

Programação Não Linear: Second Edition

Prefácio:

Programação não linear é um campo maduro, que tem experimentado grandes avanços nos últimos dez anos. O primeiro tal desenvolvimento é a fusão de algoritmos de programação linear e não-linear através da utilização de métodos de pontos interiores. Isso resultou em um profundo repensar a forma como vamos resolver problemas de programação linear, e em uma grande reavaliação da forma como tratamos as restrições na programação não-linear. Um segundo desenvolvimento, menos visível, mas ainda importante, é a maior ênfase em problemas de grande escala, e os algoritmos associados que se aproveitam da estrutura do problema, bem como hardware paralelo. Uma terceira desenvolvimento tem sido o uso extensivo de otimização sem restrições iterativo para resolver os difíceis problemas de mínimos quadrados, decorrentes do treinamento de redes neurais. Como resultado, os métodos de gradiente-like simples e regras tamanho do passo alcançaram importância crescente.

O objetivo deste livro é fornecer uma conta up-to-date, abrangente e rigorosa de programação não-linear ao nível pós-graduado começando. Além dos temas clássicos, como os algoritmos de descida, a teoria do multiplicador de Lagrange, e dualidade, cobrimos alguns dos desenvolvimentos recentes importantes: métodos de pontos interiores para programas lineares e não lineares, os principais aspectos de otimização em larga escala, e problemas de mínimos quadrados e treinamento da rede neural.

Uma outra característica notável do livro é que ele trata multiplicadores de Lagrange e dualidade utilizando duas abordagens diferentes e complementares: uma abordagem variacional baseado no teorema da função implícita, e uma abordagem de análise convexa com base em argumentos geométricos. A primeira abordagem se aplica a uma classe mais ampla de problemas, enquanto o último é mais elegante e mais poderoso para os programas convexos a que se aplica.

A descrição capítulo por capítulo do livro segue:

Capítulo 1:

Este capítulo aborda otimização sem restrições: conceitos principais, condições de otimização, e algoritmos. O material é clássico, mas há

o debate de temas frequentemente deixados sem tratamento, tais como o comportamento de algoritmos para problemas singulares, treinamento da rede neural e de tempo discreto controle ótimo.

Capítulo 2:

Este capítulo trata de otimização restrita ao longo de um conjunto convexo, sem o uso de multiplicadores de Lagrange. Eu prefiro a cobrir este material antes de lidar com a complexa maquinaria de multiplicadores de Lagrange, porque achei que os alunos absorvem facilmente algoritmos como gradiente condicional, projeção do gradiente, e coordenar a descida, que podem ser vistos como extensões naturais de algoritmos de descida sem restrições. Este capítulo contém igualmente um tratamento do método de escalonamento afim de programação linear.

Capítulo 3:

Este capítulo dá um tratamento detalhado dos multiplicadores de Lagrange, as condições necessárias e suficientes associados e análise de sensibilidade. As três primeiras seções lidar com igualdade e desigualdade não-lineares. A última seção lida com restrições lineares e desenvolve uma forma simples de teoria da dualidade para problemas com restrições lineares com custo diferenciável, incluindo programação linear e quadrática.

Capítulo 4:

Este capítulo trata de algoritmos de otimização que usam penalidades e multiplicadores de Lagrange, incluindo barreira, aumentada de Lagrange, programação quadrática seqüencial, e métodos de pontos interiores primal-dual de programação linear. O tratamento é extenso, e toma emprestado da minha monografia de 1982, a pesquisa sobre métodos de multiplicadores de Lagrange.

Capítulo 5:

Este capítulo fornece uma cobertura detalhada da teoria da dualidade (Lagrange e Fenchel). O tratamento é totalmente geométrico, e tudo é explicado em termos de figuras intuitivas.

Capítulo 6:

Este capítulo trata dos métodos de otimização de grande escala com base na dualidade. Algum material é emprestado da minha Paralelo e Distribuído Algoritmos livro (co-autoria de [John Tsitsiklis](#)), mas há também um extenso tratamento de otimização não diferenciável, incluindo subgradiente, e-subgradiente, e corte métodos de avião. Métodos de decomposição, como Dantzig-Wolfe e Benders também são discutidos.

Apêndices:

Quatro apêndices são dadas. O primeiro apresenta um resumo do cálculo, análise e resultados de álgebra linear utilizada no texto. A segunda é uma extensa conta da teoria convexidade, incluindo provas dos resultados de convexidade poliédricas básicas sobre pontos extremos e Farkas 'lema, assim os fatos básicos sobre subgradientes. O terceiro apêndice abrange métodos de minimização unidimensionais. O último apêndice discute uma implementação do método de Newton para otimização sem restrições.

Inevitavelmente, alguns compromissos de cobertura teve que ser feito. O tema da otimização não linear tem crescido tanto que deixar de fora uma série de temas importantes não poderia ser evitada. Por exemplo, não são fornecidos uma discussão das desigualdades variacional, um tratamento mais profunda das condições de otimização, e um desenvolvimento mais detalhadas de métodos de Quasi-Newton. Além disso, um número maior de aplicativos de amostra teria sido desejável. Espero que os instrutores vão complementar o livro com o tipo de exemplos práticos que seus alunos estão mais familiarizados.

O livro foi desenvolvido através de um curso de pós-graduação de primeiro ano que eu lecionava na Univ. de Illinois e no MIT ao longo de um período de 20 anos. Os pré-requisitos matemáticos são álgebra matricial-vetorial e cálculo avançado, incluindo uma boa compreensão dos conceitos de convergência. Um curso de análise e / ou álgebra linear também deve ser muito útil, e daria a maturidade matemática necessária para acompanhar e apreciar o raciocínio matemático usado no livro. Algumas das seções do livro pode ser omitido, em primeira leitura, sem perda de continuidade. Estas seções foram marcados por uma estrela. A regra seguido aqui é que o material discutido na seção estrelada não é utilizado em uma seção de não-estrelada.

O livro pode ser usado para ensinar vários tipos diferentes de cursos.

1. Um curso de dois trimestre que abrange a maioria das seções de cada capítulo.
2. Um curso de um semestre que cobre Capítulo 1, exceto para Seção 1.9, Capítulo 2, excepto para as seções 2.4 e 2.5, capítulo 3, exceto para a Seção 3.4, capítulo 4, exceto para as peças das Seções 4.2 e 4.3, as três primeiras seções do capítulo 5, e uma seleção da Seção 5.4 e Capítulo 6. Este é o curso eu costumo ensinar no MIT.
3. Um curso de um semestre que cobre a maior parte dos capítulos 1, 2 e 3, e os algoritmos selecionados a partir do capítulo 4. eu ensinei este tipo de curso várias vezes. É menos exigente dos alunos, porque ele não exige a maquinaria da análise convexa, mas ainda fornece uma versão bastante poderoso da teoria da dualidade (Seção 3.4).
4. Um curso de um trimestre que abrange partes selecionadas dos capítulos 1, 2, 3 e 4. Esta é uma versão menos abrangente de (c) acima.

5. Um curso de um trimestre em análise e otimização convexa, que começa com o Apêndice B e abrange pontos 1.1, 2.1, 3.4, e no capítulo 5.

Há uma extensa literatura sobre programação não-linear e para dar uma bibliografia completa e um relato histórico da investigação que levou à atual forma de o assunto teria sido impossível. Eu, portanto, não têm tentado compilar uma lista abrangente de contribuições originais para o campo. Citei fontes que tenho usado extensivamente, que fornecem importantes extensões para o material do livro, essa pesquisa temas importantes, ou que são particularmente bem adaptado para outras leituras. Tenho também citou seletivamente algumas fontes que são historicamente significativa, mas a lista de referência está longe de ser exaustiva a este respeito. Geralmente, para auxiliar os pesquisadores no campo, eu teria preferido citar pesquisas e livros didáticos para assuntos que são relativamente maduro, e para dar um maior número de referências para desenvolvimentos relativamente recentes.

Finalmente, eu gostaria de expressar os meus agradecimentos a um número de indivíduos por suas contribuições para o livro. O meu entendimento conceitual do assunto foi formado na Universidade de Stanford, enquanto eu interagia com David Luenberger e ensinei usando seus livros. Esta experiência teve uma influência duradoura sobre o meu pensamento. Minha colaboração de pesquisa com vários colegas, especialmente Joe Dunn, Eli Gafni, Paul Tseng, e [John Tsitsiklis](#), foram muito úteis e são refletidas no livro. Agradeço as sugestões e idéias de um número de pessoas, especialmente David Castanon, Joe Dunn, Terry Rockafellar, Paul Tseng, e John Tsitsiklis. Sou grato aos muitos alunos e colaboradores cujos comentários levaram a correções e esclarecimentos. Steve Patek, Serap Savari, e Cynara Wu foram particularmente úteis a este respeito. David Logan, Steve Patek, e Lakis Polymenakos me ajudou a gerar o gráfico da capa, que retrata a função de custo de um simples problema de treinamento da rede neural. Minha esposa Joanna me animou com a sua presença e humor durante as longas horas de escrita, como ela tem com sua companhia de mais de 30 anos. Dedico este livro a ela com o meu amor.

[Dimitri P. Bertsekas](#)
bertsekas@lids.mit.edu

Prefácio à Segunda Edição

Esta segunda edição tem se expandido por cerca de 130 páginas a cobertura do original. Cerca de 40% do material novo representa adições variadas espalhadas por todo o texto. O restante lida com três novos temas. Estes são:

(A) A nova seção no Capítulo 3, que se concentra em um tratamento simples, mas de grande alcance de Fritz John condições necessárias e qualificações de restrição, e também inclui programação semi-infinita.

(B) Uma nova seção no Capítulo 5 sobre o uso da dualidade e relaxamento Lagrangian para resolver problemas de otimização discreta. Esta seção descreve várias aplicações que motivam, e fornece um elo de ligação entre a otimização contínua e discreta.

(C) Uma nova seção no Capítulo 6 sobre os métodos de subgradientes aproximados e incrementais. Este material é o tema da investigação conjunta em curso com Angelia Geary, mas pensava-se suficientemente importante para ser incluída no resumo aqui.

Um dos objetivos da revisão foi destacar as conexões de programação não-linear com outros ramos da otimização, como a programação linear, otimização de rede e otimização discreta / inteiro. Isso deve fornecer alguma flexibilidade adicional para o uso do livro em sala de aula. Além disso, a apresentação foi melhorada, o material de fundo matemática dos apêndices foi expandido, os exercícios foram reorganizados, e adicionou-se um número substancial de novos exercícios.

Um novo recurso baseado na Internet foi acrescentado ao livro, que se estende significativamente seu alcance e cobertura. Muitos dos exercícios teóricos, muito poucos deles novos, foram resolvidos em detalhes e suas soluções foram postou em página [www](#) do livro. Estes exercícios foram marcados com o símbolo "[www](#)"

Página [www](#) do livro também contém links para recursos adicionais, tais como códigos de computador e meus transparências palestras da minha classe de programação não-linear MIT.

Eu gostaria de expressar os meus agradecimentos aos muitos colegas que contribuíram com sugestões para a melhoria da segunda edição. Eu gostaria de agradecer particularmente Angelia Geary para ela uma grande ajuda com as soluções postadas na Internet dos exercícios teóricos.

Dimitri P. Bertsekas

De Junho de 1999