UM MÉTODO QUASE-NEWTON BASEADO EM REFLEXÕES DE HOUSEHOLDER PARA OTIMIZAÇÃO IRRESTRITA

Kauan Carlos Rigoleto Fernandes (PIC/UEM), Emerson Vitor Castelani (Orientador).
E-mail: evcastelani@uem.br.

Universidade Estadual de Maringá, Centro de Ciências Exatas, Maringá, PR.

Ciências Exatas e da Terra / Matemática Aplicada.

Palavras-chave: otimização irrestrita; equação secante; householder quase-newton.

RESUMO

Este trabalho avalia um método quase-Newton para otimização irrestrita baseado em reflexões de Householder, o qual constrói a direção de busca sem atualizar ou armazenar explicitamente a matriz Hessiana, satisfazendo a equação secante de forma implícita. Para contornar a ausência de garantias de descida na direção pura, foram desenvolvidas variações com salvaguardas. Em testes com 197 problemas do repositório CUTEst (de dimensões $2 \leq \square \leq 1000$), a variante com salvaguarda por gradiente obteve 52,3% de convergência, superando o método de Cauchy (49,2%), embora abaixo de BFGS (75,1%) e LBFGS (92,9%). Os resultados confirmam a viabilidade numérica e o baixo custo por iteração do método, apesar da perda de eficiência em problemas de maior porte.

INTRODUÇÃO

Métodos de otimização irrestrita baseados em informação de segunda ordem, como o método de Newton, utilizam a matriz Hessiana da função objetivo para construir direções de busca que, sob condições adequadas, produzem convergência superlinear ou quadrática nas proximidades de um mínimo local (NOCEDAL; WRIGHT, 2006). Entretanto, o cálculo exato da Hessiana é frequentemente inviável ou custoso em problemas de grande porte, o que motivou o desenvolvimento dos métodos quase-Newton. Os métodos quase-Newton constroem, a cada iteração, uma aproximação da Hessiana (ou de sua inversa) utilizando apenas informações de gradiente sucessivas, impondo que essa aproximação satisfaça a chamada equação secante. O método BFGS é o representante mais difundido dessa classe, atualizando a aproximação por meio de correções de posto baixo (NOCEDAL; WRIGHT, 2006). Neste trabalho, propõe-se uma via alternativa para obter uma direção de busca que satisfaça implicitamente a equação secante: em vez de atualizar uma matriz de aproximação da Hessiana, constrói-se uma reflexão de Householder (GOLUB; VAN LOAN, 2013) a partir dos vetores normalizados de deslocamento e de variação do gradiente entre duas iterações consecutivas. Essa reflexão define uma direção candidata sem exigir o armazenamento ou a atualização explícita de uma matriz completa. O objetivo deste estudo é investigar a viabilidade

numérica dessa abordagem frente a métodos consolidados, como o gradiente com busca de seção áurea e o LBFGS.

MATERIAIS E MÉTODOS

Os métodos foram implementados na linguagem Julia (BEYER; BINDER; VOSS, 2017), utilizando o pacote CUTEst (GOULD; ORBAN; TOINT, 2015) para acesso aos problemas-teste e busca linear de Armijo para o controle do tamanho do passo.
A cada iteração, sejam $\square = \square_{\square+1} - \square_{\square}$ e $\square = \square\square(\square_{\square+1}) - \square\square(\square_{\square})$. Definem-se os vetores unitários $\square_{\square} = \square/||\square||$ e $\square_{\square} = \square/||\square||$, e o vetor de reflexão $\delta = \square_{\square} - \square_{\square}$, $\square = \square/||\square||$. A partir de $\square$, constrói-se o termo de reflexão $\eta = 2\square(\square^{\square}\square\square)$, utilizado para formar a direção candidata $\square = \square(\square - \square\square)$, com $\alpha = ||\square||/||\square||$, obtida de tal forma que, sob a hipótese $\alpha > 0$, a matriz implícita associada a $\square$ satisfaz a equação secante $\square\square = \square$. Como nem sempre $\square$ resulta em uma direção de descida (isto é, $\nabla\square^{\square}\square < 0$), duas variações foram estudadas:
• **Householder "puro":** utiliza $\square$ diretamente, sem qualquer verificação de descida, com passo fixo (padrão quase-Newton).
• **Householder com salvaguardas:** quando $\nabla\square^{\square}\square \geq 0$, temos duas classes de salvaguarda, a primeira recorre-se a direções de reserva construídas a partir de combinações da reflexão com a identidade para basicamente garantir que a matriz seja positiva definida dado que temos a decomposição espectral do refletor de householder bem caracterizada ($\square = \square\square - (\square + \square)\square\square$, para $\gamma \in \{1{,}0;\ 1{,}1\}$ ou, de forma mais simples, a segunda salvaguarda retorna a direção do gradiente negativo.
Esses métodos foram comparados com o método de Cauchy (gradiente com busca linear por seção áurea) e o LBFGS do pacote JSOSolvers, em problemas selecionados do CUTEst sem restrições. Foram feitos vários testes com parâmetros diferentes, porém os selecionados para este documento utilizaram como critério de parada $\square\square\square_{\square\square\square\square} = 20.000$ ou $||\square\square|| \leq 10^{-4}$ num geral de 197 problemas irrestritos de dimensões maiores ou iguais que 2 até menores ou iguais que 1000.

RESULTADOS E DISCUSSÃO

A avaliação comparativa abrangeu taxa de convergência (robustez), contagem de iterações, avaliações de função e desempenho relativo em relação à dimensão dos problemas. A Tabela 1 sintetiza os principais indicadores obtidos para cada algoritmo.

Tabela 1: Comparação de desempenho dos métodos no repositório CUTEst (197 problemas).

Método	Convergência Global	Pequenos (n ≤ 6)	Grandes (n > 6)	Iterações (Mediana)	Aval. f (Mediana)	Tempo (s) (Mediana)

Método	Convergência Global	Pequenos (n ≤ 6)	Grandes (n > 6)	Iterações (Mediana)	Aval. f (Mediana)	Tempo (s) (Mediana)
LBFGS	92,9% (183/197)	93,1%	92,7%	31,0	48,0	0,0002
BFGS	75,1% (148/197)	85,1%	64,6%	30,5	92,0	0,0001
HH Gradiente	52,3% (103/197)	65,3%	38,5%	186,0	839,0	0,0014
Cauchy	49,2% (97/197)	60,4%	37,5%	117,0	787,0	0,0005
HH Identidades	43,7% (86/197)	55,4%	31,2%	160,5	541,0	0,0003
HH Puro	37,1% (73/197)	45,5%	28,1%	22,0	24,0	0,0000

A inclusão de salvaguardas foi determinante para estabilizar a reflexão de Householder. A variante Householder Puro sofreu 25 falhas decorrentes da geração de valores não numéricos (NaN) pela ausência do critério de descida. Com a salvaguarda por gradiente (HH Gradiente), os erros não numéricos foram eliminados e a taxa de convergência subiu para 52,3%, superando o método de Cauchy (49,2%).

Note-se uma perda de eficiência em problemas de dimensão maior ($n > 6$) para todas as variantes de Householder (caindo de ~65% em problemas pequenos para ~38% em problemas maiores). Por outro lado, nos casos em que convergiu, o HH Puro apresentou a menor mediana de iterações (22,0) e de avaliações de função (24,0) de todo o benchmark, indicando reduzido custo computacional por iteração convergente.

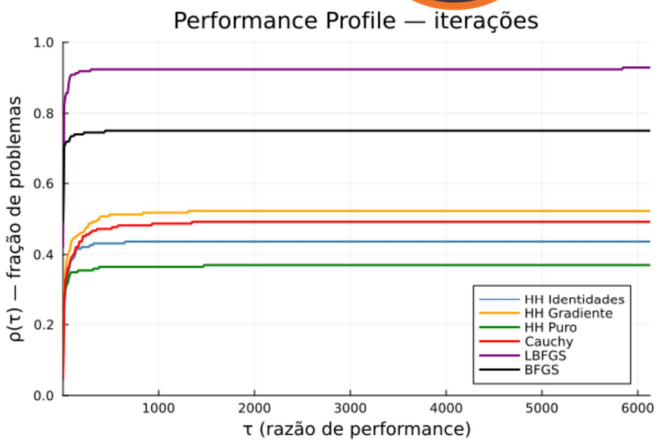


Figura 1: Perfil de Desempenho (Performance Profile) em relação ao número de iterações para os métodos comparados no repositório CUTEst.

O perfil de desempenho (Figura 1) evidencia a superioridade global do LBFGS e do BFGS em termos de robustez e velocidade. Contudo, confirma que a variante HH Gradiente é competitiva com o método de Cauchy, posicionando-se acima deste na maioria dos patamares da razão de performance.

CONCLUSÕES

Conclui-se que o método proposto é numericamente viável como alternativa de baixo custo computacional para a geração de direções quase-Newton, dispensando o armazenamento explícito de matrizes. Embora sua robustez global seja inferior à dos métodos consolidados (como LBFGS e BFGS), a variante HH Gradiente superou o método do gradiente de Cauchy em taxa de convergência (52,3% contra 49,2%), demonstrando potencial prático que pode ser aprimorado em pesquisas futuras por meio de novas estratégias de salvaguarda.

AGRADECIMENTOS

Agradeço à UEM, ao DMA, ao meu orientador, Prof. Dr. Emerson Vitor Castelani e ao Prof. Dr. Ronaldo Lopes pelas contribuições.

REFERÊNCIAS

BEYER, A.; BINDER, M.; VOSS, C. *Julia: A Fresh Approach to Numerical Computing.* 2017. Disponível em: https://julialang.org. Acesso em: 31/08/2026

GOLUB, G. H.; VAN LOAN, C. F. *Matrix Computations.* 4. ed. Baltimore: Johns Hopkins University Press, 2013.

GOULD, N. I. M.; ORBAN, D.; TOINT, PH. L. CUTEst: a constrained and unconstrained testing environment with safe threads for mathematical optimization.

Computational Optimization and Applications, New York, v. 60, n. 3, p. 545-557, 2015.

NOCEDAL, J.; WRIGHT, S. J. *Numerical Optimization.* 2. ed. New York: Springer, 2006.