Nelder-Mead Toolbox Manual – Bibliography –

Version 0.2
September 2009

Michaël BAUDIN

Contents

Chapter 1

Nelder-Mead bibliography

In this section, we present a brief overview of selected papers, sorted in chronological order, which deal with the Nelder-Mead algorithm

1.1 Spendley, Hext, Himsworth, 1962

"Sequential Application of Simplex Designs in Optimisation and Evolutionary Operation", Spendley W., Hext G. R. and Himsworth F. R., American Statistical Association and American Society for Quality, 1962

This article [21] presents an algorithm for unconstrained optimization in which a simplex is used. The simplex has a fixed, regular (i.e. all lengths are equal), shape and is made of n+1 vertices (where n is the number of parameters to optimize). The algorithm is based on the reflection of the simplex with respect to the centroid of better vertices. One can add a shrink step so that the simplex size can converge to zero. Because the simplex shape cannot change, the convergence rate may be very slow if the eigenvalues of the hessian matrix have very different magnitude.

1.2 Nelder, Mead, 1965

"A Simplex Method for Function Minimization", Nelder J. A. and Mead R., The Computer Journal, 1965

This article [12] presents the Nelder-Mead unconstrained optimization algorithm. It is based on a simplex made of n+1 vertices and is a modification of the Spendley's et al algorithm. It includes features which enables the simplex to adapt to the local landscape of the cost function. The additional steps are expansion, inside contraction and outside contraction. The stopping criterion is based on the standard deviation of the function value on the simplex.

The convergence of the algorithm is better than Spendley's et al. The method is compared against Powell's free-derivative method (1964) with comparable behavior. The algorithm is "greedy" in the sense that the expansion point is kept if it improves the best function value in the current simplex. Most Nelder-Mead variants which have been analyzed after are keeping the expansion point only if it improves over the reflection point.

1.3 Box, 1965

"A New Method of Constrained Optimization and a Comparison With Other Methods", M. J. Box, The Computer Journal 1965 8(1):42-52, 1965, British Computer Society

In this paper [2], Box presents a modification of the NM algorithm which takes into accounts for bound constraints and non-linear constraints. This variant is called the Complex method. The method expects that the initial guess satisfies the nonlinear constraints. The nonlinear constraints are supposed to define a convex set. The algorithm ensures that the simplex evolves in the feasible space.

The method to take into account for the bound constraints is based on projection of the parameters inside the bounded domain. If some nonlinear constraint is not satisfied, the trial point is moved halfway toward the centroid of the remaining points (which are all satisfying the nonlinear constraints).

The simplex may collapse into a subspace if a projection occurs. To circumvent this problem, k>=n+1 vertices are used instead of the original n+1 vertices. A typical value of k is k=2n. The initial simplex is computed with a random number generator, which takes into account for the bounds on the parameters. To take into account for the nonlinear constraints, each vertex of the initial simplex is moved halfway toward the centroid of the points satisfying the constraints (in which the initial guess already is).

1.4 Guin, 1968

"Discussion and correspondence: modification of the complex method of constrained optimization", J. A. Guin, The Computer Journal, 1968

In this article [3], Guin suggest 3 rules to improve the practical convergence properties of Box's complex method. These suggestions include the use of the next-to-worst point when the worst point does not produce an improvement of the function value. The second suggestion is to project the points strictly into the bounds, instead of projecting inside the bounds. The third suggestion is related to the failure of the method when the centroid is no feasible. In that case, Guin suggest to restrict the optimization in the subspace defined by the best vertex and the centroid.

1.5 O'Neill, 1971

"Algorithm AS47 - Function minimization using a simplex procedure", R. O'Neill, 1971, Applied Statistics

In this paper [13], R. O'Neill presents a fortran 77 implementation of the Nelder-Mead algorithm. The initial simplex is computed axis-by-axis, given the initial guess and a vector of step lengths. A factorial test is used to check if the computed optimum point is a local minimum.

1.6 Parkinson and Hutchinson, 1972

In [14], "An investigation into the efficiency of variants on the simplex method", Parkinson and Hutchinson explored several ways of improvement. First, they investigate the sensitivity of the algorithm to the initial simplex. Two parameters were investigated, i.e. the initial length and the orientation of the simplex. An automatic setting for the orientation, though very desirable, is not easy to design. Parkinson and Hutchinson tried to automatically compute the scale of the initial simplex by two methods, based on a "line search" and on a local "steepest descent". Their second investigation adds a new step to the algorithm, the unlimited expansion. After a sucessful expansion, the algorithm tries to produce an expansion point by taking the largest possible number of expansion steps. After an unlimited expansion steps is performed, the simplex is translated, so that excessive modification of the scale and shape is avoided. Combined and tested against low dimension problems, the modified algorithm, named PHS, provides typical gains of 20function evaluations.

1.7 Richardson and Kuester, 1973

"Algorithm 454: the complex method for constrained optimization", Richardson Joel A. and Kuester J. L., Commun. ACM, 1973

In this paper [18], Richardson and Kuester shows a fortran 77 implementation of Box's complex optimization method. The paper clarifies several specific points from Box's original paper while remaining very close to it. Three test problems are presented with the specific algoritmic settings (such as the number of vertices for example) and number of iterations.

1.8 Shere, 1973

"Remark on algorithm 454 : The complex method for constrained optimization", Shere Kenneth D., Commun. ACM, 1974

In this article [20], Shere presents two counterexamples where the algorithm 454, implemented by Richardson and Kuester produces an infinite loop. "This happens whenever the corrected point, the centroid of the remaining complex points, and every point on the line segment joining these two points all have functional values lower than the functional values at each of the remaining complex points.

1.9 Routh, Swartz, Denton, 1977

"Performance of the Super-Modified Simplex", M.W. Routh, P.A. Swartz, M.B. Denton, Analytical Chemistry, 1977

In this article [19], Routh, Swartz and Denton present a variant of the Nelder-Mead algorithm, which is called the Modified Simplex Method (SMS) in their paper. The algorithm is modified in the following way. After determination of the worst response (W), the responses at the centroid (C) and reflected (R) vertices are measured and a second-order polynomial curve is fitted to the responses at W, C and R. Furthermore, the curve is extrapolated beyond W and R by a percentage of the W-R vector resulting in two types of curve shapes. In the concave down case, a maximum occurs within the interval. Assuming a maximization process, evaluation of the derivative of the curve reveals the location of the predicted optimum whose response is subsequently evaluated, the new vertex is located at that position, and the optimization process is continued. In the concave up case, a response maximum does not occur within the interval so the extended interval boundary producing the highest predicted response is chosen as the new vertex location, its response is determined, and the optimization is continued. If the response at the predicted extended interval boundary location does not prove to be greater than the response at R, the vertex R may instead be retained as the new vertex and the process continued. The slope at the extended interval boundary may additionally be evaluated dictating the magnitude of the expansion coefficient, i.e. the greater the slope (indicating rapid approach to the optimum location), the smaller the required expansion coefficient and, conversely, the smaller the slope (indicating remoteness from the optimum location), the larger the required expansion coefficient.

Some additional safeguard procedure must be used in order to prevent the collapse of the simplex.

1.10 Van Der Wiel, 1980

"Improvement of the Super-Modified Simplex Optimization Procedure", P.F.A., Van Der Wiel Analytica Chimica Acta, 1980

In this article [23], Van Der Wiel tries to improve the SMS method by Routh et al.. His modifications are based on a Gaussian fit, weighted reflection point and estimation of response at the reflection point. Van Der Wiel presents a simplified pseudo-code for one algorithm The method is tested in 5 cases, where the cost function is depending on the exponential function.

1.11 Walters, Parker, Morgan and Deming, 1991

"Sequential Simplex Optimization for Quality and Productivity in Research, Development, and Manufacturing", F. S. Walters, L. R. Parker, Jr., S. L. Morgan, and S. N. Deming, 1991

In this book [24], Walters, Parker, Morgan and Deming give a broad view on the simplex methods in chemistry. The Spendley et al. and Nelder-Mead algorithms are particularily deeply analyzed, with many experiments analyzed in great detail. Template tables are given, so that an engineer can manually perform the optimization and make the necessary calculations. Practical advices are given, which allow to make a better use of the algorithms.

In chapter 5, ”Comments on Fixed-size and Variable-size Simplexes”, comparing the path of the two algorithms allows to check that a real optimum has been found. When the authors analyze the graph produced by the response depending on the number of iteration, the general behavior of the fixed-size algorithm is made of four steps. Gains in response are initially rapid, but the rate of return decreases as the simplex probes to find the ridge and then moves along the shallower ridge to find the optimum. The behavior from different starting locations is also analyzed. Varying the size of the initial simplex is also analyzed for the fixed-size simplex algorithm. The many iterations which are produced when a tiny initial simplex is used with the fixed-size simplex is emphasized.

The chapter 6, ”General Considerations”, warns that the user may setup an degenerate initial simplex, leading to a false convergence of the algorithm. Various other initial simplices are analyzed. Modifications in the algorithm to take into account for bounds contraints are presented. The behavior of the fixed-size and variable-size simplex algorithms is analyzed when the simplex converges. The ”k+1” rule, introduced by Spendley et al. to take into account for noise in the cost function is presented.

The chapter 7, ”Additional Concerns and Topics” deals with advanced questions regarding these algorithms. The variable size simplex algorithm is analyzed in the situation of a ridge. Partially oscillatory collapse of the Nelder-Mead algorithm is presented. The same behavior is presented in the case of a saddle point. This clearly shows that practionners were aware of the convergence problem of this algorithm well before Mc Kinnon presented a simple counter example (in 1998). The ”Massive Contraction” step of Nelder and Mead is presented as a solution for this oscillatory behavior. The authors present a method, due to Ernst, which allows to keep the volume of the simplex, instead of shrinking it. This method is based on a translation of the simplex. This modification requires $n + 1$ function evaluations. A more efficient method, due to King, is based on reflection with respect to the next-to-worst vertex. This modification was first suggested by Spendley et al. in their fixed-size simplex algorithm.

In the same chapter, the authors present the behavior of the algorithms in the case of multiple optima. They also present briefly other types of simplex algorithms.

A complete bibliography (from 1962 to 1990) on simplex-based optimization is given in the end of the book.

1.12 Subrahmanyam, 1989

”An extension of the simplex method to constrained nonlinear optimization”, M. B. Subrahmanyam, Journal of Optimization Theory and Applications, 1989

In this article [22], the simplex algorithm of Nelder and Mead is extended to handle nonlinear optimization problems with constraints. To prevent the simplex from collapsing into a subspace near the constraints, a delayed reflection is introduced for those points moving into the infeasible region. Numerical experience indicates that the proposed algorithm yields good results in the presence of both inequality and equality constraints, even when the constraint region is narrow.

If a vertex becomes infeasible, we do not increase the value at this vertex until the next iteration is completed. Thus, the next iteration is accomplished using the actual value of the function at the infeasible point. At the end of the iteration, in case the previous vertex is not the worst vertex, it is assigned a high value, so that it then becomes a candidate for reflection during the next iteration.

The paper presents numerical experiments which are associated with thousands of calls to the cost function. This may be related with the chosen reflection factor equal to 0.95, which probably cause a large number of reflections until the simplex can finally satisfy the constraints.

1.13 Numerical Recipes in C, 1992

”Numerical Recipes in C, Second Edition”, W. H. Press, Saul A. Teukolsky, William T. Vetterling and Brian P. Flannery, 1992

In this book [17], an ANSI C implementation of the Nelder-Mead algorithm is given. The initial simplex is based on the axis. The termination criterion is based on the relative difference of the function value of the best and worst vertices in the simplex.

1.14 Lagarias, Reeds, Wright, Wright, 1998

”Convergence Properties of the Nelder–Mead Simplex Method in Low Dimensions”, Jeffrey C. Lagarias, James A. Reeds, Margaret H. Wright and Paul E. Wright, SIAM Journal on Optimization, 1998

This paper [9] presents convergence properties of the Nelder-Mead algorithm applied to stricly convex functions in dimensions 1 and 2. Proofs are given to a minimizer in dimension 1, and various limited convergence results for dimension 2.

1.15 Mc Kinnon, 1998

”Convergence of the Nelder–Mead Simplex Method to a Nonstationary Point”, SIAM J. on Optimization, K. I. M. McKinnon, 1998

In this article [10], Mc Kinnon analyzes the behavior of the Nelder-Mead simplex method for a family of examples which cause the method to converge to a nonstationnary point. All the examples use continuous functions of two variables. The family of functions contains strictly convex functions with up to three continuous derivatives. In all the examples, the method repeatedly applies the inside contraction step with the best vertex remaining fixed. The simplices tend to a straight line which is orthogonal to the steepest descent direction. It is shown that this behavior cannot occur for functions with more than three continuous derivatives.

1.16 Kelley, 1999

"Detection and Remediation of Stagnation in the Nelder–Mead Algorithm Using a Sufficient Decrease Condition", SIAM J. on Optimization, Kelley, C. T., 1999

In this article [7], Kelley presents a test for sufficient decrease which, if passed for the entire iteration, will guarantee convergence of the Nelder-Mead iteration to a stationary point if the objective function is smooth. Failure of this condition is an indicator of potential stagnation. As a remedy, Kelley propose to restart the algorithm with an oriented simplex, smaller than the previously optimum simplex, but with a better shape and which approximates the steepest descent step from the current best point. The method is experimented against Mc Kinnon test function and allow to converge to the optimum, where the original Nelder -Mead algorithm was converging to a non-stationary point. Although the oriented simplex works well in practice, other strategies may be chosen with similar results, such as a simplex based on axis, a regular simplex (like Spendley's) or a simplex based on the variable magnitude (like Pfeffer's suggestion in Matlab's fminsearch). The paper also shows one convergence theorem which prove that if the sufficient decrease condition is satisfied and if the product of the condition of the simplex by the simplex size converge to zero, therefore, with additional assumptions on the cost function and the sequence of simplices, any accumulation point of the simplices is a critical point of f.

The same ideas are presented in the book [8].

1.17 Han, 2000

In his Phd thesis [4], Lixing Han analyzes the properties of the Nelder-Mead algorithm. Han present two examples in which the Nelder-Mead simplex method does not converge to a single point. The first example is a nonconvex function with bounded level sets and it exhibits similar nonconvergence properties with the Mc Kinnon counterexample $f(\xi_1, \xi_2) = \xi_1^2 - \xi_2(\xi_2 - 2)$. The second example is a convex function with bounded level sets, for which the Nelder-Mead simplices converge to a degenerate simplex, but not to a single point. These nonconvergent examples support the observations by some practitionners that in the Nelder-Mead simplices may collapse into a degenerate simplex and therefore support the use of a restart strategy. Han also investigates the effect of the dimensionality of the Nelder-Mead method. It is shown that the Nelder-Mead simplex method becomes less efficient as the dimension increases. Specifically, Han consider the quadratic function $\xi_1^2 + \ldots + \xi_1^n$ and shows that the Nelder-Mead method becomes less efficient as the dimension increases. The considered example offers insight into understanding the effect of dimensionnality on the Nelder-Mead method. Given all the known failures and inefficiencies of the Nelder-Mead method, a very interesting question is why it is so popular in practice. Han present numerical results of the Nelder-Mead method on the standard collection of Moré-Garbow-Hillstrom with dimensions $n \leq 6$. Han compare the Nelder-Mead method with a finite difference BFGS method and a finite difference steepest descent method. The numerical results show that the Nelder-Mead method is much more efficient than the finite difference steepest descent method for the problems he tested with dimensions $n \leq 6$. It is also often comparable with the finite difference BFGS method, which is believed to be the best derivative-free method. Some of these results are reproduced in [5] by Han and Neumann, "Effect of dimensionality on the Nelder-ĂŞMead simplex method" and in [6], "On the roots of certain polynomials arising from the analysis of the Nelder-Mead simplex method".

1.18 Nazareth, Tseng, 2001

"Gilding the Lily: A Variant of the Nelder-Mead Algorithm Based on Golden-Section Search" Computational Optimization and Applications, 2001, Larry Nazareth and Paul Tseng

The article [11] propose a variant of the Nelder-Mead algorithm derived from a reinterpretation of univariate golden-section direct search. In the univariate case, convergence of the variant can be analyzed analogously to golden-section search.

The idea is based on a particular choice of the reflection, expansion, inside and outside contraction parameters, based on the golden ratio. This variant of the Nelder-Mead algorithm is called Nelder-Mead-Golden- Ratio, or NM-GS. In one dimension, the authors exploit the connection with golden-search method and allows to prove a convergence theorem on unimodal univariate functions. This is marked contrast to the approach taken by Lagarias et al. where considerable effort is expended to show convergence of the original NM algorithm on strictly convex univariate functions. With the NM-GS variant, one obtain convergence in the univariate case (using a relatively simple proof) on the broader class of unimodal functions.

In the multivariate case, the authors modify the variant by replacing strict descent with fortified descent and maintaining the interior angles of the simplex bounded away from zero. Convergence of the modified v ariant can be analyzed by applying results for a fortified- descent simplicial search method. Some numerical experience with the variant is reported.

1.19 Perry, Perry, 2001

"A New Method For Numerical Constrained Optimization" by Ronald N. Perry, Ronald N. Perry, March 2001

In this report [15], we propose a new method for constraint handling that can be applied to established optimization algorithms and which significantly improves their ability to traverse through constrained space. To make the presentation concrete, we apply the new constraint method to the Nelder and Mead polytope algorithm. The resulting technique, called SPIDER, has shown great initial promise for solving difficult (e.g., nonlinear, nondifferentiable, noisy) constrained problems.

In the new method, constraints are partitioned into multiple levels. A constrained performance, independent of the objective function, is defined for each level. A set of rules, based on these partitioned performances, specify the ordering and movement of vertices as they straddle constraint boundaries; these rules [...] have been shown to significantly aid motion along constraints toward an optimum. Note that the new approach uses not penalty function and thus does not warp the performance surface, thereby avoiding the possible ill-conditioning of the objective function typical in penalty methods.

No numerical experiment is presented.

1.20 Andersson, 2001

"Multiobjective Optimization in Engineering Design - Application to fluid Power Systems" Johan Andersson, 2001

This PhD thesis [1] gives a brief overview of the Complex method by Box in section 5.1.

1.21 Peters, Bolte, Marschner, Nüssen and Laur, 2002

In [16], "Enhanced Optimization Algorithms for the Development of Microsystems", the authors combine radial basis function interpolation methods with the complex algorithm by Box. Interpolation with radial basis functions is a linear approach in which the model function f is generated via the weighted sum of the basis functions $\Phi_i(r)$. The parameter r describes the distance of the current point from the center x_i of the ith basis function. It is calculated via the euclidean norm. It is named ComplInt strategy. The name stands for Complex in combination with interpolation. The Complex strategy due to Box is very well suited for the combination with radial basis function interpolation for it belongs to the polyhedron strategies. The authors presents a test performed on a pratical application, which leaded them to the following comment : "The best result achieved with the ComplInt strategy is not only around 10% better than the best result of the Complex strategy due to Box, the ComplInt also converges much faster than the Complex does: while the Complex strategy needs an average of 7506, the ComplInt only calls for an average of 2728 quality function evaluations."

1.22 Han, Neumann, 2006

"Effect of dimensionality on the Nelder-Mead simplex method", L. Han and M. Neumann (2006),

In this article [5], the effect of dimensionality on the Nelder-Mead algorithm is investigated. It is shown that by using the quadratic function $f(x) = x^T * x$, the Nelder-Mead simplex method deteriorates as the dimension increases. More precisely, in dimension 1, with the quadratic function $f(x) = x^2$ and a particular choice of the initial simplex, applies inside contraction step repeatedly and the convergence rate (as the ratio between the length of the simplex at two consecutive steps) is 1/2. In dimension 2, with a particular initial simplex, the NM algorithm applies outside contraction step repeatedly and the convergence rate is $\sqrt(2)/2$.

For n>=3, a numerical experiment is performed on the quadratic function with the fminsearch algorithm from Matlab. It is shown that the original NM algorithm has a convergence rate which is converging towards 1 when n increases. For n=32, the rate of convergence is 0.9912.

1.23 Singer, Nelder, 2008

`http://www.scholarpedia.org/article/Nelder-Mead_algorithm` Singer and Nelder

This article is a complete review of the Nelder-Mead algorithm. Restarting the algorithm is adviced when a premature termination occurs.

Bibliography

[1] Johan Andersson and Linköpings Universitet. Multiobjective optimization in engineering design: Application to fluid power systems. Technical report, Department of Mechanical Engineering, Linköping University, 2001. https://polopoly.liu.se/content/1/c6/10/99/74/phdthesis.pdf, http://www.iei.liu.se/machine/johan-olvander.

[2] M. J. Box. A new method of constrained optimization and a comparison with other methods. *The Computer Journal*, 8(1):42–52, 1965.

[3] J. A. Guin. Discussion and correspondence: modification of the complex method of constrained optimization. *The Computer Journal*, 10(4):416–417, February 1968.

[4] Lixing Han. *Algorithms in Unconstrained Optimization*. Ph.D., The University of Connecticut, 2000.

[5] Lixing Han and Michael Neumann. Effect of dimensionality on the nelder-mead simplex method. *Optimization Methods and Software*, 21(1):1–16, 2006.

[6] Lixing Han, Michael Neumann, and Jianhong Xu. On the roots of certain polynomials arising from the analysis of the nelder-mead simplex method. *Linear Algebra and its Applications*, 363:109–124, 2003.

[7] C. T. Kelley. Detection and remediation of stagnation in the nelder–mead algorithm using a sufficient decrease condition. *SIAM J. on Optimization*, 10(1):43–55, 1999.

[8] C. T. Kelley. *Iterative Methods for Optimization*, volume 19. SIAM Frontiers in Applied Mathematics, 1999. http://www.siam.org/books/textbooks/fr18_book.pdf.

[9] Jeffrey C. Lagarias, James A. Reeds, Margaret H. Wright, and Paul E. Wright. Convergence properties of the nelder–mead simplex method in low dimensions. *SIAM Journal on Optimization*, 9(1):112–147, 1998. http://link.aip.org/link/?SJE/9/112/1.

[10] K. I. M. McKinnon. Convergence of the nelder–mead simplex method to a nonstationary point. *SIAM J. on Optimization*, 9(1):148–158, 1998.

[11] Larry Nazareth and Paul Tseng. Gilding the lily: A variant of the nelder-mead algorithm based on golden-section search. *Comput. Optim. Appl.*, 22(1):133–144, 2002.

[12] J. A. Nelder and R. Mead. A simplex method for function minimization. *The Computer Journal*, 7(4):308–313, January 1965.

[13] R. O'Neill. Algorithm AS47 - Function minimization using a simplex procedure. *Applied Statistics*, 20(3):338–346, 1971.

[14] Parkinson and Hutchinson. An investigation into the efficiency of variants on the simplex method. *F. A. Lootsma, editor, Numerical Methods for Non-linear Optimization*, pages 115–135, 1972.

[15] Ronald N. Perry and Ronald N. Perry. A new method for numerical constrained optimization, 2001.

[16] D. Peters, H. Bolte, C. Marschner, O. Nüssen, and R. Laur. Enhanced optimization algorithms for the development of microsystems. *Analog Integr. Circuits Signal Process.*, 32(1):47–54, 2002.

[17] W. H. Press, Saul A. Teukolsky, William T. Vetterling, and Brian P. Flannery. *Numerical Recipes in C, Second Edition.* 1992.

[18] Joel A. Richardson and J. L. Kuester. Algorithm 454: the complex method for constrained optimization. *Commun. ACM*, 16(8):487–489, 1973.

[19] M. W. Routh, P.A. Swartz, and M.B. Denton. Performance of the super modified simplex. *Analytical Chemistry*, 49(9):1422–1428, 1977.

[20] Kenneth D. Shere. Remark on algorithm 454 : The complex method for constrained optimization. *Commun. ACM*, 17(8):471, 1974.

[21] W. Spendley, G. R. Hext, and F. R. Himsworth. Sequential application of simplex designs in optimisation and evolutionary operation. *Technometrics*, 4(4):441–461, 1962.

[22] M. B. Subrahmanyam. An extension of the simplex method to constrained nonlinear optimization. *J. Optim. Theory Appl.*, 62(2):311–319, 1989.

[23] P.F.A. Van Der Wiel. Improvement of the super modified simplex optimisation procedure. *Analytica Chemica Acta*, 122:421?–433, 1980.

[24] F. S. Walters, L. R. Parker, S. L. Morgan, and S. N. Deming. *Sequential Simplex Optimization for Quality and Productivity in Research, Development, and Manufacturing.* CRC Press, Boca Raton, FL, 1991.