretende-se determinar a sucessão de controladores c_t num horizonte temporal de comprimento T , onde a optimização é efectuada relativamente a uma distribuição ideal gerada pelos modelos ideais $S(x)$ e $C(u)$. Como modelo ideal $S(x)$ para x é especificada uma distribuição bimodal obtida pela mistura de duas gaussianas

$$S(x) \stackrel{\text{def}}{=} 0.5 \mathcal{N}(-1, 0.2) + 0.5 \mathcal{N}(1, 0.2). \quad (4.113)$$

O modelo ideal $C(u)$ especificado para u é

$$C(u) \stackrel{\text{def}}{=} \mathcal{N}(0, 3). \quad (4.114)$$

Efectuando a optimização pelo método iterativo num horizonte temporal $T = 100$, obtém-se um controlo óptimo $c_t(u_t|x_t)$ caracterizado por um conjunto de 100 matrizes de transição de dimensão 256×256 , uma para cada instante de tempo $t = 0, 1, \dots, 99$. Assim, o algoritmo necessita de determinar um total de $100 \times 255 \times 256 = 6\,528\,000$ parâmetros diferentes.

Aplicando o algoritmo iterativo obtém-se o controlador $c_\infty \approx c_0$ com as probabilidades de transição que se mostram graficamente na figura 4.9 (nesta figura mostra-se uma matriz de transição 64×64 por motivos de legibilidade). Foi efectuada uma simulação estocástica em malha fechada com o controlador c_∞ . A figura 4.10 mostra o histograma normalizado

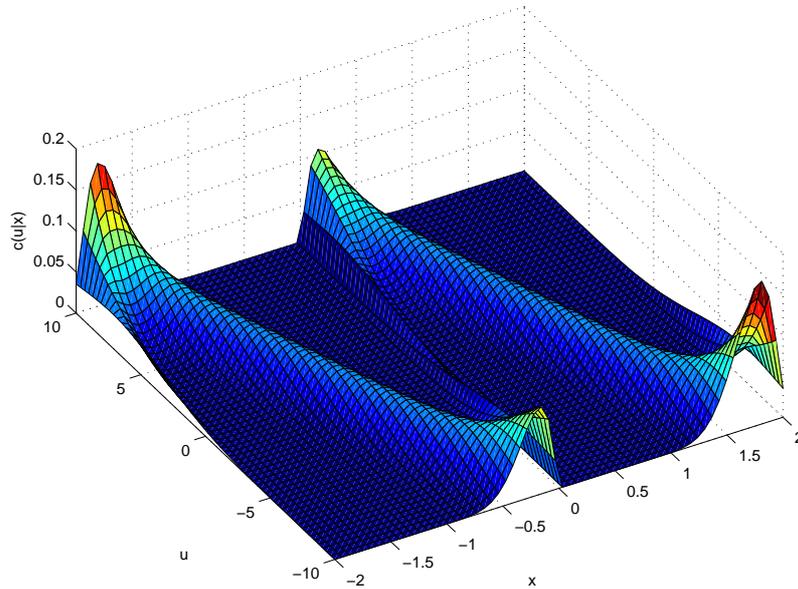


Figura 4.9: Probabilidades de transição do controlador $c_0(u|x)$.

obtido em regime estacionário num intervalo de tempo grande. Mostra-se também as curvas correspondentes às distribuições ideais $S(x)$ e $C(u)$. Conforme se pode observar, o estado x aproxima-se da distribuição ideal $S(x)$. Em geral, não é possível fazer coincidir as distribuições estacionárias com os modelos ideais $S(x)$ e $C(x)$. A solução obtida é sempre a que melhor se aproxima no sentido da divergência K-L, mas cujo desempenho está limitado pelas propriedades do sistema real $s(x_{t+1}|x_t, u_t)$ e pela família paramétrica de controladores especificada. A figura 4.11 mostra um excerto da simulação estocástica da cadeia fechada.

4.11 Aplicação à genómica

Nesta secção aplicam-se as ferramentas de controlo probabilístico desenvolvidas anteriormente em problemas com redes reguladoras de genes.

A secção 4.11.1 apresenta um resumo da genómica relevante para o problema que se pretende resolver. Para uma exposição mais detalhada consultar [13], ou livros da especialidade aí citados.

4.11.1 Introdução

Num sistema celular, cada célula desempenha funções bem definidas, determinadas pela sua informação genética. Nos casos em que a informação genética é adulterada, a função normal da célula pode ser alterada e falhar, gerando-se situações de doença. A investigação em genética tem como objectivo compreender de que forma as células executam a sua função, identificando os genes responsáveis pelos diferentes processos e mecanismos de controlo empregues.

A informação genética é armazenada numa molécula de ADN (*ácido desoxirribonucleico*⁷).

⁷DNA na designação anglosaxónica.

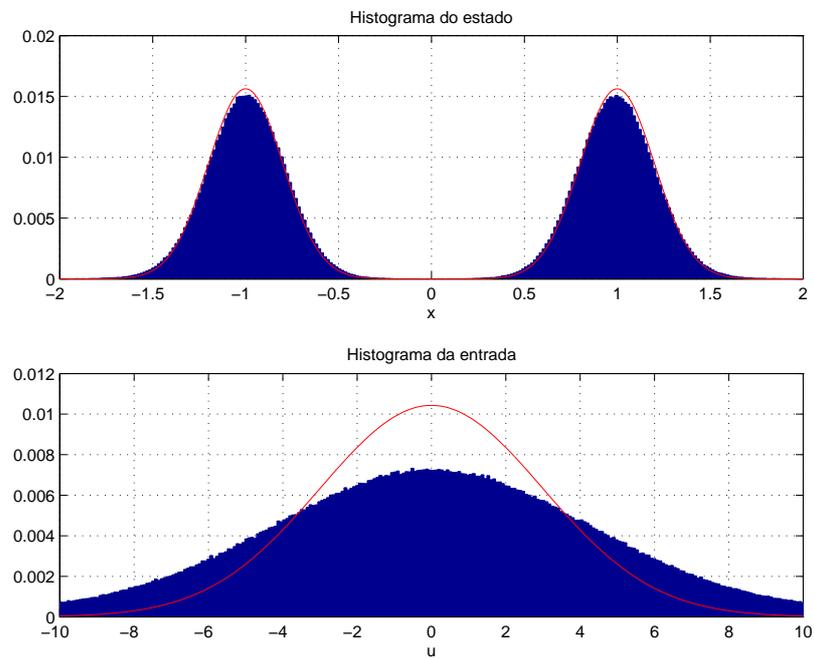


Figura 4.10: Histograma normalizado do estado x e da entrada u obtidos em simulação estocástica em regime estacionário.

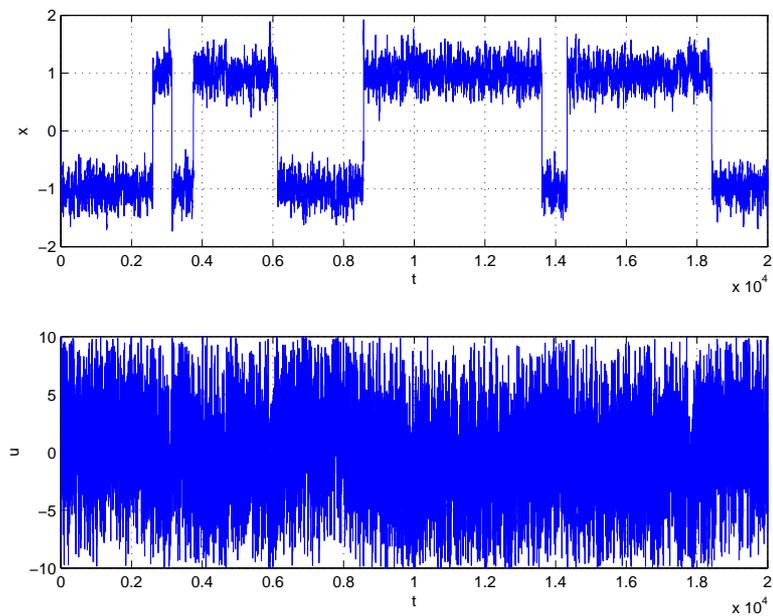


Figura 4.11: Evolução dos sinais x_t e u_t numa simulação estocástica em regime estacionário.

A molécula de ADN consiste em duas cadeias de nucleótidos organizados na forma duma dupla hélice. Os nucleótidos são de quatro tipos: Adenina (A), Guanina (G), Citosina (C) e Timina (T). Estes nucleótidos emparelham-se em pares A-T e G-C por pontes de hidrogénio. Assim, se uma cadeia é formada por uma sequência das bases A, G, C e T, a outra cadeia é-lhe complementar pelas regras de emparelhamento anteriores. O ADN codifica a informação na sequência de nucleótidos usados. Organismos diferentes têm sequências diferentes e transportam informação diferente.

Uma tira de ADN pode ser dividida em pequenas porções, chamadas *genes*, que transportam a informação necessária para produzir *proteínas*. As proteínas são formadas por uma cadeia de 20 tipos diferentes de aminoácidos. A sequência destes aminoácidos determina a forma da proteína que, por sua vez, determina a função exercida por esta. As proteínas são responsáveis por praticamente todas as funções celulares. Alguns exemplos de proteínas são as enzimas, os anticorpos, as toxinas e as hormonas. As proteínas servem também como moléculas transmissoras de sinais entre células ou como as partes móveis de máquinas moleculares existentes no interior das células. São moléculas responsáveis por praticamente todas as funções celulares.

À informação codificada no ADN de um organismo dá-se o nome de *genoma*. Durante a divisão celular, a célula copia o seu genoma para as novas células. Como as duas cadeias de nucleótidos do ADN são complementares uma da outra, cada uma destas pode servir de modelo para construir a outra. É precisamente esta propriedade que é usada para executar a replicação do ADN. As cadeias de nucleótidos são separadas, e cada cadeia é usada para sintetizar a sua complementar, obtendo-se duas cópias iguais do ADN original.

Os mecanismos usados no processo de replicação são extremamente robustos a erros, incluindo mesmo mecanismos de correcção de erros. No entanto estes não estão livres de ocorrer. Caso estes ocorram diz-se que houve *mutações* nos genes. As mutações podem dar origem a doenças hereditárias, ou cancro, mas são também o que permite a evolução das espécies.

Como referido anteriormente, o ADN transporta a informação necessária para a produção de várias proteínas, sendo essa informação codificada nos vários genes. No entanto a produção das proteínas não é efectuada directamente a partir do ADN. A porção do ADN com a informação necessária à produção de uma determinada proteína – o gene – é inicialmente copiada para um *ARNm* (ácido ribonucleico mensageiro⁸) num processo chamado *transcrição*. O ARNm é então usado para dirigir a produção da proteína noutro processo designado de *tradução*. Quando um gene é transcrito diz-se que este é *expresso* ou que está *activado*.

Durante o processo de tradução, a sequência de nucleótidos do ARNm é descodificada, traduzindo-se a linguagem baseada nos quatro nucleótidos A-G-C-U que formam o ARNm nos 20 aminoácidos que formam a linguagem das proteínas. O código que efectua esta tradução chama-se *código genético*. Este código faz corresponder a cada grupo de três nucleótidos do ARNm um aminoácido da proteína. Assim, obtêm-se $4^3 = 64$ palavras de código diferentes. De todas estas, três são usadas para marcar o início e fim da proteína, sendo as restantes 61 usadas para codificar os 20 aminoácidos. O código genético é o mesmo em toda a vida conhecida.

Os genes não estão todos activos simultaneamente. Num organismo multicelular com diversos tipos de células, embora o genoma seja o mesmo em todas elas, a expressão dos genes (activos/inibidos) é diferente entre elas. Estas diferenças de expressão levam à produção de substâncias diferente (insulina nas células- β do pâncreas, hemoglobina nos glóbulos vermelhos, *etc*) e são o que distingue a forma e função das várias células.

A divisão celular que ocorre no desenvolvimento embrionário a partir de um óvulo fer-

⁸mRNA na designação anglosaxónica.

tilizado – que é uma célula – dá origem aos diversos tipos de células de um organismo. Esta *diferenciação celular* é conseguida pelo controlo da expressão dos genes ao longo das sucessivas divisões celulares.

4.11.2 Redes Reguladoras de Genes

O mecanismo de transcrição é regulado por *proteínas reguladoras de genes* que se ligam a *sequências de ADN reguladoras*, e que podem inibir ou reforçar o ritmo a que a transcrição é efectuada. Ao serem activados determinados genes, podem ser produzidas proteínas – factores de transcrição – que inibem ou reforçam o processo de transcrição de outros genes. Assim, enquanto uns genes codificam proteínas necessárias à função da célula, outros genes têm uma função reguladora.

O conjunto dos genes conjuntamente com outras substâncias existentes na célula formam uma rede de influência mútua. Nesta rede, os nós representam o estado (activo/inactivo) dos genes, enquanto a malha representa a influência que cada gene tem nos restantes por intermédio dos factores de transcrição.

Em organismos unicelulares, estas redes podem responder ao ambiente exterior activando ou desactivando genes. Por exemplo, uma célula de levedura num ambiente rico em açúcares irá activar genes que produzem as enzimas necessárias para transformar o açúcar em álcool. A energia obtida com este processo é então usada para se multiplicar pelo processo de divisão celular.

Em organismos multicelulares o processo é semelhante. O estado de determinados genes pode depender do ambiente, pelo que a concentração de certas substâncias poderão levar à sua activação ou inibição. A existência de gradientes de concentração destas substâncias leva a que células em locais distintos tenham diferentes genes activos, adquirindo assim funções específicas diferentes. Este é o mecanismo que permite o desenvolvimento dos vários órgãos de um corpo.

A modelação matemática das redes reguladoras de genes pode ser feita usando sistemas de equações diferenciais ordinárias, onde os estados representam as concentrações das várias substâncias. Estes modelos fornecem informação detalhada sobre a evolução do sistema, mas requerem um volume muito elevado de dados para a calibração do modelo.

Um modelo mais grosseiro que requer menos informação pode ser implementado usando redes booleanas. Nestas redes, os estados (expressão dos genes) podem tomar dois valores: *on* ou *off*. A activação dos genes depende de funções Booleanas dos restantes genes e do ambiente. Mais recentemente têm sido usados processos estocásticos para modelar a expressão dos genes em redes Booleanas probabilísticas (*PBNs*).

4.11.3 Modelação de Redes Reguladoras de Genes

Considerando que um gene pode estar num dos dois estados *on* ou *off*, um conjunto de n genes tem na totalidade 2^n combinações de estados possíveis. A cada combinação pode ser atribuído um código binário, onde cada *bit* representa o estado de um gene. Como exemplo considere-se que um conjunto de três genes estão num dado momento no estado 001, *i.e.* apenas o terceiro gene (o da direita) está expresso. Se o primeiro gene é activado, o novo estado passa a ser 101. Lido em base decimal, diz-se que ocorreu uma transição do estado 1 para o 5.

A activação de um gene depende dos factores de transcrição promovidos por outros genes e do ambiente, pelo que a interacção com determinadas substâncias poderá despoletar alterações na sua expressão. Estes sinais externos podem ser emitidos por outras células vizinhas ou por moléculas extracelulares existentes no ambiente.

A literatura tem sugerido a utilização de redes Booleanas para modelar a evolução do estado de activação dos genes [13]. Neste tipo de formulação considera-se que o estado x_i do i -ésimo gene evolui de acordo com

$$x_i(t+1) = f_i(x_0, x_1, \dots, x_{n-1}),$$

onde f_i são funções Booleanas. Uma lei de evolução de tipo determinista dificilmente coincide com a realidade observada. Assim, tendo em consideração a natureza estocástica dos sistemas, é sugerida [13] a utilização de redes Booleanas probabilísticas (PBN). Nestas redes, cada função Booleana f_i é seleccionada de forma aleatória, em cada instante, de um conjunto de funções Booleanas $F_i = \{f_j^{(i)}\}_{j=1, \dots, l(i)}$. Estes conjuntos contêm os “preditores”⁹ que irão descrever a transição do estado. Isto é, para cada gene i , existe um conjunto de $l(i)$ funções Booleanas $\{f_1^{(i)}, f_2^{(i)}, \dots, f_{l(i)}^{(i)}\}$, que aplicadas ao estado actual da rede predizem o estado futuro do i -ésimo gene, tendo cada uma delas uma probabilidade c_i associada. Em seguida propõe-se um modelo alternativo para descrever a dinâmica da rede.

Admitindo a mesma codificação binária para o estado de cada gene, a evolução da rede pode ser modelada por uma cadeia de Markov $p(x_{t+1}|x_t)$, onde $x_t \in \{0, 1, \dots, 2^n - 1\}$ representa o estado da rede formado pela palavra binária correspondente ao conjunto dos genes.

Pode verificar-se que um modelo PBN determina univocamente uma cadeia de Markov. O inverso já não se verifica, podendo existir várias PBNs para a mesma cadeia de Markov.

Admita-se que a cadeia de Markov é descrita pela distribuição condicional $s(x_{t+1}|x_t, u_t)$ onde a variável u_t modela a influência que o ambiente externo à célula tem nas probabilidades de transição. Se a variável u_t depende de alguma maneira da variável x_t então obtém-se um sistema realimentado. Mais especificamente, a variável u_t pode ser descrita pela distribuição condicional $c(u_t|x_t, \theta_t)$, parametrizada por θ_t .

Um dos objectivos que se coloca é o de modificar a expressão genética de um conjunto de células. Este controlo permite, por exemplo, formar células especializadas a partir de células indiferenciadas, ou induzir apoptose em células cancerígenas.

Tendo como objectivo um determinado perfil genético final e uma sequência de transições entre estados para atingir esse objectivo, é necessário determinar que sequência de sinais u_t devem ser usados para aproximar o objectivo. A formulação do controlo probabilístico pode então ser usada para este propósito. O exemplo 11 seguinte ilustra esta abordagem sobre um sistema fictício.

Exemplo 11. *Considere-se o exemplo da figura 4.12, onde se pretende regular três genes pela adição de uma substância U . A existência em maior quantidade da substância U conduz a uma maior probabilidade de interacção com a rede de regulação de uma célula particular.*

Neste exemplo considera-se, não uma célula particular, mas uma população de células que evolui de forma estocástica. Admite-se que cada gene tem dois estados possíveis: on ou off. Admite-se também que a expressão de um gene determina a expressão de outro com probabilidade $p = 0.9$.

Num dado instante, a rede pode estar num de $2^3 = 8$ estados possíveis, numerados em binário de 000 a 111. Nesta rede, o gene 0 (zero) corresponde ao bit menos significativo enquanto o gene 2 (dois) corresponde ao bit mais significativo. As probabilidades de transição entre estados da rede dependem do modo como as influências entre os genes se dão. Para a rede da figura 4.12 determinaram-se as probabilidades de transição (4.115) e (4.116), correspondentes aos casos de presença e ausência de interacção com a substância U . Escrevem-se, com algum abuso de notação, as matrizes das probabilidades de transição $s(x_{t+1}|x_t, u_t = 0)$ e $s(x_{t+1}|x_t, u_t = 1)$, onde $x_{t+1} \in \{000, 001, \dots, 111\}$ indexa as linhas de 1 a 8, enquanto x_t

⁹Conforme nomenclatura usada em [13].

indexa as colunas de igual modo. Apresentam-se em negrito as transições mais prováveis da rede.

$$s(x_{t+1}|x_t, u_t = 0) = \begin{bmatrix} 0.009 & 0.081 & 0.081 & \mathbf{0.729} & 0.009 & 0.081 & 0.081 & \mathbf{0.729} \\ 0.081 & \mathbf{0.729} & 0.009 & 0.081 & 0.081 & \mathbf{0.729} & 0.009 & 0.081 \\ 0.001 & 0.009 & 0.009 & 0.081 & 0.001 & 0.009 & 0.009 & 0.081 \\ 0.009 & 0.081 & 0.001 & 0.009 & 0.009 & 0.081 & 0.001 & 0.009 \\ 0.081 & 0.009 & \mathbf{0.729} & 0.081 & 0.081 & 0.009 & \mathbf{0.729} & 0.081 \\ \mathbf{0.729} & 0.081 & 0.081 & 0.009 & \mathbf{0.729} & 0.081 & 0.081 & 0.009 \\ 0.009 & 0.001 & 0.081 & 0.009 & 0.009 & 0.001 & 0.081 & 0.009 \\ 0.081 & 0.009 & 0.009 & 0.001 & 0.081 & 0.009 & 0.009 & 0.001 \end{bmatrix}, \quad (4.115)$$

$$s(x_{t+1}|x_t, u_t = 1) = \begin{bmatrix} 0.009 & 0.081 & 0.081 & \mathbf{0.729} & 0.001 & 0.009 & 0.009 & 0.081 \\ 0.081 & \mathbf{0.729} & 0.009 & 0.081 & 0.009 & 0.081 & 0.001 & 0.009 \\ 0.001 & 0.009 & 0.009 & 0.081 & 0.009 & 0.081 & 0.081 & \mathbf{0.729} \\ 0.009 & 0.081 & 0.001 & 0.009 & 0.081 & \mathbf{0.729} & 0.009 & 0.081 \\ 0.081 & 0.009 & \mathbf{0.729} & 0.081 & 0.009 & 0.001 & 0.081 & 0.009 \\ \mathbf{0.729} & 0.081 & 0.081 & 0.009 & 0.081 & 0.009 & 0.009 & 0.001 \\ 0.009 & 0.001 & 0.081 & 0.009 & 0.081 & 0.009 & \mathbf{0.729} & 0.081 \\ 0.081 & 0.009 & 0.009 & 0.001 & \mathbf{0.729} & 0.081 & 0.081 & 0.009 \end{bmatrix}. \quad (4.116)$$

Suponha-se que se pretende levar o estado da rede para 001. Então, efectuando a optimização num horizonte temporal $T = 5$, com distribuições ideais

$$S(x_{t+1}) = [1 \ 10^4 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1]^T \frac{1}{10007} \quad (4.117)$$

e

$$C(u_t) = [0.5 \ 0.5]^T, \quad (4.118)$$

obtém-se uma sequência c_0, \dots, c_4 óptima. Em particular,

$$c_0(u_t|x_t) = \begin{bmatrix} 0.5000 & 0.5000 & 0.5000 & 0.5000 & 0.9999 & 1.0000 & 0.9922 & 0.9957 \\ 0.5000 & 0.5000 & 0.5000 & 0.5000 & 0.0001 & 0.0000 & 0.0078 & 0.0043 \end{bmatrix}. \quad (4.119)$$

Deste resultado observa-se que para atingir o estado desejado 001, especificado em (4.117), se distinguem duas situações:

- se o estado é 000, 001, 010 ou 011, a presença da substância U é irrelevante, obtendo-se uma probabilidade 0.5 para a sua presença ou ausência.
- se o estado é 100, 101, 110 ou 111, então deve promover-se a ausência de U .

As figuras 4.13(a) e 4.13(b) permitem confirmar este resultado de modo informal.

A figura 4.14 apresenta uma simulação da cadeia fechada durante o intervalo de tempo $t = 0, \dots, 10$, usando o controlo $c_0(u_t|x_t)$ em todos os instantes de tempo. A simulação foi inicializada com o estado inicial da rede a 110 (estado 6 na figura). Observa-se que a evolução da rede seguiu o caminho mais provável

$$110 \rightarrow 100 \rightarrow 101 \rightarrow 001$$

durante os primeiros instantes $t = 0, \dots, 3$. Tendo atingido o estado 001, o controlo gera sinais u_t a 0 (zero) ou 1 (um) com igual probabilidade durante o intervalo $t = 3, \dots, 6$. Em $t = 7$ ocorre uma alteração no estado de activação de um dos genes, passando-se de 001

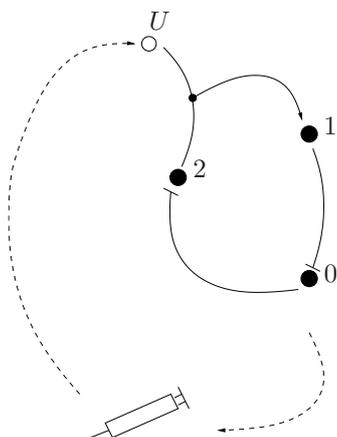


Figura 4.12: Diagrama de uma rede de regulação celular (fictícia).

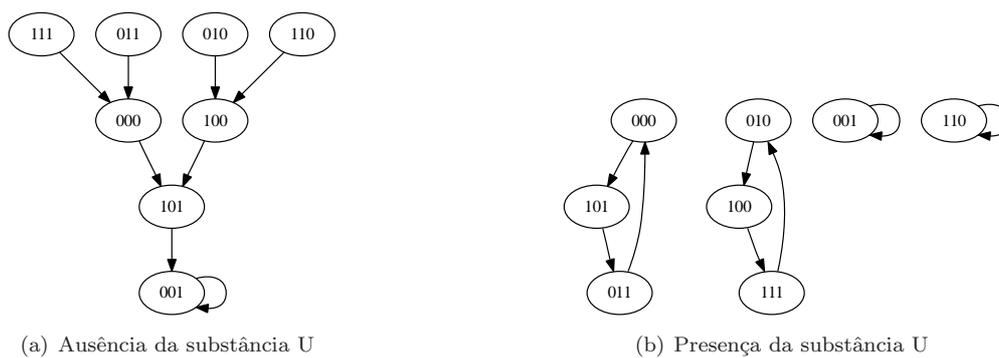


Figura 4.13: Transições de maior probabilidade entre estados da rede.

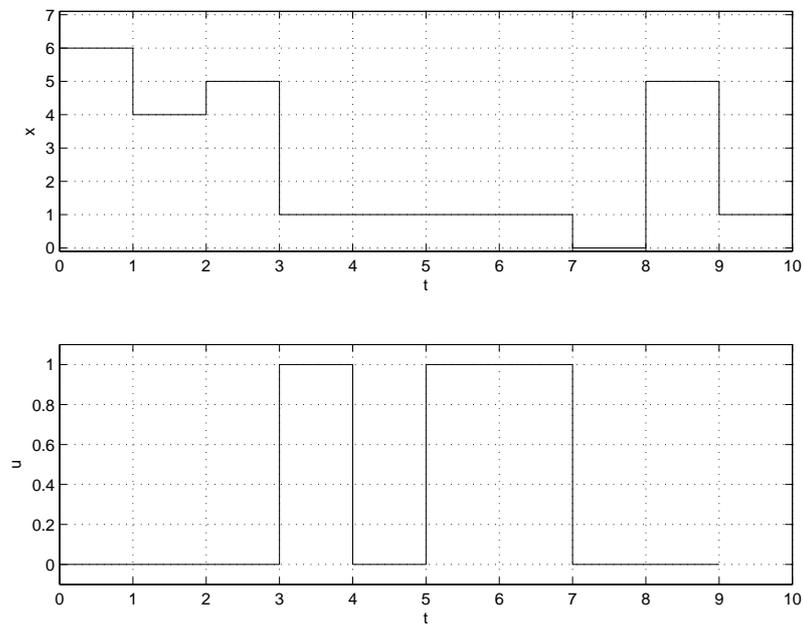


Figura 4.14: Simulação temporal.

para 000. Em consequência o controlo gera o sinal $u_t = 0$, levando novamente o estado da rede para o equilíbrio desejado. A figura 4.15 apresenta o histograma normalizado de x_t e u_t numa simulação estocástica no intervalo $t = 0, \dots, 10^5$. Este exemplo é favorável uma vez que o estado desejado 001 é simples de atingir, conforme se observa na figura 4.12. Outra situação mais difícil seria repetir o mesmo processo para o estado desejado 110. Neste caso não existe um caminho (com grande probabilidade) que leve a rede a esse estado, obtendo-se o histograma da figura 4.16. O resultado está neste caso muito longe do desejado, mas é o óptimo para esta rede.

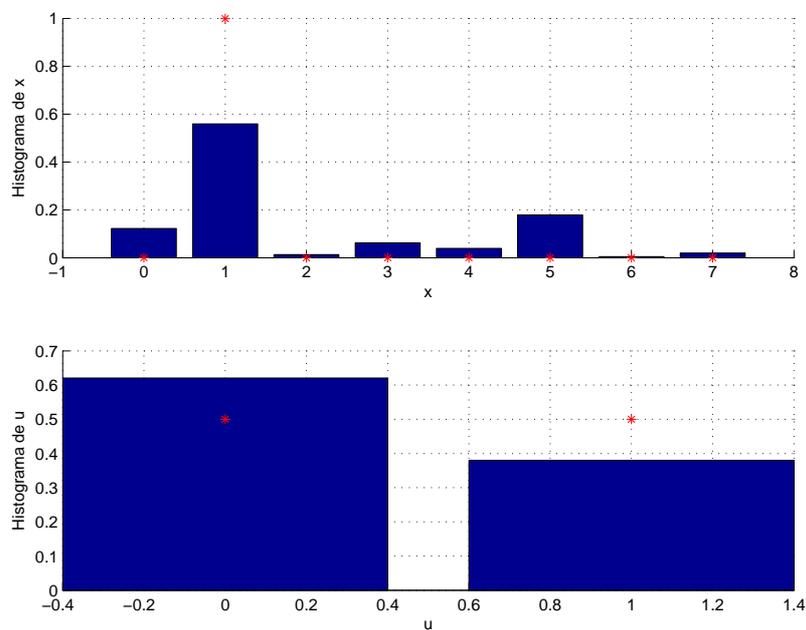


Figura 4.15: Histograma normalizado de x_t e u_t numa simulação estocástica ($T_{\text{sim}} = 10^5$).

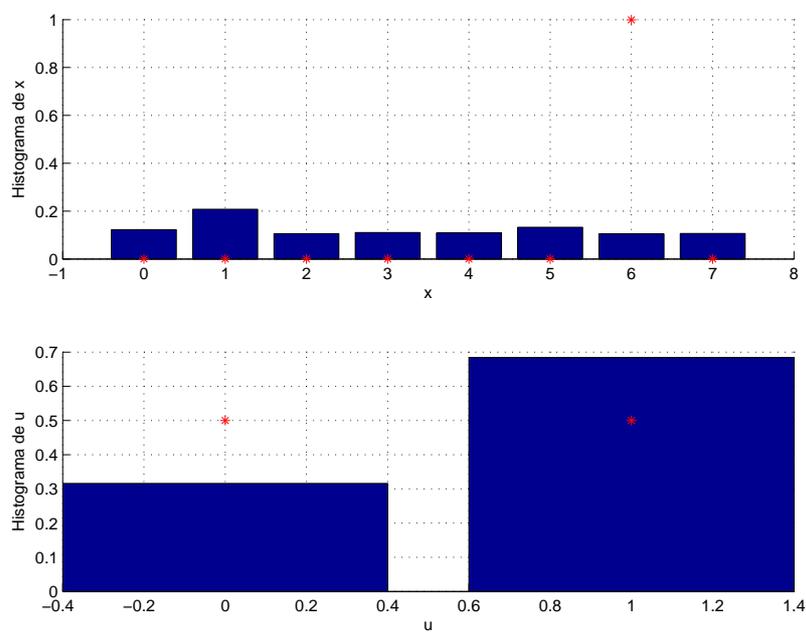


Figura 4.16: Histograma normalizado de x_t e u_t numa simulação estocástica ($T_{\text{sim}} = 10^5$).

Capítulo 5

Conclusão

Nesta tese, o problema do controlo de sistemas estocásticos em cadeia fechada é formulado de forma puramente probabilística, e são propostas soluções para esta formulação. As soluções obtidas são interpretadas geometricamente à luz das ferramentas de geometria da informação. Finalmente são apresentados resultados obtidos em simulação.

As abordagens propostas foram motivadas principalmente pelos trabalhos de Kárný para o controlo probabilístico, e pelos trabalhos de Kulhavý para a estimação recursiva não linear. No primeiro caso, o controlo é formulado como um problema de optimização, em que se pretende aproximar a distribuição de probabilidade da trajectória do sistema em cadeia fechada, a uma distribuição ideal especificada. No segundo caso, pretende-se determinar formulas recursivas para a estimação não linear, mantendo a coerência com a lei de Bayes. Nesta última abordagem são usadas ferramentas de geometria da informação. Estas ferramentas, desenvolvidas por Amari e outros, consistem essencialmente no tratamento de espaços de distribuições de probabilidade como variedades diferenciais.

Neste trabalho a formulação originalmente proposta por Kárný é reformulada de um ponto de vista geométrico. Esta reformulação requer a restrição do problema a classes paramétricas de controladores. Esta restrição tem como consequência soluções sub-óptimas, mas tratáveis, o que não acontecia com as soluções óptimas não-paramétricas.

A formulação paramétrica permite, por um lado, a análise geométrica, ao considerar os parâmetros como sistemas de coordenadas, e por outro, a viabilidade computacional *online* ao permitir uma descrição com memória finita (os parâmetros). Esta viabilidade *online* teve como contrapartida o aparecimento de um problema de optimização, a realizar *offline*, para determinar os parâmetros óptimos.

No caso de sistemas discretos, descritos por cadeias de Markov controláveis, a optimização *offline* é computacionalmente viável e eficiente. Foi desenvolvido um algoritmo baseado no gradiente natural que permitiu a optimização de sistemas de grande dimensão (milhões de parâmetros em poucos segundos em *hardware* actual). A eficiência deste algoritmo resulta do facto do método do gradiente natural aproximar assintoticamente o método de Newton, garantindo convergência quadrática. A grande vantagem do método, quando comparado com o método de Newton, reside no facto de não necessitar da Hessiana do funcional de custo, e portanto não incorrer na penalização da sua complexidade espacial e temporal elevadas. Mostrou-se que o gradiente natural pode ser calculado de forma eficiente com complexidade computacional linear, quer em tempo de cálculo quer em espaço, pelo que é aplicável a problemas de dimensão muito elevada.

Finalmente, consideraram-se vários exemplos para ilustrar os algoritmos desenvolvidos.

5.1 Problemas em aberto e trabalho futuro

Os sistemas dinâmicos, e em particular os sistemas de controlo, vistos numa perspectiva de teoria da informação e usando ferramentas geométricas, são um campo de trabalho para o qual relativamente pouca atenção tem sido dada na literatura. Este trabalho tenta cobrir uma pequena parte deste campo, sendo o caminho seguido apenas um dos muitos possíveis. Do trabalho efectuado ficaram algumas questões em aberto. Outras questões se levantaram, com vários caminhos a perspectivarem-se para trabalho futuro. Em seguida apresenta-se um conjunto de questões, ligados mais ou menos directamente ao trabalho realizado neste tese, e que se propõem como trabalho de investigação a realizar futuramente.

1. O método do gradiente natural apresentado no capítulo 3 foi aplicado a uma parametrização particular de distribuições de probabilidade discretas. Esta parametrização caracteriza a família de distribuições discretas como uma mistura, sendo os resultados posteriores dependentes desta caracterização. Esta parametrização não é, no entanto, única. Outra parametrização possível é, por exemplo, a de uma família exponencial, como se mostrou no capítulo 2. Neste caso, pode mostrar-se que a matriz de informação de Fisher coincide precisamente com a inversa da matriz de informação de Fisher da parametrização mistura. É ainda um trabalho por efectuar determinar a forma do gradiente natural na parametrização exponencial. Em particular, é relevante saber como se comparam os algoritmos do gradiente natural nos casos exponencial e mistura estudado.
2. A extensão para um horizonte temporal infinito resulta, no caso linear, num controlador com ganhos K constantes, obtidos a partir da solução da equação de Riccati algébrica. No caso probabilístico não está provada a convergência da optimização do controlador probabilístico para uma distribuição $c(u|x)$ invariante no tempo.
3. Na secção 4.4.4 propõem-se uma sequência de operações com o objectivo de eliminar os problemas numéricos na implementação da solução óptima de Kárný. Estas baseiam-se no funcional de custo para horizontes finitos, funcional esse que está na origem dos problemas numéricos. Avança-se, como especulação, que a modificação do funcional de custo da divergência K-L para o ritmo de entropia relativa, conforme proposto na extensão para horizonte temporal infinito na secção 4.4.3, poderá resolver os problemas numéricos. A razão é a seguinte: pensa-se que a propagação do *cost to go* seja substituída pela propagação do “ritmo de entropia diferencial”. Sendo esta última convergente, o argumento da exponencial não deverá divergir. Como efeito colateral, poderão surgir as condições de convergência para a distribuição $c(u|x)$, para horizonte infinito.
4. O controlo probabilístico paramétrico para variáveis contínuas requer a solução de uma equação (condição de ortogonalidade) em cada instante do horizonte temporal. Esta equação não tem, em geral, solução explícita pelo que são necessários métodos numéricos para aproximar a solução, à semelhança do que foi efectuado para o caso discreto. É necessário testar em simulação vários destes métodos.

Apêndice A

Apêndice ao capítulo: Geometria da informação

A.1 Reparametrização e métrica de Fisher

Nesta secção mostra-se de que forma uma transformação de coordenadas numa variedade estatística afecta os vários objectos nela definidos, tais como vectores tangentes e produto interno.

Suponha-se conhecida uma família paramétrica \mathcal{F} de distribuições de probabilidade, onde θ representa as coordenadas de cada distribuição $p(x|\theta) \in \mathcal{F}$. Um vector tangente num ponto θ tem representação, na forma exponencial, dada por

$$v = \sum_{i=1}^n v^i \frac{\partial \log p}{\partial \theta^i}, \quad (\text{A.1})$$

onde v^i são as componentes do vector relativamente aos vectores base $\frac{\partial \log p}{\partial \theta^i}$, $i = 1, \dots, n$. Nesta representação, a matriz de informação de Fisher $G(\theta)$ tem componentes $g_{ij}(\theta)$ definidos por

$$g_{ij}(\theta) \stackrel{\text{def}}{=} E_{p(x|\theta)} \left[\frac{\partial \log p}{\partial \theta^i} \frac{\partial \log p}{\partial \theta^j} \right]. \quad (\text{A.2})$$

O produto interno entre dois vectores arbitrários v e w , com componentes v^i e w^i , respectivamente, é

$$\begin{aligned} \langle v, w \rangle &\stackrel{\text{def}}{=} E_{p(x|\theta)} \left[\sum_{i=1}^n v^i \frac{\partial \log p}{\partial \theta^i} w^j \frac{\partial \log p}{\partial \theta^j} \right] \\ &= \sum_{i,j} v^i w^j E_{p(x|\theta)} \left[\frac{\partial \log p}{\partial \theta^i} \frac{\partial \log p}{\partial \theta^j} \right] \\ &= \sum_{i,j} v^i w^j g_{ij}(\theta). \end{aligned} \quad (\text{A.3})$$

Suponha-se que a mesma família de distribuições de probabilidade é parametrizada por outro sistema de coordenadas ξ , tal que $\xi = \varphi(\theta)$. Nestas novas coordenadas, as distribuições de probabilidade são

$$\hat{p}(x|\xi) \stackrel{\text{def}}{=} p(x|\varphi^{-1}(\xi)), \quad (\text{A.4})$$

e os vectores tangentes são

$$\hat{v} = \sum_{m=1}^n \hat{v}^m \frac{\partial \log \hat{p}}{\partial \xi^m}. \quad (\text{A.5})$$

Os vectores tangentes transitam entre os dois sistema de coordenadas por *pushforward*. As componentes dos vectores são transformadas pela matriz Jacobiana da transformação φ , como se ilustra de seguida.

$$\begin{aligned} \sum_i v^i \frac{\partial \log p(x|\theta)}{\partial \theta^i} &= \sum_i v^i \frac{1}{p(x|\theta)} \frac{\partial \hat{p}(x|\varphi(\theta))}{\partial \theta^i} \\ &= \sum_i v^i \frac{1}{\hat{p}(x|\xi)} \sum_m \frac{\partial \hat{p}(x|\xi)}{\partial \xi^m} \frac{\partial \varphi^m(\theta)}{\partial \theta^i} \\ &= \sum_m \underbrace{\left(\sum_i v^i \frac{\partial \varphi^m(\theta)}{\partial \theta^i} \right)}_{\hat{v}^m} \frac{1}{\hat{p}(x|\xi)} \frac{\partial \hat{p}(x|\xi)}{\partial \xi^m} \\ &= \sum_m \hat{v}^m \frac{\partial \log \hat{p}(x|\xi)}{\partial \xi^m}. \end{aligned} \quad (\text{A.6})$$

Em notação matricial, esta transformação pode ser escrita como

$$\hat{\mathbf{v}} = J\mathbf{v}, \quad (\text{A.7})$$

onde $\hat{\mathbf{v}} = [\hat{v}^1 \ \dots \ \hat{v}^n]^T$ e $\mathbf{v} = [v^1 \ \dots \ v^n]^T$ são colunas com as componentes dos vectores \hat{v} e v , respectivamente, e $J = \frac{\partial \varphi}{\partial \theta}$ é a matriz Jacobiana de φ .

A matriz de informação de Fisher transforma-se de forma análoga sob mudança de coordenadas. Para determinar esta transformação, observa-se de (A.6) que

$$\frac{\partial \log p}{\partial \theta^i} = \sum_m \frac{\partial \varphi^m(\theta)}{\partial \theta^i} \frac{\partial \log \hat{p}}{\partial \xi^m}. \quad (\text{A.8})$$

Usando esta identidade, e começando nas coordenadas θ , obtém-se sucessivamente que

$$\begin{aligned} g_{ij}(\theta) &= E_p \left[\frac{\partial \log p}{\partial \theta^i} \frac{\partial \log p}{\partial \theta^j} \right] \\ &= E_p \left[\sum_m \frac{\partial \varphi^m(\theta)}{\partial \theta^i} \frac{\partial \log \hat{p}}{\partial \xi^m} \sum_p \frac{\partial \varphi^p(\theta)}{\partial \theta^j} \frac{\partial \log \hat{p}}{\partial \xi^p} \right] \\ &= \sum_{m,p} \frac{\partial \varphi^m(\theta)}{\partial \theta^i} \frac{\partial \varphi^p(\theta)}{\partial \theta^j} \underbrace{E_{\hat{p}} \left[\frac{\partial \log \hat{p}}{\partial \xi^m} \frac{\partial \log \hat{p}}{\partial \xi^p} \right]}_{\stackrel{\text{def}}{=} \hat{g}_{mp}(\xi)} \\ &= \sum_{m,p} \frac{\partial \varphi^m(\theta)}{\partial \theta^i} \frac{\partial \varphi^p(\theta)}{\partial \theta^j} \hat{g}_{mp}(\xi) \end{aligned} \quad (\text{A.9})$$

Em notação matricial, esta transformação escreve-se

$$G(\theta) = J^T \hat{G}(\xi) J, \quad (\text{A.10})$$

onde G é a matriz de informação nas coordenadas θ e \hat{G} é a matriz de informação nas coordenadas ξ . Mostra-se em seguida que o produto interno definido como em A.2 e (A.3)

é invariante a transformações de coordenadas. De facto, dados dois vectores \hat{v} e \hat{w} nas coordenadas ξ , o seu produto interno calculado em notação matricial é

$$\begin{aligned}
\langle \hat{v}, \hat{w} \rangle_{\hat{g}(\xi)} &= \hat{v}^T \hat{G} \hat{w} \\
&= (J\hat{v})^T \hat{G} (Jw) \\
&= v^T (J^T \hat{G} J) w \\
&= v^T G w \\
&= \langle v, w \rangle_{g(\theta)}.
\end{aligned} \tag{A.11}$$

A.2 Conexão-0 de Amari

Prova-se de seguida que a conexão-0 de Amari é a conexão métrica obtida a partir do tensor métrico g_{ij} .

Demonstração. A conexão métrica é definida pela fórmula

$$\Gamma_{ij,k} = \frac{1}{2} \left(\frac{\partial g_{ik}}{\partial \theta^j} + \frac{\partial g_{jk}}{\partial \theta^i} - \frac{\partial g_{ij}}{\partial \theta^k} \right). \tag{A.12}$$

Calculando as derivadas individualmente, obtém-se que

$$\begin{aligned}
\frac{\partial g_{ik}}{\partial \theta^j} &= \frac{\partial}{\partial \theta^j} \int p \frac{\partial \log p}{\partial \theta^i} \frac{\partial \log p}{\partial \theta^k} \\
&= \int \frac{\partial p}{\partial \theta^j} \frac{\partial \log p}{\partial \theta^i} \frac{\partial \log p}{\partial \theta^k} + p \frac{\partial^2 \log p}{\partial \theta^j \partial \theta^i} \frac{\partial \log p}{\partial \theta^k} + p \frac{\partial \log p}{\partial \theta^i} \frac{\partial^2 \log p}{\partial \theta^j \partial \theta^k} \\
&= \int p \left(\frac{\partial \log p}{\partial \theta^j} \frac{\partial \log p}{\partial \theta^i} \frac{\partial \log p}{\partial \theta^k} + \frac{\partial^2 \log p}{\partial \theta^j \partial \theta^i} \frac{\partial \log p}{\partial \theta^k} + \frac{\partial \log p}{\partial \theta^i} \frac{\partial^2 \log p}{\partial \theta^j \partial \theta^k} \right).
\end{aligned} \tag{A.13}$$

Efectuando as permutações necessárias nos índices i, j, k obtêm-se as restantes derivadas

$$\frac{\partial g_{jk}}{\partial \theta^i} = \int p \left(\frac{\partial \log p}{\partial \theta^i} \frac{\partial \log p}{\partial \theta^j} \frac{\partial \log p}{\partial \theta^k} + \frac{\partial^2 \log p}{\partial \theta^i \partial \theta^j} \frac{\partial \log p}{\partial \theta^k} + \frac{\partial \log p}{\partial \theta^j} \frac{\partial^2 \log p}{\partial \theta^i \partial \theta^k} \right) \tag{A.14}$$

e

$$\frac{\partial g_{ij}}{\partial \theta^k} = \int p \left(\frac{\partial \log p}{\partial \theta^k} \frac{\partial \log p}{\partial \theta^i} \frac{\partial \log p}{\partial \theta^j} + \frac{\partial^2 \log p}{\partial \theta^k \partial \theta^i} \frac{\partial \log p}{\partial \theta^j} + \frac{\partial \log p}{\partial \theta^i} \frac{\partial^2 \log p}{\partial \theta^k \partial \theta^j} \right). \tag{A.15}$$

Finalmente, efectuando a soma (A.12), obtém-se que

$$\begin{aligned}
\Gamma_{ij,k} &= \frac{1}{2} \int p \left(\frac{\partial \log p}{\partial \theta^k} \frac{\partial \log p}{\partial \theta^i} \frac{\partial \log p}{\partial \theta^j} + 2 \frac{\partial^2 \log p}{\partial \theta^i \partial \theta^j} \frac{\partial \log p}{\partial \theta^k} \right) \\
&= E \left[\left(\frac{\partial^2 \log p}{\partial \theta^i \partial \theta^k} + \frac{1}{2} \frac{\partial \log p}{\partial \theta^i} \frac{\partial \log p}{\partial \theta^j} \right) \frac{\partial \log p}{\partial \theta^k} \right].
\end{aligned} \tag{A.16}$$

Este resultado corresponde a fazer $\alpha = 0$ na família de conexões- α de Amari. \square

A.3 Métrica induzida pela divergência K-L

O tensor métrico induzido pela divergência de Kullback-Leibler é

$$\begin{aligned}
\frac{\partial}{\partial \Delta \theta_i} \frac{\partial}{\partial \Delta \theta_j} D(p_{\theta+\Delta \theta} \| p_{\theta}) &= \frac{\partial}{\partial \Delta \theta_i} \frac{\partial}{\partial \Delta \theta_j} \int p_{\theta+\Delta \theta} \log \frac{p_{\theta+\Delta \theta}}{p_{\theta}} \\
&= \frac{\partial}{\partial \Delta \theta_i} \int \frac{\partial p_{\theta+\Delta \theta}}{\partial \Delta \theta_j} \log \frac{p_{\theta+\Delta \theta}}{p_{\theta}} \\
&= \underbrace{\int \frac{\partial^2 p_{\theta+\Delta \theta}}{\partial \Delta \theta_i \partial \Delta \theta_j} \log \frac{p_{\theta+\Delta \theta}}{p_{\theta}}}_{=0 \text{ quando } \Delta \theta \rightarrow 0} + \underbrace{\int p_{\theta+\Delta \theta} \frac{\partial \log p_{\theta+\Delta \theta}}{\partial \Delta \theta_i} \frac{\partial \log p_{\theta+\Delta \theta}}{\partial \Delta \theta_j}}_{=g_{ij}(\theta) \text{ quando } \Delta \theta \rightarrow 0}.
\end{aligned}$$

Tomando o limite quando $\Delta \theta \rightarrow 0$, o primeiro termo anula-se e obtém-se precisamente a métrica de Fisher $g_{ij}(\theta)$.

Da mesma forma, derivando relativamente ao segundo argumento da divergência, obtém-se

$$\begin{aligned}
\frac{\partial}{\partial \Delta \theta_i} \frac{\partial}{\partial \Delta \theta_j} D(p_{\theta} \| p_{\theta+\Delta \theta}) &= \frac{\partial}{\partial \Delta \theta_i} \frac{\partial}{\partial \Delta \theta_j} \int p_{\theta} \log \frac{p_{\theta}}{p_{\theta+\Delta \theta}} \\
&= - \underbrace{\int \frac{p_{\theta}}{p_{\theta+\Delta \theta}} \frac{\partial^2 p_{\theta+\Delta \theta}}{\partial \Delta \theta_i \partial \Delta \theta_j}}_{=0 \text{ quando } \Delta \theta \rightarrow 0} + \underbrace{\int p_{\theta} \frac{\partial \log p_{\theta+\Delta \theta}}{\partial \Delta \theta_i} \frac{\partial \log p_{\theta+\Delta \theta}}{\partial \Delta \theta_j}}_{=g_{ij}(\theta) \text{ quando } \Delta \theta \rightarrow 0}
\end{aligned}$$

onde, tomando o limite $\Delta \theta \rightarrow 0$, obtém-se novamente a métrica de Fisher $g_{ij}(\theta)$.

A.4 Conexão induzida pela divergência K-L

A conexão induzida é

$$\begin{aligned}
\frac{\partial}{\partial \Delta \theta'_k} \left(\frac{\partial}{\partial \Delta \theta_i} \frac{\partial}{\partial \Delta \theta_j} D(p_{\theta+\Delta \theta} \| p_{\theta+\Delta \theta'}) \right) &= \\
&= \frac{\partial}{\partial \Delta \theta'_k} \left(\int \frac{\partial^2 p_{\theta+\Delta \theta}}{\partial \Delta \theta_i \partial \Delta \theta_j} \log \frac{p_{\theta+\Delta \theta}}{p_{\theta+\Delta \theta'}} + g_{ij}(\theta) \right) \\
&= - \int \frac{\partial^2 p_{\theta+\Delta \theta}}{\partial \Delta \theta_i \partial \Delta \theta_j} \frac{\partial \log p_{\theta+\Delta \theta'}}{\partial \Delta \theta'_k} \\
&= - \int p_{\theta+\Delta \theta} \left(\frac{\partial^2 \log p_{\theta+\Delta \theta}}{\partial \Delta \theta_i \partial \Delta \theta_j} + \frac{\partial \log p_{\theta+\Delta \theta}}{\partial \Delta \theta_i} \frac{\partial \log p_{\theta+\Delta \theta}}{\partial \Delta \theta_j} \right) \frac{\partial \log p_{\theta+\Delta \theta'}}{\partial \Delta \theta'_k} \\
&= -E \left[\left(\frac{\partial^2 \log p_{\theta+\Delta \theta}}{\partial \Delta \theta_i \partial \Delta \theta_j} + \frac{\partial \log p_{\theta+\Delta \theta}}{\partial \Delta \theta_i} \frac{\partial \log p_{\theta+\Delta \theta}}{\partial \Delta \theta_j} \right) \frac{\partial \log p_{\theta+\Delta \theta'}}{\partial \Delta \theta'_k} \right]
\end{aligned}$$

que tomando o limite $\Delta \theta, \Delta \theta' \rightarrow 0$ resulta em

$$= -\Gamma_{ij,k}^{(-1)}.$$

A conexão induzida pela dual da divergência de Kullback-Leibler é

$$\begin{aligned} \frac{\partial}{\partial \Delta \theta'_k} \left(\frac{\partial}{\partial \Delta \theta_i} \frac{\partial}{\partial \Delta \theta_j} D(p_{\theta+\Delta\theta'} \| p_{\theta+\Delta\theta}) \right) &= \\ &= -\frac{\partial}{\partial \Delta \theta'_k} \left(\int \frac{p_{\theta+\Delta\theta'}}{p_{\theta+\Delta\theta}} \frac{\partial^2 p_{\theta+\Delta\theta}}{\partial \Delta \theta_i \partial \Delta \theta_j} + \int p_{\theta+\Delta\theta'} \frac{\partial \log p_{\theta+\Delta\theta}}{\partial \Delta \theta_i} \frac{\partial \log p_{\theta+\Delta\theta}}{\partial \Delta \theta_j} \right) \\ &= -\int p_{\theta+\Delta\theta'} \frac{\partial^2 \log p_{\theta+\Delta\theta}}{\partial \Delta \theta_i \partial \Delta \theta_j} \frac{\partial \log p_{\theta+\Delta\theta'}}{\partial \Delta \theta_k} \end{aligned}$$

que fazendo $\Delta \theta_i, \Delta \theta_j \rightarrow 0$ resulta em

$$\begin{aligned} &= E_\theta \left[\frac{\partial^2 \log p_\theta}{\partial \theta_i \partial \theta_j} \frac{\partial \log p_\theta}{\partial \theta_k} \right] \\ &= \Gamma_{ji,k}^{(1)}. \end{aligned}$$

A.5 Relação Pitagórica

Considerem-se as distribuições p , q e r , a curva mistura $m(a) \stackrel{\text{def}}{=} ap + (1-a)r$ e a curva exponencial $e(b) \stackrel{\text{def}}{=} r^{1-b}q^b/\gamma(b)$, onde $\gamma(b)$ é um factor de normalização. Mostra-se em seguida que

$$D(p\|q) = D(p\|r) + D(r\|q) - \left\langle \frac{\partial \log m(a)}{\partial a} \Big|_{a=0}, \frac{\partial \log e(b)}{\partial b} \Big|_{b=0} \right\rangle_r. \quad (\text{A.17})$$

Demonstração.

$$\begin{aligned} D(p\|q) &= \int p \log \frac{p}{q} \\ &= \int p \left(\log \frac{p}{r} + \log \frac{r}{q} \right) \\ &= D(p\|r) + \int p \log \frac{r}{q} \\ &= D(p\|r) + \int (p-r+r) \log \frac{r}{q} \\ &= D(p\|r) + D(r\|q) - \int (p-r) \log \frac{q}{r} \end{aligned} \quad (\text{A.18})$$

Os vectores tangentes às curvas tomados no ponto r são respectivamente

$$\begin{aligned} \frac{\partial \log m(a)}{\partial a} \Big|_{a=0} &= \frac{1}{m(0)} \frac{\partial m(a)}{\partial a} \Big|_{a=0} \\ &= \frac{1}{r} (p-r) \end{aligned} \quad (\text{A.19})$$

e

$$\begin{aligned} \frac{\partial \log e(b)}{\partial b} \Big|_{b=0} &= \frac{\partial}{\partial b} \Big|_{b=0} \exp \left((1-b) \log r + b \log q - \log \gamma(b) \right) \\ &= \log \frac{q}{r} - \underbrace{\frac{\partial \log \gamma(b)}{\partial b} \Big|_{b=0}}_{k \text{ (const.)}}. \end{aligned} \quad (\text{A.20})$$

O produto interno entre estes dois vectores é dado por

$$\begin{aligned}
\left\langle \frac{\partial \log m(a)}{\partial a} \Big|_{a=0}, \frac{\partial \log e(b)}{\partial b} \Big|_{a=0} \right\rangle_r &= \int r \frac{\partial \log m(a)}{\partial a} \Big|_{a=0} \frac{\partial \log e(b)}{\partial b} \Big|_{a=0} \\
&= \int r \frac{1}{r} (p-r) \left(\log \frac{q}{r} - k \right) \\
&= \int (p-r) \log \frac{q}{r} - k \underbrace{\int (p-r)}_{=0}.
\end{aligned} \tag{A.21}$$

que, substituindo em (A.18), resulta na equação pretendida. \square

A.6 Projectção ortogonal sobre famílias mistura

Uma família mistura \mathcal{M} tem como elementos distribuições obtidas por combinação de um conjunto de $n+1$ distribuições $p_0(x), p_1(x), \dots, p_n(x)$ da seguinte forma:

$$p(x) = \sum_{i=1}^n \theta_i p_i(x) + \left(1 - \sum_{i=1}^n \theta_i\right) p_0(x), \tag{A.22}$$

com $0 \leq \theta_i \leq 1$ e $\sum_{i=1}^n \theta_i \leq 1$.

Considere-se a família mistura \mathcal{M} anterior e uma distribuição $q \notin \mathcal{M}$. Então, a projectção ortogonal de q em \mathcal{M} é o ponto p^* tal que o vector tangente à curva exponencial, no ponto p^* , é ortogonal ao espaço tangente. Esta projectção determina o ponto $p \in \mathcal{M}$ que minimiza $D(p||q)$.

Demonstração. O mínimo de $D(p||q)$ é obtido pela solução da equação

$$\frac{\partial D(p||q)}{\partial \theta_i} = 0, \quad i = 1, \dots, n, \tag{A.23}$$

donde se obtém que

$$\begin{aligned}
\frac{\partial D}{\partial \theta_i} &= \frac{\partial}{\partial \theta_i} \int p \log \frac{p}{q} \\
&= \int \frac{\partial p}{\partial \theta_i} \log \frac{p}{q} + \frac{\partial p}{\partial \theta_i} \\
&= \int \frac{\partial p}{\partial \theta_i} \log \frac{p}{q} \\
&= \int p \frac{\partial \log p}{\partial \theta_i} \log \frac{p}{q}.
\end{aligned} \tag{A.24}$$

Tendo em consideração que $\frac{\partial \log p}{\partial \theta_i}$ são os vectores da base do espaço tangente, e que o vector tangente à curva exponencial $e(a) \stackrel{\text{def}}{=} p^{1-a} q^a / \gamma(a)$ tomado no ponto p é

$$\frac{\partial \log e(a)}{\partial a} \Big|_{a=0} = \log \frac{q}{p} + D(p||q), \tag{A.25}$$

obtém-se que

$$\begin{aligned}
\frac{\partial D}{\partial \theta_i} &= \int p \frac{\partial \log p}{\partial \theta_i} \log \frac{p}{q} \\
&= \int p \frac{\partial \log p}{\partial \theta_i} \left(\frac{\partial \log e(a)}{\partial a} \Big|_{a=0} - D(p||q) \right) \\
&= \int p \frac{\partial \log p}{\partial \theta_i} \frac{\partial \log e(a)}{\partial a} \Big|_{a=0} - \underbrace{\int p \frac{\partial \log p}{\partial \theta_i} D(p||q)}_{=0} \\
&= \left\langle \frac{\partial \log p}{\partial \theta_i}, \frac{\partial \log e(a)}{\partial a} \Big|_{a=0} \right\rangle_p,
\end{aligned} \tag{A.26}$$

e, portanto, o mínimo de $D(p||q)$ é atingido para o ponto p tal que, o vector tangente à curva exponencial é ortogonal a todos os vectores da base do espaço tangente. \square

Apêndice B

Apêndice ao capítulo: Otimização de distribuições de probabilidade

B.1 Sobre a definição de gradiente natural

Considere-se o funcional

$$L(v, \lambda) \stackrel{\text{def}}{=} dF(v) + \lambda(1 - \langle v, v \rangle). \quad (\text{B.1})$$

Os pontos de estacionariedade são as soluções da equação $dL(v, \lambda) = \mathbf{0}$. Resolvendo esta equação, obtém-se que

$$\begin{aligned} dL(v, \lambda) &= \begin{bmatrix} \frac{\partial L}{\partial v} & \frac{\partial L}{\partial \lambda} \end{bmatrix} \\ &= \begin{bmatrix} \frac{\partial}{\partial v} dF(v) - \lambda \frac{\partial}{\partial v} \langle v, v \rangle & 1 - \langle v, v \rangle \end{bmatrix} \\ &= \begin{bmatrix} dF - \lambda v^T G & 1 - \langle v, v \rangle \end{bmatrix} \\ &= \begin{bmatrix} 0 & 0 \end{bmatrix}. \end{aligned} \quad (\text{B.2})$$

Ao igualar cada componente a zero, obtém-se que

$$\lambda v^T G = dF, \quad (\text{B.3})$$

e

$$\langle v, v \rangle = 1. \quad (\text{B.4})$$

Multiplicando a equação (B.3) à direita por um vector arbitrário w , e definindo o vector gradiente como $\nabla F \stackrel{\text{def}}{=} \lambda v$, obtém-se

$$\nabla F^T G w = dF w, \quad (\text{B.5})$$

que surge escrito habitualmente na literatura como

$$\langle \nabla F, w \rangle = dF(w), \quad \forall w \neq \mathbf{0}. \quad (\text{B.6})$$

A equação (B.3) mostra também como determinar explicitamente o gradiente. Tendo em consideração que $G = G^T$, a transposição da equação resulta em

$$G\nabla F = dF^T, \quad (\text{B.7})$$

donde se obtém

$$\nabla F = G^{-1} \begin{bmatrix} \frac{\partial F}{\partial \theta^1} \\ \vdots \\ \frac{\partial F}{\partial \theta^n} \end{bmatrix}. \quad (\text{B.8})$$

B.2 Amplificação de uma direcção numa transformação linear

Para observar este efeito, decompõem-se o vector x como a soma de uma componente na direcção $\sqrt{\theta}$ com o respectivo complemento ortogonal z , *i.e.*,

$$x = \alpha\sqrt{\theta} + z, \quad (\text{B.9})$$

onde

$$\alpha \stackrel{\text{def}}{=} \frac{\sqrt{\theta}^T x}{\|\sqrt{\theta}\|^2}, \quad z \stackrel{\text{def}}{=} x - \alpha\sqrt{\theta}. \quad (\text{B.10})$$

Então,

$$\begin{aligned} \mathbf{H}_0 x &= \left(\mathbf{I} - (1 - \theta^0) \frac{\sqrt{\theta}}{\|\sqrt{\theta}\|} \frac{\sqrt{\theta}^T}{\|\sqrt{\theta}\|} \right) x \\ &= \left(\mathbf{I} + (\theta^0 - 1) \frac{\sqrt{\theta}}{\|\sqrt{\theta}\|} \frac{\sqrt{\theta}^T}{\|\sqrt{\theta}\|} \right) (\alpha\sqrt{\theta} + z) \\ &= (\alpha\sqrt{\theta} + z) + (\theta^0 - 1)\alpha\sqrt{\theta} \overbrace{\frac{\sqrt{\theta}^T \sqrt{\theta}}{\|\sqrt{\theta}\|^2}}^{=1} + \\ &\quad + (\theta^0 - 1) \frac{\sqrt{\theta}}{\|\sqrt{\theta}\|^2} \underbrace{\sqrt{\theta}^T z}_{=0} \\ &= (\theta^0 \alpha)\sqrt{\theta} + z, \end{aligned} \quad (\text{B.11})$$

de onde se observa que a amplificação θ^0 afecta apenas a componente de x que está sobre a direcção $\sqrt{\theta}$.

Se $\theta^0 = 0$, então a componente sobre a direcção $\sqrt{\theta}$ é anulada. Nesse caso, a transformação \mathbf{H}_0 efectua uma projecção sobre o complemento ortogonal de $\sqrt{\theta}$. Este tipo de transformações é usualmente construído como $\mathbf{P} \stackrel{\text{def}}{=} \mathbf{I} - \mathbf{u}\mathbf{u}^T$, onde \mathbf{u} é um vector unitário. Outra transformação relacionada é a transformação de Householder definida por $\mathbf{H} \stackrel{\text{def}}{=} \mathbf{I} - 2\mathbf{u}\mathbf{u}^T$. Esta transformação efectua uma reflexão na direcção do vector \mathbf{u} .

B.3 Intervalos admissíveis para o passo η no método do gradiente natural

Considere-se a equação de actualização dos parâmetros no método do gradiente natural

$$\theta_{\tau+1} = \theta_{\tau} - \eta_{\tau} \tilde{\nabla} F. \quad (\text{B.12})$$

Para η positivo ou negativo esta equação corresponde, respectivamente, aos procedimentos de minimização ou maximização. Pretende-se determinar um intervalo I_{τ} tal que se $\eta_{\tau} \in I_{\tau}$ então os parâmetros actualizados $\theta_{\tau+1}$ são admissíveis. Assim, deve garantir-se em cada passo do algoritmo que

$$\theta_{\tau+1}^i > 0, \quad i = 1, \dots, n, \quad (\text{B.13})$$

$$\sum_{i=1}^n \theta_{\tau+1}^i < 1. \quad (\text{B.14})$$

Obtém-se para a restrição (B.13) que

$$\theta_{\tau}^i - \eta_{\tau} (\tilde{\nabla} F)^i > 0 \quad (\text{pela equação (B.12)}) \quad (\text{B.15})$$

$$\Leftrightarrow \eta_{\tau} \left(G^{-1} \nabla F \right)^i < \theta^i \quad (\text{pela definição de } \tilde{\nabla} F) \quad (\text{B.16})$$

$$\Leftrightarrow \eta_{\tau} \left(\theta^i (\nabla F)^i - \theta^i (\theta \cdot \nabla F) \right) < \theta^i \quad (\text{pela fórmula rápida}) \quad (\text{B.17})$$

$$\Leftrightarrow \eta_{\tau} \left((\nabla F)^i - \theta \cdot \nabla F \right) < 1 \quad (\text{porque } \theta_{\tau}^i > 0) \quad (\text{B.18})$$

de onde se obtém

$$\Rightarrow \eta_{\tau} < \frac{1}{(\nabla F)^i - \theta \cdot \nabla F}, \quad \text{se } (\nabla F)^i > \theta \cdot \nabla F, \quad (\text{B.19})$$

$$\text{e } \eta_{\tau} > \frac{1}{(\nabla F)^i - \theta \cdot \nabla F}, \quad \text{se } (\nabla F)^i < \theta \cdot \nabla F. \quad (\text{B.20})$$

A condição (B.20) é irrelevante em problemas de minimização, uma vez que nestes $\eta > 0$. Por outro lado, a condição (B.19) deve ser satisfeita para todos os índices i para os quais $(\nabla F)^i > \theta \cdot \nabla F$. Ou seja,

$$\eta_{\tau} < \min_i \frac{1}{(\nabla F)^i - \theta \cdot \nabla F} = \frac{1}{\max_i (\nabla F)^i - \theta \cdot \nabla F}. \quad (\text{B.21})$$

Em problemas de maximização, onde $\eta < 0$, a condição (B.19) é irrelevante, enquanto a equação (B.20) deve ser satisfeita para todos os índices tais que $(\nabla F)^i < \theta \cdot \nabla F$. Ou seja,

$$\eta_{\tau} > \max_i \frac{1}{(\nabla F)^i - \theta \cdot \nabla F} = \frac{1}{\min_i (\nabla F)^i - \theta \cdot \nabla F}. \quad (\text{B.22})$$

As condições (B.19) e (B.20) garantem a restrição (B.13). Para satisfazer a restrição (B.14) procede-se de forma análoga. Obtém-se neste caso

$$\sum_{i=1}^n \left(\theta^i - \eta (G^{-1} \nabla F)^i \right) < 1 \quad (\text{pela equação (B.12)}) \quad (\text{B.23})$$

$$\Leftrightarrow \eta \sum_{i=1}^n (G^{-1} \nabla F)^i > \sum_{i=1}^n \theta^i - 1 \quad (\text{B.24})$$

$$\Leftrightarrow \eta \sum_{i=1}^n \left(\theta^i (\nabla F)^i - \theta^i (\theta \cdot \nabla F) \right) > \sum_{i=1}^n \theta^i - 1 \quad (\text{pela fórmula rápida}) \quad (\text{B.25})$$

$$\Leftrightarrow \eta \underbrace{\sum_{i=1}^n \theta^i (\nabla F)^i}_{=\theta \cdot \nabla F} - \eta \sum_{i=1}^n \theta^i (\theta \cdot \nabla F) > \sum_{i=1}^n \theta^i - 1 \quad (\text{B.26})$$

$$\Leftrightarrow \eta (\theta \cdot \nabla F) \left(1 - \sum_{i=1}^n \theta^i \right) > - \left(1 - \sum_{i=1}^n \theta^i \right) \quad (\text{B.27})$$

$$\Leftrightarrow \eta (\theta \cdot \nabla F) > -1 \quad (\text{porque } \sum_i \theta^i < 1) \quad (\text{B.28})$$

de onde se obtém

$$\Rightarrow \eta < \frac{1}{-\theta \cdot \nabla F}, \quad \text{se } \theta \cdot \nabla F < 0, \quad (\text{B.29})$$

$$\text{e } \eta > \frac{1}{-\theta \cdot \nabla F}, \quad \text{se } \theta \cdot \nabla F > 0. \quad (\text{B.30})$$

A condição (B.30) é irrelevante em problemas de minimização, enquanto a (B.29) é irrelevante em problemas de maximização.

A condição (B.19) pode ser combinada com a condição (B.29) da seguinte maneira:

$$\eta < \frac{1}{\max \alpha - \theta \cdot \nabla F}, \quad \alpha \stackrel{\text{def}}{=} \{0, (\nabla F)^1, \dots, (\nabla F)^n\}. \quad (\text{B.31})$$

Da mesma forma, combinando as condições (B.20) e (B.30), obtém-se:

$$\eta > \frac{1}{\min \alpha - \theta \cdot \nabla F}. \quad (\text{B.32})$$

As restrições no passo podem ser escritas na forma de um intervalo:

$$\eta \in \left] \frac{1}{\min \alpha - \theta \cdot \nabla F}, \frac{1}{\max \alpha - \theta \cdot \nabla F} \right[, \quad (\text{B.33})$$

com

$$\alpha \stackrel{\text{def}}{=} \{0, (\nabla F)^1, \dots, (\nabla F)^n\}. \quad (\text{B.34})$$

B.4 Notas sobre distribuições e vectores não normalizados

Uma distribuição de probabilidade normalizada $p(x)$ é definida de modo a satisfazer

$$\int p(x) dx = 1. \quad (\text{B.35})$$

Esta restrição permite definir uma classe de equivalência de distribuições da seguinte maneira. Duas distribuições p e \tilde{p} dizem-se equivalentes, $p \sim \tilde{p}$, se $p = \tilde{p}/a$ para alguma constante $a > 0$.

Dada uma distribuição não normalizada \tilde{p} , é possível obter a distribuição normalizada equivalente p fazendo

$$p(x) = \frac{\tilde{p}(x)}{\int \tilde{p}(x) dx}. \quad (\text{B.36})$$

De forma mais geral, um modelo paramétrico não normalizado $\tilde{p}(x|\theta)$ pode ser normalizado dividindo \tilde{p} pela função $a(\theta)$:

$$p(x|\theta) = \frac{\tilde{p}(x|\theta)}{a(\theta)}, \quad a(\theta) = \int \tilde{p}(x|\theta) dx \quad (\text{B.37})$$

A utilização de distribuições não normalizadas pode ser estendida para vectores tangentes não normalizados. Se $\tilde{p}(x|\theta)$ é uma distribuição não normalizada, os vectores tangentes não normalizados são

$$\frac{\partial \log \tilde{p}(x|\theta)}{\partial \theta^i} \quad (\text{B.38})$$

Para obter o vector tangente normalizado que lhe corresponde, pode decompor-se \tilde{p} no produto $a(\theta)p(x|\theta)$, obtendo-se

$$\begin{aligned} \frac{\partial \log \tilde{p}(x|\theta)}{\partial \theta^i} &= \frac{\partial \log (a(\theta)p(x|\theta))}{\partial \theta^i} \\ &= \frac{\partial \log a(\theta)}{\partial \theta^i} + \frac{\partial \log p(x|\theta)}{\partial \theta^i} \\ &= b(\theta) + \frac{\partial \log p(x|\theta)}{\partial \theta^i} \end{aligned} \quad (\text{B.39})$$

Como se sabe que o valor esperado de um vector tangente é nulo,

$$E_\theta \left[\frac{\partial \log p(x|\theta)}{\partial \theta^i} \right] = 0, \quad (\text{B.40})$$

então o valor esperado de um vector tangente não normalizado é

$$E_\theta \left[\frac{\partial \log \tilde{p}(x|\theta)}{\partial \theta^i} \right] = b(\theta). \quad (\text{B.41})$$

Dado um vector não normalizado a sua normalização pode ser efectuada da seguinte maneira:

$$\frac{\partial \log p(x|\theta)}{\partial \theta^i} = \frac{\partial \log \tilde{p}(x|\theta)}{\partial \theta^i} - E_\theta \left[\frac{\partial \log \tilde{p}(x|\theta)}{\partial \theta^i} \right]. \quad (\text{B.42})$$

Assim, a classe de equivalência das distribuições induz uma classe de equivalência de espaços tangentes. Dois vectores tangentes são equivalentes se a sua diferença é uma função não dependente das variáveis aleatórias.

Em problemas de optimização da divergência de Kullback-Leibler $D(p||q)$, libertando a restrição de normalização de p e substituindo-a por uma não normalizada \tilde{p} , não resulta em geral na mesma distribuição óptima. Derivando a divergência $D(\tilde{p}_\theta||q)$ relativamente aos parâmetros, obtém-se

$$\frac{\partial D(\tilde{p}_\theta||q)}{\partial \theta^i} = \frac{\partial a(\theta)}{\partial \theta^i} (D(p||q) + \log a + 1) + a \frac{\partial D(p_\theta||q)}{\partial \theta^i} \quad (\text{B.43})$$

donde se conclui que o mínimo de $D(\tilde{p}_\theta \| q)$ coincide com o mínimo de $D(p_\theta \| q)$ apenas quando $a(\theta) = a$ é uma constante. Como consequência, qualquer procedimento de optimização iterativo deve ser efectuado com a constante, ou seja, a actualização da distribuição de probabilidade não deve alterar o valor de $\int \tilde{p}(x) dx$.

Apêndice C

Apêndice ao capítulo: Controle probabilístico

C.1 Controle probabilístico paramétrico

A demonstração consiste em calcular sucessivamente as derivadas parciais

$$\frac{\partial J}{\partial \theta_{T-1}}, \frac{\partial J}{\partial \theta_{T-2}}, \dots, \frac{\partial J}{\partial \theta_0} \quad (\text{C.1})$$

usando para o efeito (4.69). Obtém-se para o instante $T - 1$ que

$$\begin{aligned} \frac{\partial J}{\partial \theta_{T-1}} &= \frac{\partial}{\partial \theta_{T-1}} E_{p(x_{T-1})} [D(p(x_T, u_{T-1}|x_{T-1})||q(x_T, u_{T-1}|x_{T-1}))] \\ &= E_{p(x_{T-1})} \left[\frac{\partial}{\partial \theta_{T-1}} \int s_T c_{T-1} \log \frac{s_T c_{T-1}}{S_T C_{T-1}} dx_T du_{T-1} \right] \\ &= E_{p(x_{T-1})} \left[\int s_T \frac{\partial c_{T-1}}{\partial \theta_{T-1}} \log \frac{s_T c_{T-1}}{S_T C_{T-1}} dx_T du_{T-1} \right] \\ &= E_{p(x_{T-1})} \left[E_{c_{T-1}} \left[\left(D(s_T||S_T) + \log \frac{c_{T-1}}{C_{T-1}} \right) \frac{\partial \log c_{T-1}}{\partial \theta_{T-1}} \right] \right]. \end{aligned} \quad (\text{C.2})$$

Esta expressão anula-se para todo o $p(x_{T-1})$ se

$$E_{c_{T-1}} \left[\left(D(s_T||S_T) + \log \frac{c_{T-1}}{C_{T-1}} \right) \frac{\partial \log c_{T-1}}{\partial \theta_{T-1}} \right] = 0. \quad (\text{C.3})$$

Prosseguindo para $t = T - 2$, obtém-se

$$\frac{\partial J_{0:T}}{\partial \theta_{T-2}} = \frac{\partial}{\partial \theta_{T-2}} E_{p(x_{T-2})} [D(p(-|x_{T-2})||q(-|x_{T-2}))] \quad (\text{C.4})$$

onde o traço denota as v.a. até ao final do horizonte, neste caso,

$$\begin{aligned} p(-|x_{T-2}) &= p(x_{T-1:T}, u_{T-2:T-1}|x_{T-2}) \\ &= s_{T-1} c_{T-2} p(x_T, u_{T-1}|x_{T-1}) \end{aligned} \quad (\text{C.5})$$

e

$$\begin{aligned} q(-|x_{T-2}) &= q(x_{T-1:T}, u_{T-2:T-1}|x_{T-2}) \\ &= S_{T-1} C_{T-2} q(x_T, u_{T-1}|x_{T-1}). \end{aligned} \quad (\text{C.6})$$

Calculando a derivada parcial da expressão (C.4), vem

$$\begin{aligned}
\frac{\partial J_{0:T}}{\partial \theta_{T-2}} &= E_{p(x_{T-2})} \left[\frac{\partial}{\partial \theta_{T-2}} \int s_{T-1} c_{T-2} p(-|x_{T-1}) \log \frac{s_{T-1} c_{T-2} p(-|x_{T-1})}{S_{T-1} C_{T-2} q(-|x_{T-1})} \right] \\
&= E_{p(x_{T-2})} \left[\int \frac{\partial c_{T-2}}{\partial \theta_{T-2}} \int s_{T-1} p(-|x_{T-1}) \left(\log \frac{s_{T-1}}{S_{T-1}} + \log \frac{c_{T-2}}{C_{T-2}} + \right. \right. \\
&\quad \left. \left. + \log \frac{p(-|x_{T-1})}{q(-|x_{T-1})} \right) dx_T du_{T-1} dx_{T-1} du_{T-2} \right] \\
&= E_{p(x_{T-2})} \left[\int \frac{\partial c_{T-2}}{\partial \theta_{T-2}} \left(\int s_{T-1} \log \frac{s_{T-1}}{S_{T-1}} dx_{T-1} + \log \frac{c_{T-2}}{C_{T-2}} + \right. \right. \\
&\quad \left. \left. + \int s_{T-1} \int p(-|x_{T-1}) \log \frac{p(-|x_{T-1})}{q(-|x_{T-1})} dx_T du_{T-1} dx_{T-1} \right) du_{T-2} \right] \quad (C.7) \\
&= E_{p(x_{T-2})} \left[E_{c_{T-2}} \left[\left(D(s_{T-1} \| S_{T-1}) + \log \frac{c_{T-2}}{C_{T-2}} + \right. \right. \right. \\
&\quad \left. \left. + E_{s_{T-1}} [D(p(-|x_{T-1}) \| q(-|x_{T-1}))] \right) \frac{\partial \log c_{T-2}}{\partial \theta_{T-2}} \right] \right]
\end{aligned}$$

Prosseguindo de igual modo para os restantes instantes de tempo t , obtém-se que

$$\begin{aligned}
\frac{\partial J_{0:T}}{\partial \theta_t} &= E_{p(x_t)} \left[E_{c_t} \left[\left(D(s_{t+1} \| S_{t+1}) + \log \frac{c_t}{C_t} + \right. \right. \right. \\
&\quad \left. \left. + E_{s_{t+1}} [D(p(-|x_{t+1}) \| q(-|x_{t+1}))] \right) \frac{\partial \log c_{T-2}}{\partial \theta_{T-2}} \right] \right]. \quad (C.8)
\end{aligned}$$

que igualando a zero, resulta na equação

$$E_{c_t} \left[\left(D(s_{t+1} \| S_{t+1}) + \log \frac{c_t}{C_t} + E_{s_{t+1}} [D(p(-|x_{t+1}) \| q(-|x_{t+1}))] \right) \frac{\partial \log c_{T-2}}{\partial \theta_{T-2}} \right] = 0. \quad (C.9)$$

Definindo

$$-\log \gamma_{t+1}(x_{t+1}) \stackrel{\text{def}}{=} D(p(-|x_{t+1}) \| q(-|x_{t+1})) \quad (C.10)$$

e

$$\gamma_T(x_T) = 1 \quad (C.11)$$

obtém-se a equação (4.71) do algoritmo. A equação (4.72) consiste em observar que

$$\begin{aligned}
-\log \gamma_t(x_t) &= D(p(-|x_t) \| q(-|x_t)) \\
&= \int s_{t+1} c_t p(-|x_{t+1}) \log \frac{s_{t+1} c_t p(-|x_{t+1})}{S_{t+1} C_t q(-|x_{t+1})} dx_{t+1:T} du_{t:T-1} \\
&= \int c_t \left(\int s_{t+1} \log \frac{s_{t+1}}{S_{t+1}} dx_{t+1} + \log \frac{c_t}{C_t} + \right. \\
&\quad \left. + \int s_{t+1} \left(\int p(-|x_{t+1}) \log \frac{p(-|x_{t+1})}{q(-|x_{t+1})} dx_{t+1} \right) du_t \right) \\
&= E_{c_t} \left[D(s_{t+1} \| S_{t+1}) + \log \frac{c_t}{C_t} + E_{s_{t+1}} [-\log \gamma_{t+1}(x_{t+1})] \right]. \quad (C.12)
\end{aligned}$$

Finalmente, admitindo $q(x_0) = p(x_0)$, verifica-se que o custo total é

$$J_{0:T} = E_{p(x_0)} [D(p(-|x_0) \| q(-|x_0))] = E_{p(x_0)} [-\log \gamma_0(x_0)]. \quad (C.13)$$

C.2 Solução da formulação linear Gaussiana paramétrica

Considera-se o sistema “real” probabilístico e o controlo definidos por

$$s(x_{t+1}|x_t, u_t) = \mathcal{N}(Ax_t + Bu_t, Q_t), \quad (\text{C.14})$$

$$c(u_t|x_t) = \mathcal{N}(K_t x_t, R_t), \quad (\text{C.15})$$

e as distribuições ideais de média nula

$$S(x_{t+1}) = \mathcal{N}(0, \bar{Q}_t), \quad (\text{C.16})$$

$$C(u_t) = \mathcal{N}(0, \bar{R}_t). \quad (\text{C.17})$$

Pretende-se minimizar a divergência de K-L entre a distribuição real e ideal $D(p_{0:T}||q_{0:T})$. Para o efeito considera-se a família de controladores parametrizada pela matriz de ganhos K_t , e aplica-se a solução obtida no algoritmo do teorema 10.

Para obter a solução da condição de optimalidade é necessário obter expressões para os termos envolvidos nas expressões (4.71) e (4.72). Assim,

$$\log \frac{c}{C} = -\frac{1}{2}(u - Kx)^T R^{-1}(u - Kx) + \frac{1}{2}u^T \bar{R}^{-1}u + \frac{1}{2} \log \frac{|\bar{R}|}{|R|}, \quad (\text{C.18})$$

$$D(s||S) = \frac{1}{2}(Ax + Bu)^T \bar{Q}^{-1}(Ax + Bu) + \frac{1}{2} \log \frac{|\bar{Q}|}{|Q|} + \frac{1}{2} \text{Tr}(\bar{Q}^{-1}Q) - \frac{N}{2}, \quad (\text{C.19})$$

$$\frac{\partial \log c}{\partial K} = x(u - Kx)^T R^{-1}. \quad (\text{C.20})$$

Os índices temporais são omitidos das expressões para maior clareza, embora se deva ter em atenção que eles estão presentes. Notando que a condição de optimalidade resulta em

$$E_c \left[\underbrace{\left(D(s||S) + \log \frac{c}{C} + E_s[-\log \gamma] \right)}_{\text{factor escalar}} x(u - Kx)^T R^{-1} \right] = 0 \quad (\text{C.21})$$

$$\Leftrightarrow x E_c \left[(\dots)(u - Kx)^T \right] R^{-1} = 0 \quad (\text{C.22})$$

$$\Leftrightarrow E_c \left[(\dots)(u - Kx)^T \right] = 0, \quad (\text{C.23})$$

e admitindo que o factor escalar se pode escrever como um polinómio em $u - Kx$, então separando a expressão em duas parcelas, uma para os termos de ordem par (t.o.p.) e outra para os de ordem ímpar, obtém-se a equação

$$E_c \left[(\text{termos de ordem ímpar})(u - Kx)^T \right] + \underbrace{E_c \left[(\text{termos de ordem par})(u - Kx)^T \right]}_{=0} = 0. \quad (\text{C.24})$$

Os termos de ordem par são anulados por um argumento de simetria/antisimetria da distribuição gaussiana e do argumento do valor esperado. A condição de optimalidade reduz-se ao termo da esquerda de (C.24). Determina-se em seguida esse termo.

Da expressão (C.18) apenas a parcela $u^T \bar{R}^{-1}u$ contribui, parcialmente, com termos de ordem ímpar. Escrevendo

$$\begin{aligned} \frac{1}{2}u^T \bar{R}^{-1}u &= \frac{1}{2}(u - Kx + Kx)^T \bar{R}^{-1}(u - Kx + Kx) \\ &= \frac{1}{2}(u - Kx)^T \bar{R}^{-1}(u - Kx) + \underbrace{\frac{1}{2}(Kx)^T \bar{R}^{-1}Kx + (u - Kx)^T \bar{R}^{-1}Kx}_{\text{termos de ordem par (a eliminar)}} \end{aligned} \quad (\text{C.25})$$

Da expressão (C.19) apenas a fórmula quadrática contribui com termos de ordem ímpar. Escrevendo

$$\begin{aligned} \frac{1}{2}(Ax + Bu)^T \bar{Q}^{-1}(Ax + Bu) &= \\ &= \frac{1}{2}u^T B^T \bar{Q}^{-1}Bu + u^T B^T \bar{Q}^{-1}Ax + \frac{1}{2}x^T A^T \bar{Q}^{-1}Ax \\ &= (u - Kx)^T B^T \bar{Q}^{-1}BKx + (u - Kx)^T B^T \bar{Q}^{-1}Ax + \text{t.o.p.} \end{aligned} \quad (\text{C.26})$$

Reunindo os termos de ordem ímpar de (C.25) e (C.26) na condição de optimalidade, obtém-se

$$E_c \left[\underbrace{(u - Kx)^T \left(B^T \bar{Q}^{-1}A + (B^T \bar{Q}^{-1}B + \bar{R}^{-1})K \right) x}_{\text{factor escalar}} (u - Kx)^T R^{-1} \right] = 0. \quad (\text{C.27})$$

Esta equação tem como solução a matriz dos ganhos

$$K = -(B^T \bar{Q}^{-1}B + \bar{R}^{-1})^{-1} B^T \bar{Q}^{-1}A. \quad (\text{C.28})$$

Calcula-se em seguida a expressão (4.72) que propaga o *cost to go*. Esta envolve o cálculo dos seguintes valores esperados:

$$E_c \left[\log \frac{c}{C} \right] = \frac{1}{2}x^T K^T \bar{R}^{-1}Kx + \frac{1}{2} \log \frac{|\bar{R}|}{|R|} + \frac{1}{2} \text{Tr}(\bar{R}^{-1}R) - \frac{M}{2}, \quad (\text{C.29})$$

$$E_c [D(s||S)] = \frac{1}{2} E_c [(Ax + Bu)^T \bar{Q}^{-1}(Ax + Bu)] + \frac{1}{2} \text{Tr}(\bar{Q}^{-1}Q) + \frac{1}{2} \log \frac{|\bar{Q}|}{|Q|} - \frac{N}{2}, \quad (\text{C.30})$$

onde

$$\begin{aligned} E_c [(Ax + Bu)^T \bar{Q}^{-1}(Ax + Bu)] &= \\ &= E_c \left[u^T B^T \bar{Q}^{-1}Bu + 2u^T B^T \bar{Q}^{-1}Ax + x^T A^T \bar{Q}^{-1}Ax \right] \\ &= E_c \left[(u - Kx)^T B^T \bar{Q}^{-1}B(u - Kx) \right] - \overbrace{E_c \left[2(u - Kx)^T B^T \bar{Q}^{-1}BKx \right]}^{=0} + \\ &\quad + x^T K^T B^T \bar{Q}^{-1}BKx + 2x^T K^T B^T \bar{Q}^{-1}Ax + x^T A^T \bar{Q}^{-1}Ax \\ &= \text{Tr}(B^T \bar{Q}^{-1}BR) + x^T \left(K^T B^T \bar{Q}^{-1}BK + 2K^T B^T \bar{Q}^{-1}A + A^T \bar{Q}^{-1}A \right) x \\ &= \text{Tr}(B^T \bar{Q}^{-1}BR) + x^T (A + BK)^T \bar{Q}^{-1}(A + BK)x. \end{aligned} \quad (\text{C.31})$$

Somando (C.29) e (C.30) obtém-se a soma de uma forma quadrática em x com um escalar:

$$\begin{aligned} \frac{1}{2}x^T \left((A + BK)^T \bar{Q}^{-1}(A + BK) + K^T \bar{R}^{-1}K \right) x + \\ + \frac{1}{2} \text{Tr}(B^T \bar{Q}^{-1}BR) + \frac{1}{2} \text{Tr}(\bar{Q}^{-1}Q) + \frac{1}{2} \text{Tr}(\bar{R}^{-1}R) + \frac{1}{2} \log \frac{|\bar{Q}| \cdot |\bar{R}|}{|Q| \cdot |R|} - \frac{N + M}{2}. \end{aligned} \quad (\text{C.32})$$

O *cost to go* está definido no instante de tempo terminal por $-\log \gamma_T = 0$. Este é usado para obter os ganhos K_{T-1} , conforme a equação (C.28). Em $t = T - 1$, o *cost to go* é dado pela expressão (C.32). Pode-se então escrever que

$$-\log \gamma_{T-1}(x_{T-1}) = \frac{1}{2}x_{T-1}^T \Sigma_{T-1}^{-1}x_{T-1} + \omega_{T-1}, \quad (\text{C.33})$$

onde Σ_{T-1}^{-1} e ω_{T-1} são, respectivamente, uma matriz quadrada e um termo escalar apropriados obtidos de (C.32). Colocando a hipótese de o *cost to go* ter uma expressão desta forma para todos os instantes temporais do horizonte anteriores a $T - 1$, calcula-se o termo

$$\begin{aligned} E_{s_{t+1}} \left[-\log \gamma_{t+1}(x_{t+1}) \right] &= \frac{1}{2} E_{s_{t+1}} \left[x_{t+1}^T \Sigma_{t+1}^{-1} x_{t+1} \right] + \omega_{t+1} \\ &= \frac{1}{2} \left(\text{Tr}(\Sigma_{t+1}^{-1} Q) + (Ax_t + Bu_t)^T \Sigma_{t+1}^{-1} (Ax_t + Bu_t) \right) + \omega_{t+1}, \end{aligned} \quad (\text{C.34})$$

e ainda

$$\begin{aligned} E_{c_t} \left[E_{s_{t+1}} \left[-\log \gamma_{t+1}(x_{t+1}) \right] \right] &= \frac{1}{2} E_{c_t} \left[(Ax_t + Bu_t)^T \Sigma_{t+1}^{-1} (Ax_t + Bu_t) \right] + \\ &\quad + \frac{1}{2} \text{Tr}(\Sigma_{t+1}^{-1} Q) + \omega_{t+1} \\ &= \frac{1}{2} x_t^T (A + BK_t)^T \Sigma_{t+1}^{-1} (A + BK_t) x_t + \\ &\quad + \frac{1}{2} \text{Tr}(B^T \Sigma_{t+1}^{-1} BR) + \frac{1}{2} \text{Tr}(\Sigma_{t+1}^{-1} Q) + \omega_{t+1}. \end{aligned} \quad (\text{C.35})$$

Obtém-se finalmente que $-\log \gamma_t(x_t)$, calculado como a soma das parcelas (C.29), (C.30) e (C.34), resulta em

$$\begin{aligned} -\log \gamma_t(x_t) &= \frac{1}{2} x_t^T \left((A + BK_t)^T (\bar{Q}^{-1} + \Sigma_{t+1}^{-1}) (A + BK_t) + K_t^T \bar{R}^{-1} K_t \right) x_t + \\ &\quad + \frac{1}{2} \text{Tr}(B^T (\bar{Q}^{-1} + \Sigma_{t+1}^{-1}) BR) + \frac{1}{2} \text{Tr}((\bar{Q}^{-1} + \Sigma_{t+1}^{-1}) Q) + \\ &\quad + \frac{1}{2} \text{Tr}(\bar{R}^{-1} R) + \frac{1}{2} \log \frac{|\bar{Q}| \cdot |\bar{R}|}{|Q| \cdot |R|} - \frac{N + M}{2} + \omega_{t+1}, \end{aligned} \quad (\text{C.36})$$

donde se conclui que a hipótese colocada é satisfeita, isto é,

$$-\log \gamma_t(x_t) = \frac{1}{2} x_t^T \Sigma_t^{-1} x_t + \omega_t, \quad (\text{C.37})$$

com

$$\Sigma_t^{-1} = (A + BK_t)^T (\bar{Q}^{-1} + \Sigma_{t+1}^{-1}) (A + BK_t) + K_t^T \bar{R}^{-1} K_t \quad (\text{C.38})$$

$$\begin{aligned} \omega_t &= \frac{1}{2} \text{Tr}(B^T (\bar{Q}^{-1} + \Sigma_{t+1}^{-1}) BR) + \frac{1}{2} \text{Tr}((\bar{Q}^{-1} + \Sigma_{t+1}^{-1}) Q) + \frac{1}{2} \text{Tr}(\bar{R}^{-1} R) + \\ &\quad + \frac{1}{2} \log \frac{|\bar{Q}| \cdot |\bar{R}|}{|Q| \cdot |R|} - \frac{N + M}{2} + \omega_{t+1}. \end{aligned} \quad (\text{C.39})$$

Definindo a matriz $\bar{\Sigma}^{-1} \stackrel{\text{def}}{=} \bar{Q}^{-1} + \Sigma^{-1}$, as equações (C.38)-(C.39) podem ser simplificadas para

$$\bar{\Sigma}_t^{-1} = (A + BK_t)^T \bar{\Sigma}_{t+1}^{-1} (A + BK_t) + K_t^T \bar{R}^{-1} K_t + \bar{Q}^{-1} \quad (\text{C.40})$$

$$\begin{aligned} \omega_t &= \frac{1}{2} \text{Tr}(B^T \bar{\Sigma}_{t+1}^{-1} BR) + \frac{1}{2} \text{Tr}(\bar{\Sigma}_{t+1}^{-1} Q) + \frac{1}{2} \text{Tr}(\bar{R}^{-1} R) + \\ &\quad + \frac{1}{2} \log \frac{|\bar{Q}| \cdot |\bar{R}|}{|Q| \cdot |R|} - \frac{N + M}{2} + \omega_{t+1}. \end{aligned} \quad (\text{C.41})$$

Tendo determinado $-\log \gamma_t(x_t)$ para todo o horizonte $t = 0, \dots, T-1$, os ganhos K_t correspondentes podem ser calculados resolvendo a condição de optimalidade (4.71). Para inserir $E_{s_{t+1}}[-\log \gamma_{t+1}(x_{t+1})]$ na condição de optimalidade consideram-se apenas os termos de ordem ímpar, tal como anteriormente. Isto é, rescreve-se (C.34) da seguinte maneira:

$$E_{s_{t+1}}[-\log \gamma_{t+1}(x_{t+1})] = (u_t - K_t x_t)^T B^T \bar{\Sigma}_{t+1}^{-1} B K_t x_t + (u_t - K_t x_t)^T B^T \bar{\Sigma}_{t+1}^{-1} A x + \text{t.o.p.} \quad (\text{C.42})$$

Obtém-se então a condição de optimalidade

$$E_{c_t} \left[\underbrace{(u_t - K_t x_t)^T \left(B^T \bar{\Sigma}_{t+1}^{-1} A + (B^T \bar{\Sigma}_{t+1}^{-1} B + \bar{R}^{-1}) K_t \right) x_t}_{\text{factor escalar}} (u_t - K_t x_t)^T R^{-1} \right] = 0, \quad (\text{C.43})$$

que tem como solução os ganhos

$$K_t = -(B^T \bar{\Sigma}_{t+1}^{-1} B + \bar{R}^{-1})^{-1} B^T \bar{\Sigma}_{t+1}^{-1} A. \quad (\text{C.44})$$

C.3 Alguns resultados úteis

C.3.1 Cálculo matricial

Sejam x, y matrizes coluna, e A, B matrizes de dimensões apropriadas. Então

$$\frac{\partial}{\partial x} (x^T A x) = x^T (A^T + A) \quad (\text{C.45})$$

$$\frac{\partial}{\partial A} (y^T A x) = x y^T \quad (\text{C.46})$$

$$\frac{\partial}{\partial A} (y^T A^T B A x) = x y^T A^T B + y x^T A^T B^T. \quad (\text{C.47})$$

C.3.2 Função densidade de probabilidade Gaussiana

Considere-se que X tem densidade de probabilidade Gaussiana

$$p(x) = \mathcal{N}(\mu, \Sigma) = \frac{1}{\sqrt{(2\pi)^n |\Sigma|}} e^{-\frac{1}{2}(x-\mu)^T \Sigma^{-1} (x-\mu)}. \quad (\text{C.48})$$

Então:

$$E_{p(x)}[Ax + c] = A\mu + c, \quad (\text{C.49})$$

$$E_{p(x)}[(x-b)^T Q(x-b)] = \text{Tr}(Q\Sigma) + (\mu-b)^T Q(\mu-b). \quad (\text{C.50})$$

C.3.3 Entropia de uma distribuição Gaussiana

A entropia diferencial de uma distribuição Gaussiana é

$$h(X) = \frac{1}{2} \log \left((2\pi e)^n |\Sigma| \right). \quad (\text{C.51})$$

C.3.4 Divergência de K-L entre distribuições Gaussianas

Se $p(x)$ e $q(x)$ são duas distribuições Gaussianas multivariadas

$$p(x) = \mathcal{N}(\mu_p, \Sigma_p), \quad (\text{C.52})$$

$$q(x) = \mathcal{N}(\mu_q, \Sigma_q). \quad (\text{C.53})$$

Então a divergência de Kullback-Leibler entre $p(x)$ e $q(x)$ é

$$D(p\|q) = \frac{1}{2} \left[\left(\log \frac{\det \Sigma_q}{\det \Sigma_p} \right) + \text{tr}(\Sigma_q^{-1} \Sigma_p) + (\mu_q - \mu_p)^T \Sigma_q^{-1} (\mu_q - \mu_p) - N \right]. \quad (\text{C.54})$$

Bibliografia

- [1] S. Amari and S.C. Douglas. Why natural gradient? In *Acoustics, Speech, and Signal Processing, 1998. ICASSP '98. Proceedings of the 1998 IEEE International Conference on*, volume 2, pages 1213 – 1216 vol.2, 12-15 May 1998.
- [2] S. Amari, S.C. Douglas, A. Cichocki, and H.H. Yang. Multichannel blind deconvolution and equalization using the natural gradient. In *Signal Processing Advances in Wireless Communications, 1997 First IEEE Signal Processing Workshop on*, pages 101 – 104, April 1997.
- [3] S. Amari, H. Park, and K. Fukumizu. Adaptive method of realizing natural gradient learning for multilayer perceptrons. *Neural Computation*, 12(6):1399–1409, June 2000.
- [4] Shun-Ichi Amari. *Differential-Geometric Methods in Statistics*. Lecture Notes in Statistics. Springer-Verlag, 1985.
- [5] Shun-Ichi Amari. Natural gradient works efficiently in learning. *Neural Computation*, 10(2):251–276, 1998.
- [6] Shun-Ichi Amari and Hiroshi Nagaoka. *Methods of Information Theory*, volume 191 of *Translations of Mathematical Monographs*. American Mathematical Society, 2000.
- [7] Suguru Arimoto. An algorithm for computing the capacity of arbitrary discrete memoryless channels. *Information Theory, IEEE Transactions on*, 18(1):14 – 20, Jan 1972.
- [8] M. Sanjeev Arulampalam, Simon Maskell, Neil Gordon, and Tim Clapp. A tutorial on particle filters for online nonlinear/non-gaussian bayesian tracking. *IEEE Transactions on Signal Processing*, 50(2):174–188, Feb. 2002.
- [9] M. S. Bazaraa, Hanif D. Sherali, and C. M. Shetty. *Nonlinear programming: theory and algorithms*. Wiley-Interscience series in discrete mathematics and optimization. Wiley, New York, 2nd ed. edition, 1993. Includes bibliographical references (p. 576-625) and index.
- [10] Richard E. Blahut. Computation of channel capacity and rate-distortion functions. *Information Theory, IEEE Transactions on*, 18(4):460 – 473, Jul 1972.
- [11] William M. Boothby. *An Introduction to Differentiable Manifolds and Riemannian Geometry*, volume 120 of *Pure and Applied Mathematics*. Academic Press, second edition, 1986.
- [12] Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory*. Wiley Series in Telecommunications. Wiley, 1991.

- [13] Aniruddha Datta and Edward R. Dougherty. *Introduction to Genomic Signal Processing with Control*. 2007.
- [14] Manfredo Perdigão do Carmo. *Geometria Riemanniana*. Projecto Euclides. IMPA, second edition, 1988.
- [15] Y. Ephraim and N. Merhav. Hidden markov processes. *IEEE Transactions on Information Theory*, 48(6):1518 – 1569, June 2002.
- [16] Daniel Gabay. Minimizing a differentiable function over a differential manifold. Technical Report 9, INRIA, February 1980.
- [17] P. E. Gill and W. Murray, editors. *Numerical methods for constrained optimization*, London, 1974. Academic Press. Bibliography: p. 261-268.
- [18] Shun ichi Amari. Neural learning in structured parameter spaces — natural riemannian gradient. In Michael C. Mozer, Michael I. Jordan, and Thomas Petsche, editors, *Advances in Neural Information Processing Systems*, volume 9, page 127. The MIT Press, 1997.
- [19] Rudolf Kulhavý. *Recursive Nonlinear Estimation: A Geometric Approach*. Lecture Notes in Control and Information Sciences. Springer, 1996.
- [20] Rudolf Kulhavý. Recursive nonlinear estimation - a geometric approach. *Automatica*, 26(3):545–555, May 1990.
- [21] M Kárný. Towards fully probabilistic control design. *Automatica*, 32(12):1719–1722, Dec. 1996.
- [22] Frank L. Lewis and Vassilis L. Syrmos. *Optimal Control*. Second edition edition, 1995.
- [23] Xue-Bin Liang. A new iterative algorithm for calculating the capacity of discrete memoryless channels. In *Proceedings of the IEEE GLOBECOM*, pages 1–5, 2006.
- [24] J. H. Manton. Optimization algorithms exploiting unitary constraints. *IEEE Transactions on Signal Processing*, 50(3):635–650, Mar. 2002.
- [25] G. Matz and P. Duhamel. Information geometric formulation and interpretation of accelerated blahut-arimoto-type algorithms. In *Information Theory Workshop, 2004. IEEE*, pages 66 – 70, Oct. 2004.
- [26] Michael K. Murray and John W. Rice. *Differential Geometry and Statistics*. Monographs on Statistics and Applied Probability. Chapman & Hall, 1993.
- [27] E. Nováková and M. Kárný. Fully probabilistic control design for markov chains. In *European Control Conference*, 1997.
- [28] V. Peterka. Bayesian system identification. *Automatica*, 17(1):41–53, 1981.
- [29] S. S. Rao. *Optimization: theory and applications*. Wiley Eastern Ltd., New Delhi, 1978. Includes bibliographical references.
- [30] Claude E. Shannon. A mathematical theory of communication. *The Bell System Technical Journal*, 27:379–423,623–656, 1948.