Aluno: André Luiz Correa Vianna Filho
Tema: Região de Confiança

Resumo

Para resolver problemas de otimização irrestritos pode ser usada a estratégia de, a cada iteração, aproximar a função localmente por um modelo, e então minimizar este modelo com o objetivo de encontrar um ponto no qual o valor da função seja menor. Neste trabalho será usado o modelo quadrático. Considere a função $f : \mathbb{R}^n \to \mathbb{R}$ de classe C^2 e o modelo quadrático em torno do ponto x_k da forma:

$$m_k(x) = f(x_k) + \nabla f(x_k)^T(x - x_k) + \frac{1}{2}(x - x_k)^T B_k(x - x_k), \tag{1}$$

onde B_k é uma matriz simétrica. Se $B_k = \nabla^2 f(x_k)$ temos a aproximação de Taylor de segunda ordem, e o método de Newton.

O método de Newton puro em geral depende muito do ponto inicial, mas há formas de contornar esta dificuldade, uma delas é utilizar busca linear. A busca linear usa o modelo para determinar uma boa direção, ou seja, uma direção de descida, e então calcula o tamanho do passo que resulta em uma redução satisfatória no valor da função, que pode ser determinada com o uso da condição de Armijo. Outra alternativa são os chamados métodos de região de confiança. Estes métodos consistem em determinar uma região na qual o modelo é de fato uma boa aproximação da função, seja Δ_k o tamanho de tal região, e então resolver a cada iteração:

$$\begin{cases} \min m_k(x_k + p) \\ s.a. \ ||p|| \leq \Delta_k \end{cases} \tag{2}$$

Neste caso, a direção e o tamanho do passo são calculados ao mesmo tempo. O tamanho da região deve ser atualizado a cada iteração, porque se o modelo não aproxima bem a função na região, o tamanho deve ser reduzido, e (2) deve ser resolvido novamente. E caso a aproximação seja muito boa, a solução de (2) deve ser aceita e o tamanho da região aumentado para permitir que o método faça avanços mais significativos.

Para os métodos de região de confiança, dados o ponto inicial e o tamanho da região, temos basicamente o seguinte formato:

1. Calcular $f(x_k)$ e o modelo $m_k(x_k + p)$;
2. Encontrar o passo p_k resolvendo (2);
3. Aceitar ($x_{k+1} = x_k + p_k$) ou rejeitar o passo ($x_{k+1} = x_k$) e atualizar o tamanho da região de acordo com a qualidade da solução de (2);
4. Se a condição de parada não for satisfeita, voltar ao item 1.

Em geral, a diferença entre os métodos de região de confiança está no item 2, no método utilizado para resolver (2). É comum utilizar o método dos gradientes conjugados.

Neste trabalho será usado o método de região de confiança para resolver problemas com restrição do tipo caixa. Seja f a função definida acima, e sejam $x = (x_1, \ldots, x_n)$, $l = (l_1, \ldots, l_n)$, e $u = (u_i, \ldots, u_n)$ vetores de $\mathbb{R}^n$, o problema é da forma:

$$\begin{cases} \min f(x) \\ s.a. \ l_i \leq x_i \leq u_i \end{cases} \tag{3}$$

O objetivo é mostrar como estas restrições são tratadas na resolução de (3) e comparar dois métodos: *TRON* e *LANCELOT*. Mais precisamente, comparar a forma como os métodos resolvem (2). Esta comparação é motivada pela comparação feita em [1], na qual, devido à diferença na forma de resolver (2), *TRON* precisava de muito menos execuções do método dos gradientes conjugados do que *LANCELOT*, resultando em um tempo de execução bem melhor.

Referências

[1] LIN, C.; MORÉ, J. *Newton's Method for Large Bound-Constrained Optimization Problems*. SIAM, 1991.

[2] SARTENAER, A. *Large-Scale Nonlinear Optimization and the LANCELOT Package*. 1995.

[3] NOCEDAL, J.; WRIGHT, S. *Numerical Optimization*. Springer, Segunda edição, 2006.