

Programação Dinâmica e Controle Ótimo

Prefácio:

Este livro de dois volumes é baseado em um curso de pós-graduação de primeiro ano em programação dinâmica e controle ótimo que ensinei por mais de vinte anos na Universidade de Stanford, da Universidade de Illinois, e do Instituto de Tecnologia de Massachusetts. O curso tem sido tipicamente com a participação de estudantes de engenharia, pesquisa operacional, economia e matemática aplicada. Assim, um dos objetivos principais do livro tem sido proporcionar um tratamento unificado do tema, adequado para um público amplo. Em particular, os problemas com um caráter contínuo, tais como problemas de controle estocásticos, populares em teoria de controle moderno, são tratados simultaneamente com problemas com um personagem discreto, tais como problemas de decisão markovianos, popular em pesquisa de operações. Além disso, muitas aplicações e exemplos, retirados de uma ampla variedade de campos, são discutidos.

O livro pode ser visto como uma versão muito expandida e pedagogicamente melhorada do meu livro de 1987, "Dinâmica de programação: Modelos determinísticos e estocásticos", publicado pela Prentice-Hall. Eu incluí muito material novo sobre os problemas do caminho mais curto determinísticos e estocásticos, bem como um novo capítulo sobre problemas de controle ótimo de tempo contínuo e o Princípio da Máxima Pontryagin, desenvolvido a partir de um ponto de vista de programação dinâmica. Tenho também acrescentou uma extensa exposição de técnicas de aproximação à base de simulação para programação dinâmica. Essas técnicas, que são muitas vezes referida como "programação neuro-dinâmica" ou "aprendizado por reforço," representam um avanço na aplicação prática de programação dinâmica para problemas complexos que envolvem a dupla maldição de grande dimensão e falta de um modelo matemático preciso. Outros materiais também foi aumentada, substancialmente modificado e atualizado.

Com o novo material, no entanto, o livro cresceu tanto em tamanho, tornou-se necessário que a dividi-la em dois volumes: um no horizonte finito, e o outro em problemas de horizonte infinito. Esta divisão não era natural em termos de tamanho, mas também em termos de estilo e orientação. O primeiro volume é mais orientado para a modelagem, ea segunda é mais orientado para a análise matemática e computação. Para tornar o primeiro volume independente para instrutores que desejam cobrir uma quantidade modesta de material horizonte inifinite em um curso que é orientada principalmente para modelagem, conceituação, e problemas horizonte finito, eu adicionei um capítulo final que fornece um tratamento introdutório de problemas horizonte infinitos.

Muitos tópicos do livro são relativamente independente dos outros. Por exemplo, o Capítulo 2 do Vol. I em mais curtos problemas de caminho pode ser ignorada sem perda de continuidade, eo mesmo é verdade para o capítulo 3 do Vol. I, que trata de tempo contínuo de controle ótimo. Como resultado, o livro pode ser usado para ensinar vários tipos diferentes de cursos.

1. Um curso de dois semestres que abrange tanto os volumes
2. Um curso de um semestre focado principalmente em problemas de horizonte finito que cobre a maior parte do primeiro volume
3. Um curso de um semestre focado no controle ótimo estocástico que abrange os capítulos 1, 4, 5 e 6 do Vol. I, e os capítulos 1, 2 e 4 de Vol. II
4. Um curso de um semestre que cobre o Capítulo 1, cerca de 50% dos capítulos 2 a 6 de Vol. I, e cerca de 70% dos capítulos 1, 2 e 4 do Vol. II. Este é o curso eu costumo ensinar no MIT.
5. Um curso de engenharia de um quarto, que abrange os três primeiros capítulos e partes dos capítulos 4 a 6 do Vol. EU.
6. A um quarto curso orientado matematicamente focada em problemas de horizonte infinito, que abrange Vol. II.

O pré-requisito matemática para o texto é o conhecimento de cálculo avançado, teoria da probabilidade de introdução, e matriz-vetor álgebra. Um resumo deste material é fornecido

nos apêndices. Naturalmente, exposição prévia a teoria dinâmica do sistema, controle, otimização, ou operações de investigação será útil para o leitor, mas com base na minha experiência, o material dado aqui é razoavelmente auto-suficiente.

O livro contém um grande número de exercícios, e o leitor sério irá beneficiar grandemente, passando por eles. Soluções para todos os exercícios são compilados em um manual que está disponível para instrutores do autor. Muitos agradecimentos são devido às várias pessoas que passaram longas horas contribuindo para este manual, particularmente Steven Shreve, Eric Loiederman, Lakis Polymenakos, e Cynara Wu.

A programação dinâmica é uma técnica conceitualmente simples que pode ser adequadamente explicado por meio de análise elementar. No entanto, um tratamento matematicamente rigorosa da programação dinâmica geral requer a maquinaria complicada de probabilidade medida teórica. Minha escolha foi a de ignorar as complicada matemática, desenvolvendo o tema na generalidade, ao reivindicar rigor somente quando os espaços de probabilidade subjacentes são contáveis. Um tratamento matematicamente rigorosa do sujeito é realizado em minha monografia " Stochastic Optimal Controle: The Time caso discreto ", Academic Press, 1978, co-autoria de Steven Shreve. Esta monografia complementa o presente texto e fornece uma base sólida para os assuntos desenvolvidos tanto informalmente aqui.

Finalmente, eu sou grato a um número de indivíduos e instituições por suas contribuições para o livro. Minha compreensão do assunto foi aguçado enquanto eu trabalhava com Steven Shreve em nosso 1978 monografia. Minha interação e colaboração com [John Tsitsiklis](#) em caminhos mais curtos estocásticos e programação dinâmica aproximada ter sido mais valioso. Michael Caramanis, Emmanuel Fernandez-Gaucherand, Pierre Humblet, Lennart Ljung, e John Tsitsiklis ensinou a partir de versões do livro, e contribuiu com vários substantiva comentários e problemas de casa. Uma série de colegas ofereceram valiosos insights e informações, especialmente David Castanon, Eugene Feinberg, e Krishna

Pattipati. NSF apoiou a investigação. Prentice-Hall permitiu graciosamente a utilização de material do meu livro de 1987. Ensinar e interagindo com os alunos do MIT têm mantido meu interesse e entusiasmo para o assunto.

Prefácio à Segunda Edição: Volume 1

Esta segunda edição expandiu por quase 30% da cobertura do original. A maior parte do novo material é concentrado em quatro áreas:

(A) No capítulo 4, a seção foi adicionada na estimativa e controle de sistemas com uma descrição não-probabilística (associação do conjunto) de incerteza. Este assunto, um favorito pessoal do autor, uma vez que foi o tema de seu Ph.D. 1971 tese no MIT, tornou-se popular, como minimax e métodos de controle H-infinito para a crescente proeminência.

(B) Capítulo 6 foi dobrado de tamanho, para refletir a popularidade de controle abaixo do ideal e métodos de programação neuro-dinâmica. Em particular, a cobertura da segurança equivalente, e os métodos de verificação à frente limitados foi substancialmente ampliado. Além disso, uma nova seção foi adicionada sobre a programação neuro-dinâmica e algoritmos de rollout, e suas aplicações em otimização combinatória e controle ótimo estocástico.

(C) No Capítulo 7, uma introdução à de tempo contínuo, problemas de decisão semi-Markov foi adicionado em uma última seção separada.

(D) Um novo anexo foi incluído, que lida com várias formulações de problemas de decisão em situações de incerteza. As fundações do minimax e abordagens de utilidade esperada são enquadrados dentro de um contexto mais amplo, e alguns dos aspectos da teoria da utilidade são discutidos.

Prefácio à Segunda Edição: Volume 2

Esta segunda edição do Vol. II deve ser visto como uma revisão relativamente menor do original. A cobertura foi ampliada em algumas áreas da seguinte forma:

(A) Em Capítulo 1, o material foi adicionado em variantes do método política iteração.

(B) No Capítulo 2, o material sobre os métodos de programação neuro-dinâmica foi atualizada e expandida para refletir alguns desenvolvimentos recentes.

(C) No Capítulo 4, material foi adicionado em alguns novos métodos valor de iteração.

(D) No Capítulo 5, o material sobre os problemas semi-Markov foi revista, com uma parcela significativa simplificado e deslocou-se para Volume I. \ smskip

Há também diversas adições e melhorias espalhadas por todo o texto, e uma cobertura mais detalhada de problemas determinísticos é dada no Capítulo 1. Por fim, um novo recurso baseado na Internet foi acrescentado ao livro, que estende seu alcance e cobertura. Muitos dos exercícios teóricos foram resolvidos em detalhes e suas soluções foram postou em página [www](#) do livro. Estes exercícios foram marcados com o símbolo [www](#).

Eu gostaria de expressar os meus agradecimentos aos muitos colegas que contribuíram com sugestões para a melhoria da segunda edição.

Prefácio à terceira edição: Volume 1

A terceira edição contém uma quantidade substancial de material novo, especialmente em programação dinâmica

aproximada, que agora se tornou um dos principais pontos focais do livro. Em particular:

(A) O tema do controle minimax foi desenvolvido em maior detalhe, incluindo uma nova seção no Capítulo 1, que se conecta com o novo material no Capítulo 6.

(B) A seção sobre algoritmos de leilão para caminhos mais curtos no Capítulo 2 foi eliminado. Estes métodos não são atualmente utilizados em programação dinâmica, e uma discussão detalhada foi fornecida em um capítulo a partir do autor Otimização de Rede livro. Este capítulo pode ser baixado gratuitamente a partir de <http://web.mit.edu/dimitrib/www/net.html>

(C) A seção foi adicionada no Capítulo 2 em programação dinâmica e algoritmos de menor caminho para problemas restritos e multiobjetivo.

(D) O material em estatísticas suficientes e problemas de decisão Markov parcialmente observáveis na Seção 5.4 foi reestruturado e ampliado.

(D) o material novo considerável foi adicionado no Capítulo 6:

(1) uma discussão alargada das políticas de verificação à frente de uma etapa e barrancos de desempenho associados na Seção 6.3.1.

(2) Uma discussão sobre métodos de agregação e discretização de \ problemas estatais -nuous Conti (ver subsecção 6.3.4).

(3) Uma discussão de modelo de controle preditivo e sua relação com outros métodos de controle abaixo do ideal (ver subsecção 6.5.2).

(4) Um tratamento expandido de open-loop de controle de feedback e métodos relacionados com base em uma estrutura reduzida (ver subsecção 6.5.3).

Tenho também acrescentou alguns exercícios, e revisto algumas seções, preservando o seu conteúdo essencial. Agradecimentos são devidos a Haixia Lin, que

elaboraram vários exercícios, e Janey Yu, que revisou algumas das novas seções e me deu um feedback valioso.

[Dimitri P. Bertsekas](#)

verão 2005

Prefácio à terceira edição: Volume 2

Esta é uma grande revisão da 2ª edição, e contém uma quantidade substancial de novo material, bem como a reorganização do material antigo. O comprimento do texto aumentou em mais do que 50%, e mais de metade do material velho foi renovado e / ou revisado. A maior parte do material é adicionado em quatro zonas.

(A) A cobertura do problema do custo médio do Capítulo 4 aumentou muito em extensão e profundidade. Em particular, há agora uma análise completa de problemas multi-cadeia, bem como uma análise mais extensa de espaços infinitos problemas (Seção 4.6).

(B) O material em programação dinâmica aproximada foi recolhido no Capítulo 6. Ele foi bastante expandido para incluir uma nova pesquisa, completando assim o livro de 1996 " ` ` Programação Neuro-Dynamic. "

(C) mapeamentos de contração e seu papel em várias análises foram destaque na material novo sobre infinitos problemas de espaço estado (pontos 1.4, 2.5 e 4.6), e em seu uso no material de programação dinâmica aproximada do Capítulo 6.

(D) As questões de medida de teoria matemática que devem ser abordados para uma teoria rigorosa da programação dinâmica estocástica foram ilustradas e resumidos num apêndice para o benefício do leitor orientada matematicamente.

Também foram adicionados alguns exercícios e algumas seções foram revistos, preservando o seu conteúdo essencial.

Eu gostaria de expressar os meus agradecimentos a muitos colegas que contribuíram valiosos comentários. Estou particularmente grato a Ciamac Moallemi, Steven Shreve, John Tsitsiklis, e Ben Van Roy, que analisou alguns dos novos materiais e cada um contribuiu com várias sugestões substanciais. Gostaria de agradecer especialmente Janey Yu por sua ajuda extraordinária. Janey ler com muito cuidado e olho grande parte afiados do livro, contribuíram análise importante e muitas, comentários substantivos incisivas, e também colaborou comigo na pesquisa que foi incluída no Capítulo 6.

[Dimitri P. Bertsekas](#)

outono de 2006

Prefácio à quarta edição: Volume 2

Esta é uma grande revisão da 3ª edição e contém uma quantidade substancial de novo material, bem como a reorganização do material antigo. O comprimento aumentou em mais de 60% a partir da 3ª edição, e a maior parte do material antigo foi reestruturado e / ou revisado. Volume 2 agora conta com mais de 700 páginas e é maior em tamanho do que Vol. 1. Ele pode sem dúvida ser vista como um livro de novo!

Aproximado DP tornou-se o ponto focal central de Vol. II, e ocupa mais de metade do livro (os dois últimos capítulos, e uma grande parte dos capítulos 1-3). Assim, pode-se também visualizar Vol. \ II como uma continuação ao meu livro de 1996 `` Programação Neuro-Dynamic "(co-autoria com John Tsitsiklis). O presente livro centra-se em grande parte em uma nova pesquisa que se tornou disponível a partir de 1996. Por outro lado, o estilo clássico do livro foi preservada, e algum material foi explicado a um nível intuitivo ou informal,

referindo-se à literatura de jornal ou o livro Programação Neuro-Dynamic para um tratamento mais matemática.

No processo de expansão e reorganização, o desenho do livro se torna mais modular e adequado para uso em sala de aula. O material do núcleo, que pode ser percorrida em cerca de um terço a metade de um semestre é o capítulo 1 (exceto para as seções específicas de aplicativos 1.3 e 1.4), Capítulo 2, e no Capítulo 6, que são auto-suficientes quando tomados em conjunto. Este material se concentra em problemas com desconto, e pode ser completada por partes de capítulos 3 e 7.1 na estocásticos menores problemas de caminho. Na verdade, este compreende a metade do que eu cubro na minha classe MIT (a outra metade vem do Volume I, incluindo o capítulo 6 do que o volume que lida com horizonte finito aproximado DP). O material sobre os problemas de custo médio, dada no Capítulo 5, e as seções 7.2 e 7.4, e o material avançado em modelos positivos e negativos DP (Capítulo 4), e Monte Carlo álgebra linear (Seção 7.3) são temas de terminais que podem ser abrangidos pelo critério do instrutor.

Como o foco do livro mudou, eu coloquei uma maior ênfase na pesquisa nova ou recente na aproximado DP e métodos baseados em simulação, bem como sobre os métodos iterativos assíncronas, tendo em conta o papel central da simulação, que é por natureza assíncrona. Muito desse material é uma consequência da minha pesquisa e da pesquisa dos meus colaboradores, realizados nos seis anos desde a edição anterior. Alguns dos destaques, pela ordem que aparecem no livro, são:

(A) Métodos computacionais para generalizadas descontados DP (Seções 2.5 e 2.6), incluindo limites de erro para aproximações na Seção 2.5, e os métodos de iteração assíncrona política otimista das seções 2.6.2 e 2.6.3, e sua aplicação ao jogo e problemas minimax , iteração política restrita, e Q-learning.

(B) os métodos Política de iteração (incluindo versões otimistas assíncronas) para estocásticos problemas caminho mais curto que envolvem políticas inadequadas (Seção 3.4).

(C) novo material extensivo em vários métodos baseados em simulação, o valor aproximado de iteração e política nas Seções 6,3-6,6 (equação e de agregação de métodos projetados).

(D) Novos algoritmos Q-learning confiáveis para otimista iteração política (Seções 2.6.3 e 6.6.2).

(E) técnicas de simulação para Novos métodos de várias etapas, como de forma livre geométrica e amostragem (Seções 6.4.1 e 7.3.3).

(F) o novo material extensivo em Monte Carlo álgebra linear na Seção 7.3 (principalmente o baseado em simulação e solução aproximada de grandes sistemas de equações lineares), que estende a metodologia DP da avaliação de políticas aproximada.

Grande parte da investigação em (a) - (e) é baseado no meu trabalho com Janey (Huizhen) Yu, enquanto a maioria das pesquisas em (f) é baseado no meu trabalho com Janey Yu e Mengdi Wang. Minha colaboração com Janey e Mengdi teve um forte impacto sobre o livro, e é muito apreciada. Alguns de nossos trabalhos foi apresentado em resumo só, e foi adaptado para se ajustar ao estilo e propósitos deste livro; naturalmente, quaisquer deficiências na sua apresentação são de minha inteira responsabilidade. O leitor é remetido para os nossos documentos conjuntos e individuais, que descrevem mais plenamente a nossa investigação, incluindo material que não pode ser abrangido por este livro.

Quero manifestar o meu apreço aos colegas e colaboradores na pesquisa aproximada DP, que contribuíram para o livro em várias formas, principalmente Vivek Borkar, Angelia Nedic, e Ben Van Roy. Um agradecimento especial vai para John Tsitsiklis, com quem interagi extensivamente através da colaboração e partilha de ideias sobre DP e algoritmos assíncronos por mais de 30 anos. Gostaria também de reconhecer interações votos com muitos colegas, incluindo Vivek Farias, Eugene Feinberg, Warren Powell, Martin Puterman, Uriel Rothblum, e Bruno Scherrer. Por último, gostaria de agradecer aos muitos alunos em minhas aulas DP

da última década, que trabalha pacientemente com um livro em desenvolvimento, e contribuíram com suas idéias e experiências através de seus projetos de pesquisa a partir de uma ampla variedade de campos de aplicação.

[Dimitri P. Bertsekas](#)

Junho 2012