Comparison of Optimization Algorithms for the Kowalik Problem

Doğuhan Kayra Güneş
Department of Mathematical Engineering
Yildiz Technical University
Istanbul, Türkiye
kayra.gunes@std.yildiz.edu.tr

***Abstract*—This paper presents a numerical comparison of four optimization algorithms – Newton-Raphson, Hestenes-Stiefel, Polak-Ribière, and Fletcher-Reeves – applied to the Kowalik benchmark problem. The algorithms were implemented in MATLAB and tested with three different initial guesses. The performance of each algorithm was evaluated based on the number of iterations, execution time, and the final function value achieved. The impact of the initial guess on the convergence behavior of each algorithm is also discussed.**

***Index Terms*—Optimization algorithms, Kowalik problem, Newton-Raphson, Conjugate Gradient, Hestenes-Stiefel, Polak-Ribière, Fletcher-Reeves, Numerical comparison, MATLAB**

I. Introduction

The optimization of nonlinear functions is a crucial task in various fields of engineering and science. To evaluate and compare the effectiveness of different optimization algorithms, benchmark functions with known properties are often employed.

This study focuses on the Kowalik problem [1], a well-known least-squares benchmark function, and investigates the performance of four distinct optimization algorithms: the second-order Newton-Raphson method and three first-order conjugate gradient methods (Hestenes-Stiefel, Polak-Ribière, and Fletcher-Reeves).

The algorithms were implemented in MATLAB, and their performance was assessed based on their ability to find the minimum of the Kowalik function from different starting points. This report details the implementation, results, and a comparative analysis of these algorithms.

II. Kowalik Problem

The Kowalik problem [1] is a benchmark function defined as the sum of squares of eleven nonlinear terms:

$$f(x) = \sum_{i=1}^{11} \left(a_i - \frac{x_1(1 + x_2 b_i)}{1 + x_3 b_i + x_4 b_i^2} \right)^2$$

where $x = [x_1, x_2, x_3, x_4]^T$ are the optimization variables, and a_i and b_i are constants with values provided in [1], Table 7. The problem is a four-dimensional optimization problem with a known global minimum value of approximately 3.0748×10^{-4} around the point $x^* \approx [0.1928, 0.1908, 0.1231, 0.1358]^T$. The domain of interest for the variables is typically $0 \leq x_i \leq 0.42$ for $i = 1, 2, 3, 4$.

TABLE I
Constants for the Kowalik Problem

i	a_i	b_i
1	0.1957	4.0
2	0.1947	2.0
3	0.1735	1.0
4	0.1600	0.5
5	0.0844	0.25
6	0.0627	0.167
7	0.0456	0.125
8	0.0342	0.1
9	0.0323	0.0833
10	0.0235	0.0714
11	0.0246	0.0625

III. Implemented Optimization Algorithms

A. Newton-Raphson

The Newton-Raphson method is an iterative optimization algorithm that uses the gradient and the Hessian matrix of the objective function to find stationary points. The update rule is given by:

$$x_{k+1} = x_k - H_k^{-1} \nabla f(x_k)$$

where $\nabla f(x_k)$ is the gradient and H_k is the Hessian matrix of the objective function at the current point x_k.

B. Hestenes-Stiefel

The Hestenes-Stiefel (HS) method is a conjugate gradient method where the search direction p_k is updated using the previous search direction p_{k-1} and a scalar β_k^{HS}:

$$p_k = -\nabla f(x_k) + \beta_k^{HS} p_{k-1}$$

The parameter β_k^{HS} is calculated as:

$$\beta_k^{HS} = \frac{\nabla f(x_k)^T (\nabla f(x_k) - \nabla f(x_{k-1}))}{(\nabla f(x_{k-1}) - \nabla f(x_{k-2}))^T p_{k-1}}$$

with $p_0 = -\nabla f(x_0)$.

C. Polak-Ribière

The Polak-Ribière (PR) method is another conjugate gradient method with a different formula for β_k^{PR}:

$$p_k = -\nabla f(x_k) + \beta_k^{PR} p_{k-1}$$

where

$$\beta_k^{PR} = \frac{\nabla f(x_k)^T(\nabla f(x_k) - \nabla f(x_{k-1}))}{\|\nabla f(x_{k-1})\|^2}$$

again with $p_0 = -\nabla f(x_0)$. To ensure descent, β_k^{PR} is often taken as $\max(0, \beta_k^{PR})$.

D. Fletcher-Reeves

The Fletcher-Reeves (FR) method is a classic conjugate gradient method where β_k^{FR} is defined as:

$$p_k = -\nabla f(x_k) + \beta_k^{FR} p_{k-1}$$

with

$$\beta_k^{FR} = \frac{\|\nabla f(x_k)\|^2}{\|\nabla f(x_{k-1})\|^2}$$

and $p_0 = -\nabla f(x_0)$.

IV. Numerical Results and Discussion

The four optimization algorithms were implemented in MATLAB and tested with three different initial guesses, generated randomly within the bounds $[0, 0.42]^4$. The stopping criteria were set to $||\nabla f(x_k)|| \leq 10^{-4}$ and $|f(x_{k+1}) - f(x_k)| \leq 10^{-4}$, with a maximum of 5000 iterations. The numerical results for each algorithm and initial guess are summarized in the above tables.

TABLE II
Benchmarking Results for Newton-Raphson

Initial Guess	Iterations	Execution Time (s)	Final $f(x)$
1	5000	31.7569	1.7311×10^{-3}
2	5000	32.0646	9.0897×10^{-4}
3	6	0.0100	3.0749×10^{-4}

TABLE III
Benchmarking Results for Hestenes-Stiefel

Initial Guess	Iterations	Execution Time (s)	Final $f(x)$
1	168	0.1493	3.0754×10^{-4}
2	81	0.0658	3.0795×10^{-4}
3	104	0.0946	3.0755×10^{-4}

TABLE IV
Benchmarking Results for Polak-Ribière

Initial Guess	Iterations	Execution Time (s)	Final $f(x)$
1	163	0.2922	3.0822×10^{-4}
2	87	0.1665	3.0886×10^{-4}
3	209	0.3864	3.0877×10^{-4}

TABLE V
Benchmarking Results for Fletcher-Reeves

Initial Guess	Iterations	Execution Time (s)	Final $f(x)$
1	626	0.8942	3.0806×10^{-4}
2	204	0.2528	3.0767×10^{-4}
3	513	0.7490	3.0768×10^{-4}

The number of iterations and execution times varied significantly across the algorithms and initial guesses. Newton-Raphson showed very fast convergence (6 iterations) for one initial guess but failed to converge within the maximum iterations for the other two, indicating its sensitivity to the starting point. The conjugate gradient methods (Hestenes-Stiefel, Polak-Ribière, and Fletcher-Reeves) generally took more iterations to converge but were more consistent in reaching a final function value close to the known global minimum. The execution time per iteration was higher for Newton-Raphson due to the Hessian computation.

The convergence behavior clearly depends on the initial conditions. For Newton-Raphson, a good initial guess led to rapid convergence, while poor initial guesses resulted in stagnation. The conjugate gradient methods were more robust to the initial guesses tested, consistently finding a good approximation of the minimum, although the number of iterations varied.

The trade-off between the algorithms is evident. Newton-Raphson can be very efficient if it converges, but its reliability is questionable without a good starting point. Conjugate gradient methods are generally more robust for this problem, providing a good balance between computational cost per iteration and the number of iterations required.

Tighter stopping criteria would likely increase the number of iterations for all algorithms and potentially prevent premature convergence to a suboptimal point. Looser criteria might lead to faster termination but with a less accurate solution.

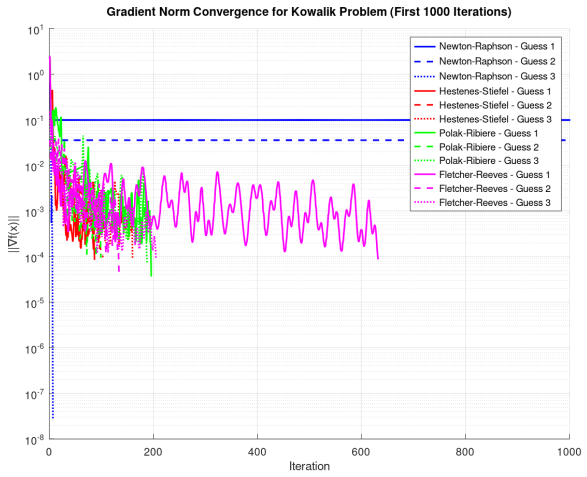

Fig. 1. Function Value Convergence for Kowalik Problem

Figure 1 shows the convergence of the function value over iterations for each algorithm and initial guess. Newton-Raphson exhibits a sharp decrease initially in one case but plateaus at higher values in others. The conjugate gradient methods show a more gradual but consistent decrease towards lower function values.

Figure 2 illustrates the convergence of the gradient norm. For Newton-Raphson, the gradient norm decreases rapidly in the successful case but stagnates at a higher level in the unsuccessful runs. The conjugate gradient methods show a more steady decrease in the gradient norm, eventually reaching the stopping tolerance.

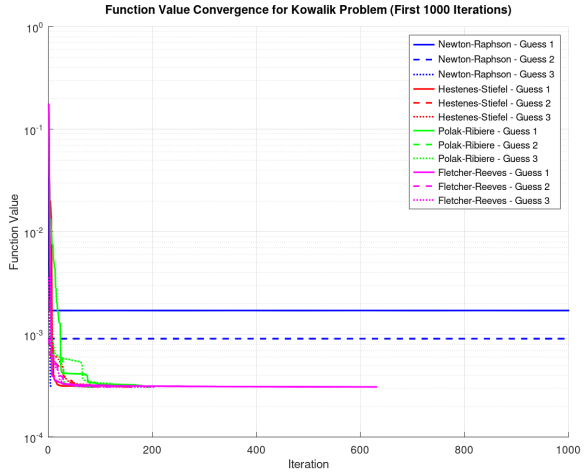

Fig. 2. Gradient Norm Convergence for Kowalik Problem

V. Conclusion

This study compared the performance of Newton-Raphson and three conjugate gradient optimization algorithms on the Kowalik problem. The numerical results demonstrated that the conjugate gradient methods (Hestenes-Stiefel, Polak-Ribière, and Fletcher-Reeves) were more robust in consistently converging to a solution close to the global minimum from different initial guesses, although they often required a larger number of iterations compared to Newton-Raphson in the successful case. Newton-Raphson exhibited fast convergence for one specific initial guess but struggled with others, highlighting its sensitivity to the starting point and potential issues with local minima or the Hessian matrix. The choice of optimization algorithm depends on the specific problem characteristics and the desired balance between convergence speed, computational cost per iteration, and robustness to the initial guess.

References

[1] M. M. Ali, C. Khompatraporn, and Z. B. Zabinsky, "A numerical evaluation of several stochastic algorithms on selected continuous global optimization test problems," *Journal of Global Optimization*, vol. 31, no. 4, pp. 635–672, 2005.

[2] IEEE, "Manuscript Templates for Conference Proceedings." [Online]. Available: https://www.ieee.org/conferences/publishing/templates.html.