HIGH SCHOOL EDITION
Journal of Student Research
Volume 12 Issue 4 (2023)

A Comparison of Nonlinear Optimization Methods for Rosenbrock, Booth, and Matyas Functions

Niah Goudar[1] and Kristen Skaff[1#]

[1]Newbury Park High School
[#]Advisor

ABSTRACT

There is increasing interest in the application of nonlinear optimization techniques across a broad range of real-world problems that is linked to rapid advances in novel technologies and renewed interest in areas like sustainable development. While many algorithms for nonlinear optimization have been developed, they differ widely in their methods of convergence to the solution. Consequently, matching the function to be optimized with the right optimization algorithm is essential. To further explore this relationship, we studied two nonlinear optimization methods – the Nelder-Mead simplex method that only performs function evaluations and the quasi-Newton method which requires estimation of derivatives. These algorithms were used to find the global minima of the Rosenbrock, Booth, and Matyas functions from multiple starting points. The convergence paths for each starting point across both optimization methods for all 3 functions were visualized on contour plots. While convergence on the global minima was observed in all instances, our analysis indicated that the quasi-Newton method was consistently more efficient and needed fewer iterations than the Simplex method. This was especially pronounced for the Booth and Matyas functions where ~10-fold and 20-fold fewer iterations, respectively, were necessary for convergence. Our analysis reinforces the need to carefully match properties of the function being minimized with the performance characteristics of the optimization approach to obtain fast convergence on the global minimum.

Introduction

The general approach of optimization has broad practical applications across many disciplines ranging from engineering to finance (Bartholomew-Biggs, 2005, 2008). As many real-world systems are nonlinear, optimization of nonlinear functions becomes important to minimize waste, maximize efficiency, and to design an optimal manufacturing plant for the product of choice (Cui et al., 2017; Edgar & Himmelblau, 1987). Extensive work has been done in the area of nonlinear optimization and many different approaches are available to solve these problems (Beck, 2014).

Every optimization problem has a function that needs to be optimized and a choice must be made on the optimization method that will be applied to the function. Also, because nonlinear optimization is iterative, the starting point needs to be specified using which the optimization method will advance towards the solution. It is important to choose starting points that are close to the global minimum of the function being optimized. Starting with initial values that are far from the global minimum can lead to the optimization method converging on undesirable local minima and thus lead to an incorrect solution.

To better understand the performance of optimization methods, three functions with different properties were chosen. They included the Rosenbrock function (Rosenbrock, 1960) which is non-convex with a global minimum of (1, 1) located in a long, narrow, and flat parabolic (banana-type) valley (Figure 1). The second was the Booth function (Salleh et al., 2022) characterized by a long flat valley (Figure 2) with the global minimum located at (1, 3). Finally, the Matyas function (Wang & Luo, 2021) was chosen for its plate shaped surface (Figure 3) and has a global minimum at (0, 0). The diversity and uniqueness of surface properties across these three functions make them good candidates to evaluate optimization algorithms.

A range of initial estimates were evaluated around the true global minima for the 3 functions and the number of iterations necessary for convergence was determined for the two nonlinear optimization approaches. While convergence was reached in all cases, substantial differences in the performance of the optimization methods were observed, reinforcing the need to carefully match functional properties with the optimization algorithm.

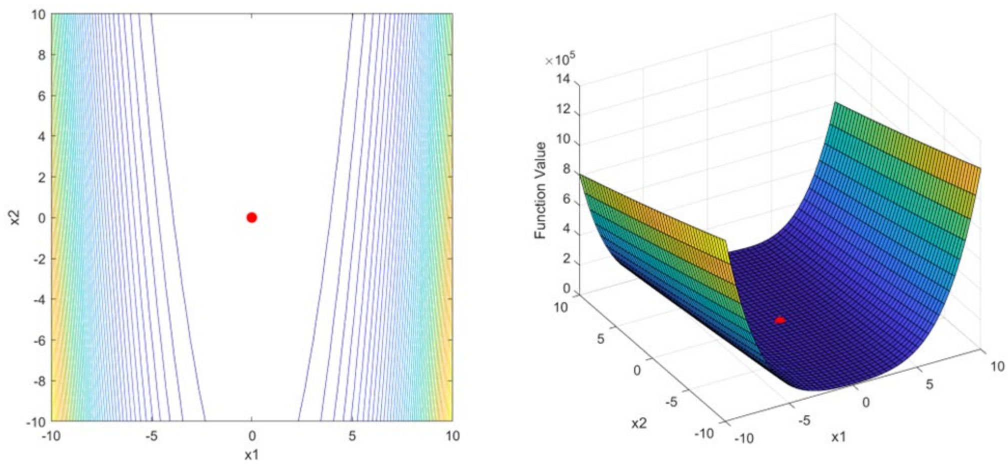

Figure 1. Contour and surface plots for the Rosenbrock function for x_1 and x_2 values in the -10 to 10 range. The function minimum of 0 corresponding to $x_1 = 1$ and $x_2 = 1$ is shown as a red dot.

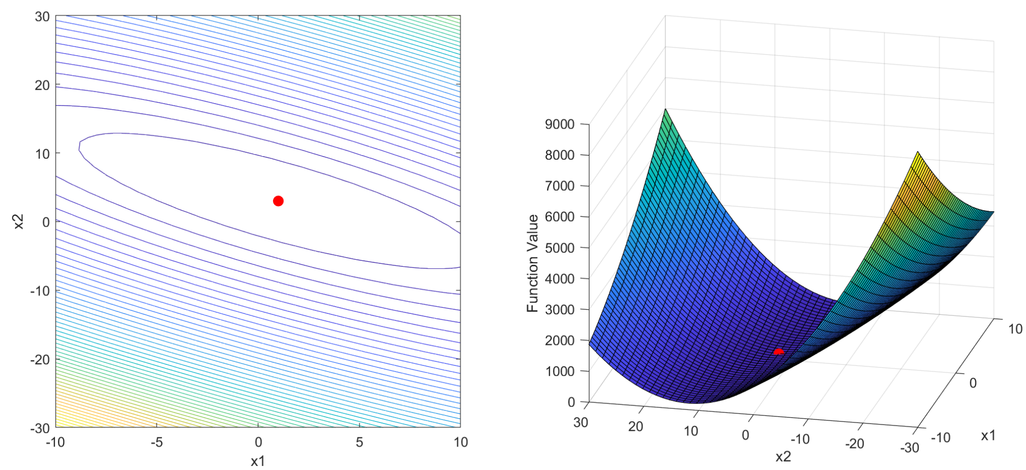

Figure 2. Contour and surface plots for the Booth function for x_1 and x_2 values in the -10 to 10 and -30 to 30 ranges, respectively. The function minimum of 0 corresponding to $x_1 = 1$ and $x_2 = 3$ is shown as a red dot.

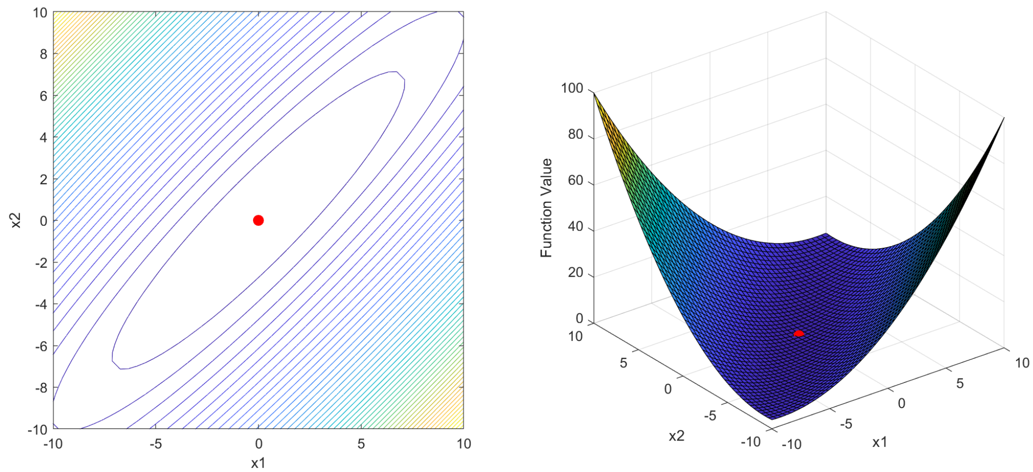

Figure 3. Contour and surface plots for the Matyas function for x_1 and x_2 values in the -10 to 10 range. The function minimum of 0 corresponding to $x_1 = 0$ and $x_2 = 0$ is shown as a red dot.

Methods

Optimization Algorithms Tested

As the purpose of this investigation was to compare optimization methods across different functions, algorithm choice is an essential component. The Nelder-Mead Simplex Method (Nelder & Mead, 1965) has gained much importance since its introduction and utilizes a polygon that converges on the global minima through reflection, contraction, and expansion motions. Because this method only uses function evaluations and does not require the calculation of derivatives, it has broad applicability over a range or problems. Conversely, the quasi-Newton method creates a quadratic model problem using curvature information at each iteration. Through numerical differentiation, gradient information is collected, and the method follows a direction of descent to reach the minima (Edgar & Himmelblau, 1987). While both methods are appropriate for finding the minima of nonlinear functions, they take very different approaches to arrive at the solution.

Choice of Starting Points

It is well known that the choice of initial estimates is critical to ensure convergence during minimization of nonlinear functions (Beck, 2014). Convergence is faster and more likely when the initial estimates are closer to the values that correspond to the global minimum. In this study, parameter values corresponding to function minima were known and were (1,1), (1,3), and (0,0), for the Rosenbrock, Booth, and Matyas functions, respectively. For robust testing of the performance of the optimization algorithms, 6 sets of initial estimates were used for each function, and these included both positive and negative values as shown in Table 1.

Table 1. Choice of initial parameter estimates and their values at the minimum for the 3 functions evaluated in this study.

Function	Values at f(x) = 0	Initial Estimates Evaluated
Rosenbrock	(1, 1)	(-10, -10), (-5, -5), (-2, -2), (2, 2), (5, 5), (10, 10)
Booth	(1, 3)	(-10, -30), (-5, -15), (-2, -6), (2, 6), (5, 15), (10, 30)
Matyas	(0, 0)	(-10, -10), (-5, -5), (-2, -2), (2, 2), (5, 5), (10, 10)

Computer Programs

Optimization of all 3 functions was performed using MATLAB R2023a (The MathWorks, Inc., Natick, MA). The Simplex method was implemented using fminsearch while the quasi-Newton method was implemented using the fminunc function. The termination and convergence criteria defaults were 1 x 10^{-4} for both the algorithms evaluated in this study.

Results

Rosenbrock Function

A total of 6 initial parameter estimates were used to find the minima of the Rosenbrock function using the Simplex and quasi-Newton methods. Convergence was reached in all 6 instances and the associated number of iterations are shown in Table 2. The convergence path is shown in Figure 4 over the contour plots for the two extreme starting conditions of (10, 10) and (-10, -10). The iteration trajectories for the other initial estimates in Table 2 were generally like those shown in Figure 4.

Across both methods, there was a general increase in the number of iterations as the initial estimates moved further away from the optimal values of (1, 1). Additionally, within each method, positive initial estimates were associated with a higher number of iterations than the corresponding negative initial estimates. The most striking difference from Table 2 is the efficiency of the quasi-Newton method where a 3 – 4-fold reduction in the number of iterations is seen across the 6 initial estimates evaluated in this study.

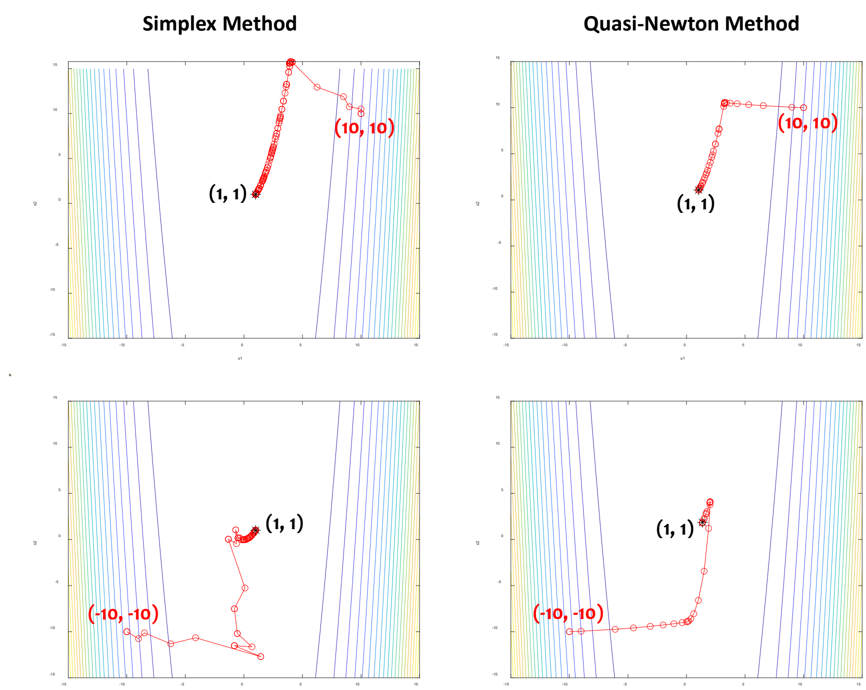

Figure 4. Contour plots showing the iteration trajectories for the Rosenbrock function for the Simplex and quasi-Newton methods at initial estimates of (10, 10) and (-10, -10). Convergence to the optimal value of (1, 1) was achieved in all instances.

Table 2. Comparison of convergence between the Simplex and quasi-Newton method for the Rosenbrock function.

Initial Estimates	Number of Iterations	
(x_1, x_2)	Simplex Method	Quasi-Newton Method
(2, 2)	62	23
(5, 5)	106	36
(10, 10)	129	44
(-2, -2)	79	19
(-5, -5)	78	27
(-10, -10)	89	28

Booth Function

The Booth function has a global minimum of (1, 3) and like the approach for the Rosenbrock function, initial estimates were chosen that were 2-fold, 5-fold and 10-fold of the global minima in both positive and negative directions. These values are shown in Table 3 along with the corresponding numbers of iterations for the Simplex and quasi-Newton methods. Figure 5 shows the iteration trajectories starting from (10, 30) and (-10, -30), the two extreme initial estimates on contour plots. Convergence to the optimal values was achieved in all instances with the quasi-Newton method rapidly reaching the global minimum compared to the Simplex method.

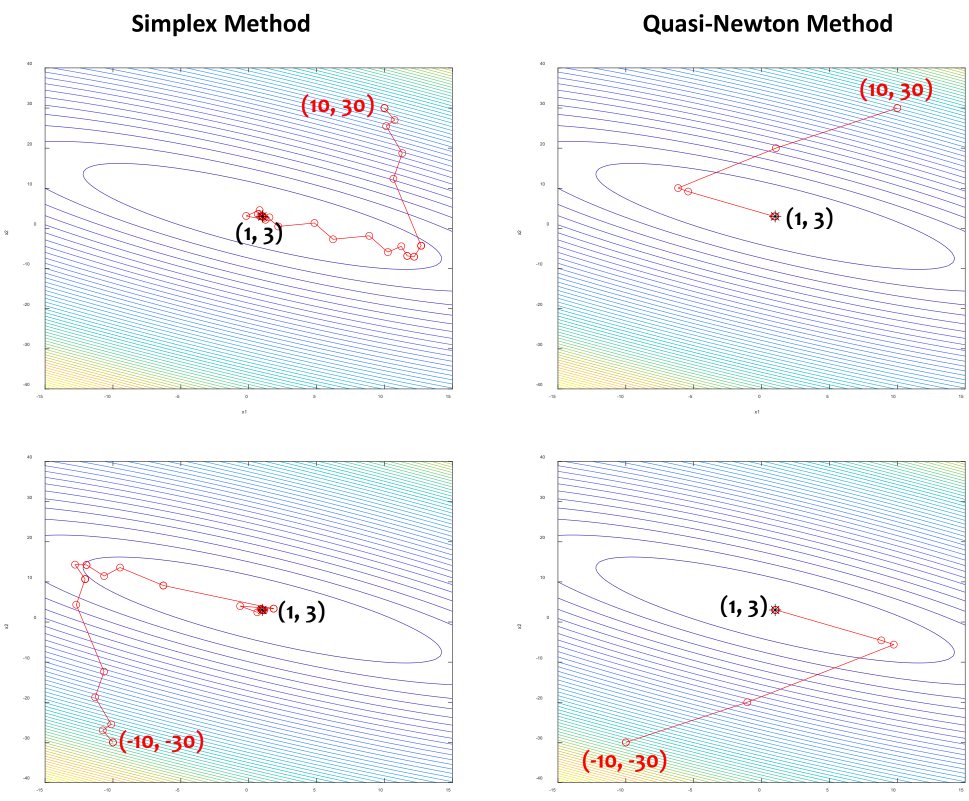

Figure 5. Contour plots showing the iteration trajectories for the Booth function for the Simplex and quasi-Newton methods at initial estimates of (10, 30) and (-10, -30). Convergence to the optimal value of (1, 3) was achieved in all instances.

Table 3. Comparison of convergence between the Simplex and quasi-Newton method for the Booth function.

Initial Estimates	**Number of Iterations**	
(x_1, x_2)	**Simplex Method**	**Quasi-Newton Method**
(2, 6)	46	6
(5, 15)	54	6
(10, 30)	55	6
(-2, -6)	58	6
(-5, -15)	56	8
(-10, -30)	55	6

Compared to Table 2 data for the Rosenbrock function, the number of iterations did not vary significantly for the Booth function as the initial estimates were changed (Table 3) across both optimization approaches. However, the difference in performance was very significant with the Simplex method converging in 46 – 58 iterations compared to 6 – 8 iterations for the quasi-Newton method. This translates to almost an order of magnitude reduction in the number of iterations for the quasi-Newton method.

Matyas Function

The Matyas function has a global minimum of (0, 0) and the initial estimates of x_1 and x_2 used to find the function minimum and the associated number of iterations for the Simplex and quasi-Newton methods are shown in Table 4. Like the Booth function, there was minimal influence of the initial estimates on the number of iterations with the Simplex method converging in 36 – 44 iterations while the quasi-Newton method only needed 2 iterations for convergence across all conditions.

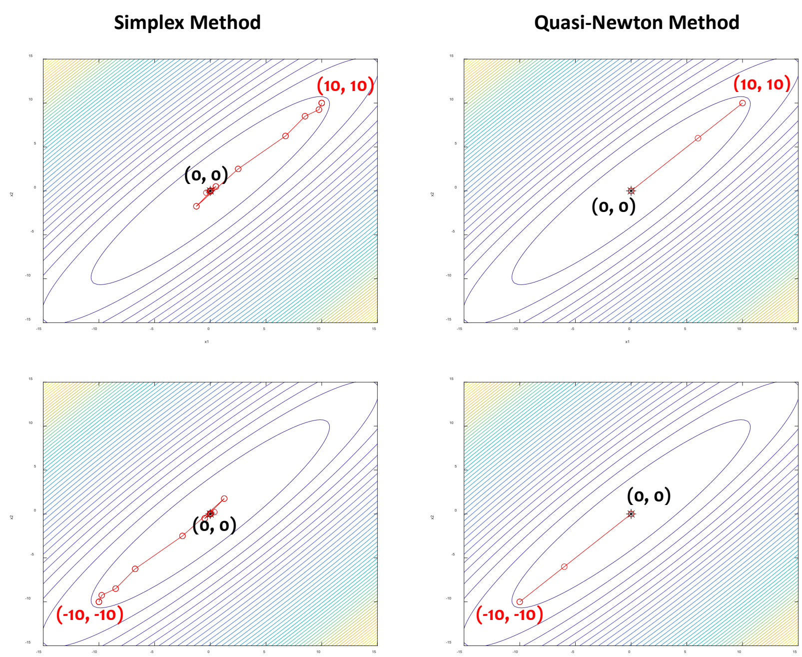

Figure 6. Contour plots showing the iteration trajectories for the Matyas function for the Simplex and Quasi-Newton methods at initial estimates of (10, 10) and (-10, -10). Convergence to the optimal value of (0, 0) was achieved in all instances.

Figure 6 shows the convergence path for both optimization approaches at the extreme initial estimates of (10, 10) and (-10, -10). The flat profile of the Matyas function resulted in a linear convergence path for both optimization approaches and the quasi-Newton method was ~20-fold more efficient than the Simplex method (2 vs ~40 iterations).

Table 4. Comparison of convergence between the Simplex and quasi-Newton method for the Matyas function.

Initial Estimates	**Number of Iterations**	
(x_1, x_2)	**Simplex Method**	**Quasi-Newton Method**
(2, 2)	39	2
(5, 5)	42	2
(10, 10)	44	2
(-2, -2)	36	2
(-5, -5)	42	2
(-10, -10)	44	2

Conclusion

When the Simplex and quasi-Newton optimization methods were used to find the global minima of the Rosenbrock, Booth, and Matyas functions from varying starting points, significant differences in the number of iterations necessary for convergence were seen, especially for the Booth and Matyas functions. While all initial estimates resulted in convergence across both optimization methods, the quasi-Newton method needed ~10-fold and ~20-fold fewer iterations for the Booth and Matyas functions, respectively. It is very likely the long flat valley of the Booth function and the plate-like profile of the Matyas function were well suited for rapid convergence of the quasi-Newton method. The difference in performance was not as high for the Rosenbrock function where the global minimum lies inside a long and narrow parabolic valley. Our work reinforces past observations that while many optimization methods may lead to the desired solution, their performance is dependent on the function being evaluated. Matching properties of the function to be minimized with the features of the optimization method can help choose the most appropriate method to determine the global minimum of a function.

References

Bartholomew-Biggs, M. (2005). *Nonlinear Optimization with Financial Applications*. Springer. https://doi.org/10.1007/b102601

Bartholomew-Biggs, M. (2008). *Nonlinear Optimization with Engineering Applications*. Springer.

Beck, A. (2014). *Introduction to Nonlinear Optimization: Theory, Algorithms, and Applications with MATLAB*. SIAM-Society for Industrial and Applied Mathematics.

Cui, Y., Geng, Z., Zhu, Q., & Han, Y. (2017). Review: Multi-objective optimization methods and application in energy saving. *Energy*, *125*, 681–704. https://doi.org/10.1016/j.energy.2017.02.174

Edgar, T. F., & Himmelblau, D. M. (1987). *Optimization of Chemical Processes*. McGraw-Hill College.

Nelder, J. A., & Mead, R. (1965). A Simplex Method for Function Minimization. *The Computer Journal*, *7*(4), 308–313. https://doi.org/10.1093/comjnl/7.4.308

Rosenbrock, H. H. (1960). An Automatic Method for Finding the Greatest or Least Value of a Function. *The Computer Journal*, *3*(3), 175–184. https://doi.org/10.1093/comjnl/3.3.175

Salleh, Z., Almarashi, A., & Alhawarat, A. (2022). Two efficient modifications of AZPRP conjugate gradient method with sufficient descent property. *Journal of Inequalities and Applications*, *2022*. https://doi.org/10.1186/s13660-021-02746-0

Wang, Z., & Luo, Q. (2021). Hybrid metaheuristic algorithm using butterfly and flower pollination base on mutualism mechanism for global optimization problems. *Engineering with Computers*, *37*. https://doi.org/10.1007/s00366-020-01025-8