TÓPICOS EM CIÊNCIA DA COMPUTAÇÃO

PLANO DE ESTUDOS: OTIMIZAÇÃO E ANÁLISE CONVEXA

ELLEN HIDEMI FUKUDA E PAULO JOSÉ DA SILVA E SILVA

RESUMO. Este é o plano de estudos da disciplina MAC5701 – Tópicos em Ciência da Computação – da aluna de mestrado Ellen Hidemi Fukuda. Tal estudo será realizado no primeiro semestre de 2005 sob a supervisão do professor Dr. Paulo José da Silva e Silva. O objetivo deste trabalho é estudar os principais conceitos de Análise Convexa, uma subárea de Otimização Contínua que utiliza ferramentas como conjuntos convexos, funções convexas, subdiferenciais e conjugações.

1. INTRODUÇÃO

Consideremos o seguinte problema de Programação Linear (PL) de grande escala:

$$(1) \qquad \begin{cases} \min & \langle c, x \rangle \\ \text{s.a} & Ax = b \\ & Dx = e \\ & x \geq 0, \end{cases}$$

em que $c \in \mathbb{R}^n$, $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, $D \in \mathbb{R}^{d \times n}$, $e \in \mathbb{R}^d$ e D tem posto d. Consideramos que as restrições definidas pela matriz A são "difíceis", ou seja, que o problema (1) sem tais restrições é de fácil solução.

O algoritmo do volume (VA) foi proposto recentemente como alternativa para solução de (1), explorando a relaxação lagrangeana das restrições "difíceis" [2]. O método pode ser visto como um método de subgradientes com algumas características de métodos de feixe. Infelizmente a sua convergência não foi demonstrada, mas resultados numéricos encorajadores foram apresentados.

O algoritmo do volume foi recentemente tema de um outro artigo em que os autores propõem modificações que permitem demonstrar resultados de convergência [1]. Isto é obtido com pequenas alterações no VA de modo a ser possível interpretá-lo como um método de feixe.

O projeto do mestrado propõe o estudo dos métodos descritos acima, bem como sua implementação e comparação numérica. Como tópico avançado deseja-se explorar melhor a relação entre esses algoritmos, métodos extra-gradiente e métodos de ponto proximal. No entanto, todo esse estudo não será possível sem estudar inicialmente os principais conceitos de Análise Convexa. Este será justamente o assunto a ser abordado na disciplina MAC5701.

2. O PROJETO DE MESTRADO

Retomemos ao problema (1). Uma alternativa para a solução do problema, evitando a dificuldade gerada pelas restrições $Ax = b$, é atacar o seu dual obtido por relaxação lagrangeana. Para isso definamos a função lagrangeana, considerando apenas as restrições "difíceis",

$$L(x, \mu) \stackrel{\text{def}}{=} \langle c, x \rangle + \langle Ax - b, \mu \rangle, \qquad \mu \in \mathbb{R}^m.$$

Chamemos de Ψ o conjunto de restrições do problema "fácil":

$$\Psi \stackrel{\text{def}}{=} \{x \mid x \geq 0,\, Dx = e\}.$$

Sabemos então que vale a seguinte relação de dualidade fraca

$$\inf_{x\in\Psi} \sup_{\mu\in\mathbb{R}^m} L(x,\mu) \geq \sup_{\mu\in\mathbb{R}^m} \inf_{x\in\Psi} L(x,\mu). \tag{2}$$

O problema da esquerda tem o mesmo valor ótimo que (1). O problema à direita, em função das variáveis μ, é conhecido como problema dual. Nesse caso o problema é

$$\max_{\mu\in\mathbb{R}^m} \theta(\mu) \qquad \text{com} \qquad \theta(\mu) \stackrel{\text{def}}{=} \inf_{x\in\Psi} L(x,\mu). \tag{3}$$

A função dual θ, é côncava (por ser ínfimo de lineares), mas não-diferenciável.

Muitas vezes é possível provar que a desigualdade (2) é na realidade uma igualdade, fato conhecido como ausência de folga de dualidade. Nesse caso, as resoluções dos problemas dual e primal estão fortemente relacionadas. Note ainda que a avaliação da função objetivo dual corresponde à resolução de (1) sem as restrições "difíceis". Ainda, o cômputo de θ através do problema de otimização que a define produz naturalmente não só seu valor funcional, mas também um supergradiente. Isso é resumido na seguinte equação generalizada:

$$\forall x^* \in \operatorname*{argmin}_{x\in\Psi} L(x,\mu), \qquad Ax^* - b \in \partial\theta(\mu).$$

Isso torna possível a resolução direta do problema dual.

Existem vários métodos de otimização não-diferenciável que dependem apenas da possibilidade de valores funcionais e subgradientes. Por exemplo, há os métodos de subgradientes, planos de corte e os métodos de feixe [6, 3]. Os métodos de subgradientes destacam-se por sua simplicidade, trazendo uma analogia direta com métodos de Cauchy para otimização diferenciável. Porém, eles não possuem critérios de parada claros, sendo ainda lentos e muitas vezes pouco precisos. Os métodos de feixe, por outro lado, são bastante robustos e possuem bons critérios de parada. Porém, cada passo requer a solução de um problema quadrático que pode ser computacionalmente custoso.

Recentemente, Barahona e Anbil propuseram um novo método de subgradientes para a resolução do problema dual chamado de algoritmo do volume (VA) [2]. Esse método se caracteriza por tentar reconciliar as características de robustez e velocidade do método de feixes, com a simplicidade dos métodos de subgradientes. Outra característica marcante é o cômputo de uma solução primal ao mesmo tempo que se resolve o dual. Esse ponto primal é obtido estimando-se o volume associado a faces ativas das restrições "difíceis", o que explica o nome do algoritmo. Em seu artigo, Barahona e Anbil não provam a convergência do método, mas buscam validá-lo a partir de numerosas experiências computacionais com problemas lineares obtidos de relaxações de problemas combinatórios. Os autores também conjecturam que o algoritmo proposto deve ser especialmente eficiente nesse caso. Os resultados computacionais são bastante encorajadores.

Em seguida, Bahiense *et.al.* escreveram um interessante artigo explorando melhor a relação existente entre o VA e métodos de feixe [1]. Em seu trabalho, eles mostraram que o VA pode ser visto como um método extra-gradiente baseado em ϵ-subgradientes [5, 8, 7]. Para isso Bahiense *et.al.* apresentaram um novo ponto dual, p_t, que é gerado implicitamente pelo algoritmo a cada iteração e mostram que a direção tomada em cada passo do algoritmo é um ϵ-subgradiente nesse ponto. O valor de ϵ também foi estimado.

De posse dessa nova interpretação, Bahiense *et.al* buscaram então provar a convergência do algoritmo do volume. Para isso, os autores observaram que o critério de aceitação presente no VA para um próximo iterado é muito permissivo. Ele aceita qualquer passo que leve a uma melhora na função objetivo dual, não sendo capaz de evitar passos curtos demais.[1] Para contornar esse problema, eles propõem um modelo para aproximar a função objetivo dual, permitindo a estimação do quão "bom" é o decréscimo obtido com o candidato a novo ponto. Isso é feito através de uma condição do tipo Armijo e de uma avaliação do ganho que deveria ser obtido na direção extra-gradiente que se quer seguir. Essa modificação também permite uma escolha mais livre dos tamanhos de passo. O modelo de feixes obtido é bastante econômico, mantendo apenas dois planos de corte.

Usando o modelo para função dual, os autores sugerem então duas modificações do VA e analisam as suas propriedades de convergência. O primeiro método, chamado de Algoritmo do Volume Revisado (RVA), apenas aproxima o modelo de feixe descrito acima. Isso ocorre para evitar a necessidade de se computar a função objetivo dual e seu subgradiente no ponto implícito p_t, o que implicaria na solução de um novo problema linear "fácil" em cada iteração. O método obtido apresenta boas propriedades de convergência; em particular, prova-se que, sob condições favoráveis, a seqüência de pontos implícitos $\{p_t\}$ é maximizante.

A seguir, eles apresentam outra modificação que transforma o algoritmo do volume em um método de feixe, chamado de BVA. Nesse caso, teoremas de convergência conhecidos se aplicam automaticamente [3]. Entretanto, há necessidade de se resolver o problema linear "fácil" no ponto implícito p_t.

Por fim, Bahiense *et.al.* apresentam resultados computacionais para problemas de Steiner obtidos a partir de problemas reais de VLSI. Os resultados são interessantes, com o algoritmo original VA sendo o mais rápido, seguido do RVA e do BVA. Por outro lado, o RVA foi capaz de resolver instâncias mais difíceis, apresentando uma qualidade superior a obtida pelos outros métodos.

3. A Análise Convexa e a Relação Com o Mestrado

A Análise Convexa é a disciplina da Matemática que estuda conjuntos e funções convexas e, em sua versão moderna, remonta aos estudos de Minkowski (1864-1909), Fenchel (1905-1988), Moreau(1923-) e Rockafellar(1935-). Sua compreensão é fundamental no estudo de problemas de otimização, particularmente em teoria de dualidade. Nesse caso, é natural a construção de funções e conjuntos convexos através de processos extremais como o conjunto de minimizadores, e as funções de valor ótimo e o objetivo dual. Esses objetos convexos são então naturalmente não finitos e/ou não suaves, exigindo emprego de noções como cones duais, subdiferenciabilidade e conjugação para sua compreensão.

Conforme visto anteriormente, o projeto de mestrado versará sobre os métodos de feixe e o algoritmo do volume para a solução do problema dual de um problema de programação linear. A função objetivo do problema em questão é convexa, linear por partes e conseqüentemente não suave. Os algoritmos de descida são

[1] Compare com uma tentativa de demonstrar a convergência de um método tipo Cauchy sem o auxílio de uma regra de tamanho de passo como Armijo, Wolfe ou um esquema de regiões de confiança.

portanto baseados em subgradientes e em operações de cálculo generalizado. Assim, as ferramentas da Análise Convexa formam o vocabulário básico para a análise do comportamento de tais métodos.

4. Metodologia do Estudo

O estudo se baseia principalmente no livro *Fundamentals of Convex Analysis* [4] de Jean-Baptiste Hiriart-Urruty e Claude Lemaréchal. Este livro é uma versão reduzida de dois volumes mais clássicos *Convex Analysis and Minimization Algorithms I e II* [3], onde foram extraídos os principais conceitos. Ele é essencialmente teórico, sem lidar com algoritmos numéricos. Tal livro será abordado integralmente durante o semestre, com estudos, discussões e realização dos exercícios. Ao final, será redigido uma monografia com os tópicos estudados.

5. Tópicos a Serem Estudados

Os principais tópicos de estudo são listados abaixo. Os seis itens principais referem-se aos capítulos do nosso livro texto.

0. Introdução: notação e resultados elementares
    - Limitantes inferiores e superiores
    - Conjunto dos números reais extendidos
    - Álgebra linear e bilinear
    - Diferenciabilidade no espaço euclidiano
    - Análise multivalorada
    - Funções convexas com variáveis reais

A. Conjuntos convexos
    - Definições básicas
    - Conjuntos convexos unidos a um conjunto convexo
    - Projeção em conjuntos convexos fechados
    - Separação e aplicações
    - Aproximações cônicas de conjuntos convexos

B. Funções convexas
    - Definições básicas e exemplos
    - Operações funcionais que preservam convexidade
    - Comportamento local e global de funções convexas
    - Diferenciabilidade de primeira e segunda ordem

C. Sublinearidade e funções suporte
    - Funções sublineares
    - Funções suporte de conjuntos não vazios
    - Correspondência entre conjuntos convexos e funções sublineares

D. Subdiferenciais de funções convexas finitas
    - Subdiferenciabilidade: definições e interpretações
    - Propriedades locais de um subdiferencial
    - Regras de cálculo com subdiferenciais
    - O subdiferencial como uma função múltipla

E. Conjugação em análise convexa
- Conjugação convexa de uma função
- Regras de cálculo com operações de conjugação
- Diferenciabilidade de funções de conjugação

Referências

[1] L. Bahiense, N. Maculan, and C. Sagastizábal. The volume algorithm revisited: relation with bundle methods. *Mathematical Programming*, 94:41–69, 2002.

[2] F. Barahona and R. Anbil. The volume algorithm: producing primal solutions with a subgradient method. *Mathematica Programming*, 87:385–399, 2000.

[3] J.-B. Hiriart-Urruty and C. Lemaréchal. *Convex Analysis and Minimization Algorithms*. Springer Verlag, Heidelberg, 1996. Two volumes - 2nd printing.

[4] J.-B. Hiriart-Urruty and C. Lemaréchal. *Fundamentals of Convex Analysis*. Springer Verlag, Heidelberg, 2001.

[5] G.M. Korpolevich. The extragradient method for finding saddle points and other problems. *Matecon*, 12:747–756, 1976.

[6] N. Z. Shor. *Minimization methods for non-differentiable functions*. Springer, 1985.

[7] M.V. Solodov and B.F. Svaiter. A hybrid approximate extragradient-proximal point algorithm using the enlargement of an operator. *Set Valued Analysis*, 24(7):323–345, 1999.

[8] M.V. Solodov and B.F. Svaiter. A hybrid projection-proximal point algorithm. *Journal of Convex Analysis*, 6(1):59–70, 1999.