Possíveis projetos de Otimização Irrestrita para a disciplina CM106 - Otimização I

Abel Soares Siqueira

1 Otimização Irrestrita

Na otimização irrestrita, estamos buscando um minimizador local de $f(x)$, $f : \mathbb{R}^n \to \mathbb{R}$. Existem alguns métodos para esse problema, e podemos considerar casos específicos deste problema.

1.1 Métodos para minimização irrestrita geral

Para

$$\min \quad f(x),$$

estudamos alguns métodos, com algumas técnicas de busca linear. São eles:

→ Método do Gradiente;

→ Método de Newton;

→ Método de Newton Modificado;

→ Métodos Quase-Newton;

com busca

→ Exata, utilizando o Método da Seção Áurea;

→ Inexata, utilizando Armijo com Backtracking;

→ Inexata, utilizando Wolfe.

Algumas combinações destes métodos são triviais de se implementar, e já fizemos, porém algumas são possíveis de gerar projeto:

- Gradiente com Seção Áurea, mostrando a dificuldade do método em encontrar soluções rapidamente;
- Newton Modificado com Armijo, utilizando a decomposição de Cholesky para a solução do sistema;

- BFGS para a inversa da Hessiana com memória limitada e Armijo, ignorando direções com $s^T y \leqslant 0$. Neste problema, o problema será a implementação do BFGS com memória limitada;
- BFGS para a inversa da Hessiana com Wolfe. Para Wolfe, não se usa backtracking, então a implementação de Wolfe que precisa ser estudada, e porque o BFGS funciona bem. Aqui pode-se usar um código pronto para o BFGS.

1.2 Minimização sem derivadas

Se f não é diferenciável, ou suas derivadas não são fáceis de se obter, podemos utilizar métodos sem derivadas. Uma alternativa seria aproximar a derivada utilizando diferenças finitas ou outras técnicas, mas neste caso estamos considerando métodos que não aproximam a derivada.

O método mais básico destes é o de Busca Direcional: Dado x^0, à cada iteração olhamos nas n direções coordenadas, com uma distância $\alpha_k > 0$. Se algum dos pontos for menor, seguimos nessa direção, caso contrário, diminuímos α_k e repetimos. Se α_k ficar pequeno o suficiente, paramos o método e declaramos convergência.

1. Dado x^0, $\alpha_0 > 0$, $k = 0$
2. Se existe $d \in \mathcal{D} = \{e_1, \dots, e_n, -e_1, \dots, -e_n\}$, tal que $f(x^k + \alpha_k d) < f(x^k)$, definimos $x^{k+1} = x^k + \alpha_k d$, incrementamos k e repetimos 2.
3. Caso contrário, nenhuma direção fez decréscimo, então escolhemos $0 < \alpha_{k+1} < \alpha_k$.
4. Incremente k e vá à 2.

Esse método é bastante simples, e uma variação também simples é considerar um conjunto $\mathcal{D}$ diferente que gere positivamente o $\mathbb{R}^n$, por exemplo $\{e_1, \dots, e_n, -\frac{1}{n}\sum_{i=1}^n e_i\}$.

Um outro método bastante importante, consideravelmente melhor que o método acima, é o método de Nelder-Mead, que consiste em utilizar um conjunto de $n+1$ pontos do $\mathbb{R}^n$, remover o pior destes pontos, substituindo por outro ponto, ou possivelmente "encolher" o conjunto, aproximando todos os pontos. O método é bastante visual, e de fácil implementação, por isso é um favorito.

O possível projeto para esta parte seria

- Implementar o método de buscas direcionais e o Nelder-Mead, e compará-los.

1.3 Quadrados Mínimos Não-Lineares

O problema de quadrados mínimos não-lineares acontece quando f tem uma forma específica:

$$f(x) = \|F(x)\|^2 = \sum_{i=1}^{n} F_i(x)^2,$$

onde $F : \mathbb{R}^n \to \mathbb{R}^m$, e F_i é a i-ésima componente de F. Um caso particular deste problema é $F(x) = Ax - b$, A matriz e b vetor. Esse problema se chama quadrados mínimos lineares, e é um caso particular de programação quadrática. Um dos problemas com esse tipo de função é que a sua Hessiana contém uma soma de Hessianas,

$$\nabla^2 f(x) = J(x)^T J(x) + \sum_{i=1}^{n} F_i(x) \nabla^2 F_i(x),$$

e isso é muito caro computacionalmente, e não agrega muito à função. Dessa maneira, o método de Newton fica mais lento, e perde muito de sua força. Uma alternativa é aproximar o somatório por μI, onde $\mu \geqslant 0$ é um parâmetro adaptado a cada iteração.

O projeto para esse tipo de problema seria:

- Explicar e implementar o método acima, usando diferentes μ.

1.4 Newton Inexato

O Método de Newton com busca linear consiste em encontrar a solução de

$$\nabla^2 f(x^k) d_k = -\nabla f(x^k),$$

e atualizar $x^{k+1} = x^k + t_k d_k$, com t_k obtido por busca linear. No entanto, $\nabla^2 f(x^k)$ pode ser definida negativa, e além disso, é trabalhoso resolver o sistema. Uma possibilidade de resolução é considerar o método de Gradientes Conjugados para resolver o sistema de Newton, mas parar com uma tolerância maior. Se paramos antes, fazemos menos iterações do Gradientes Conjugados, e convergimos mais rápido. Além disso, se encontrarmos $d^T \nabla^2 f(x^k) d \leqslant 0$ em alguma iteração do método de Gradientes Conjugados, paramos a execução e achamos a direção encontrada até o momento.

O projeto seria:

- Implementação do Newton Inexato com busca de Armijo.