 ISSN 1984-820X

Complexidade Algorítmica de Métodos de primeira ordem

Adriano R. Delfino*   **Elizabeth W. Karas**
Departamento de Matemática,UFPR,
81531-980, Curitiba, PR
E-mail: adriano.delfino1@gmail.com , ewkaras@ufpr.br.

RESUMO

Neste trabalho, discutiremos problemas de minimização irrestrita de uma função $f : \mathbb{R}^n \to \mathbb{R}$ convexa e diferenciável. Para resolver este problema, certamente existem vários algoritmos ou métodos diferentes na literatura. Mas qual desses métodos é melhor? Qual desses métodos tem um melhor desempenho em uma determinada classe de funções? É possível melhorar um algoritmo de modo que a convergência seja mais rápida?

Resolver o problema significa achar uma solução $\bar{x}$ aproximada usando algum método $\mathcal{M}$ de primeira ordem com uma tolerância $\varepsilon > 0$, ou seja, achar $\bar{x}$ tal que $f(\bar{x}) - f^* < \varepsilon$, onde f^* é o valor ótimo.

Podemos analisar o desempenho de um certo algoritmo contando, por exemplo, o número de iterações, ou o tempo gasto, ou ainda o número de operações aritméticas usado para resolver determinado problema com precisão ε.

Neste trabalho, discutiremos a complexidade algorítmica relacionada com o número de iterações que é requerido para resolver o problema com tolerância ε. Definimos a *cota de complexidade inferior* k de uma classe de funções $\mathcal{P}$ como sendo o número tal que para qualquer método $\mathcal{M}$, existe uma função $\bar{f} \in \mathcal{P}$ tal que o método $\mathcal{M}$ gasta pelo menos k iterações para minimizar $\bar{f}$ com precisão ε.

Já para um determinado método $\mathcal{M}$ utilizado para minimizar uma classe de funções $\mathcal{P}$, definimos sua *cota de complexidade superior* como sendo o número máximo de iterações que o método $\mathcal{M}$ gasta para minimizar qualquer função da classe $\mathcal{P}$ com precisão ε. Um método $\mathcal{M}$ é considerado ótimo para minimizar uma classe de funções, quando a sua cota de complexidade superior é proporcional à cota de complexidade inferior da classe de funções.

Estamos particularmente interessados na complexidade algorítmica de métodos de primeira ordem para minimizar funções f pertencentes à classe de funções convexas ($\mathcal{F}$) e fortemente convexas ($\mathcal{S}$), ambas com o gradiente Lipschitz, ou seja, existe $L > 0$ tal que para todo $x, y \in \mathbb{R}^n$,

$$\|\nabla f(x) - \nabla f(y)\| \leq L\|x - y\| \ .$$

Nesterov mostra em [2] que a cota de complexidade inferior para a classe $\mathcal{F}$ é $\sqrt{\frac{3L\|x_0-x^*\|^2}{32\varepsilon}} - 1$ onde $x_0 \in \mathbb{R}^n$ é o ponto inicial e x^* é o minimizador de f. Nesterov também mostra que $\frac{\sqrt{Q_f}-1}{4}[ln\frac{1}{\varepsilon} + ln\frac{\mu}{2} + 2ln\|x_0 - x^*\|]$ é a cota de complexidade inferior para a classe $\mathcal{S}$, onde $\mu > 0$ é o parâmetro de convexidade e $Q_f = \frac{L}{\mu} > 1$. Por outro lado, a cota de complexidade superior do método do gradiente é $\frac{2L\|x_0-x^*\|^2}{\varepsilon} - 4$, se $f \in \mathcal{F}$. Isto garante que o método do gradiente não é ótimo para esta classe de funções.

---
*bolsista do Programa de Pós Graduação em Matemática Aplicada Prof/Capes

Nesterov propõe em [2] um método ótimo para as classes $\mathcal{F}$ e $\mathcal{S}$, ou seja, cuja cota de complexidade superior é proporcional à cota inferior destas classes. Nesterov mostra que a complexidade desse método para $f \in \mathcal{S}$ é $\sqrt{Q_f}[ln\frac{1}{\varepsilon} + lnL + 2ln\|x_0 - x^*\|]$, ou seja, o termo principal $\sqrt{Q_f}ln\frac{1}{\varepsilon}$ é proporcional à cota de complexidade de $\mathcal{S}$.

O método proposto depende do conhecimento prévio da constante de Lipchitz L para o gradiente. Gonzaga e Karas [1] e Nesterov [4] obtém algoritmos ótimos com a mesma complexidade do método proposto por Nesterov, mas que não exige o conhecimento prévio de L. Ambos os artigos [1, 4] foram feitos independentemente um do outro.

Nosso objetivo é estudar com detalhes esses algoritmos, implementá-los e compará-los. Nossa intenção é aplicá-los em otimização de funções não suaves [3] e otimização semi-definida [5].

Palavras-chave: *Otimização Convexa, Teoria de Complexidade, Métodos Ótimos*

Referências

[1] C. C. Gonzaga and E. W. Karas. Optimal steepest descent algorithms for unconstrained convex problems: Fine tuning Nesterov's method. Technical report, Dep. Mathematics, Federal University of Paraná, Brazil, 2008.

[2] Y. Nesterov. *Introductory Lectures on Convex Optimization. A basic course.* Kluwer Academic Publishers, Boston, 2004.

[3] Y. Nesterov. Smooth minimization of non-smooth functions. *Mathematical Programming*, 103(1):127 – 152, 2005.

[4] Y. Nesterov. Gradient methods for minimizing composite objective function. Discussion paper 76, CORE, UCL, Belgium, 2007.

[5] Y. Nesterov. Smoothing technique and its applications in semidefinite optimization. *Mathematical Programming*, 110(2):245 – 259, 2007.