Nelder-Mead Simplex Algorithm: Effect of Dimensionality and New Implementation

Colloquium Talk

Department of Mathematics

Southern Illinois University-Carbondale

March 7, 2013

Lixing Han

University of Michigan-Flint

Derivative-free optimization

We consider solving the unconstrained optimization problem

$$\text{minimize } f(x)$$

numerically, where the objective function $f : R^n \to R$ is continuous.

- The derivatives of the objective function are unavailable or expensive to compute.
- Derivative-free algorithms do not use derivatives or attempt to approximate derivatives.

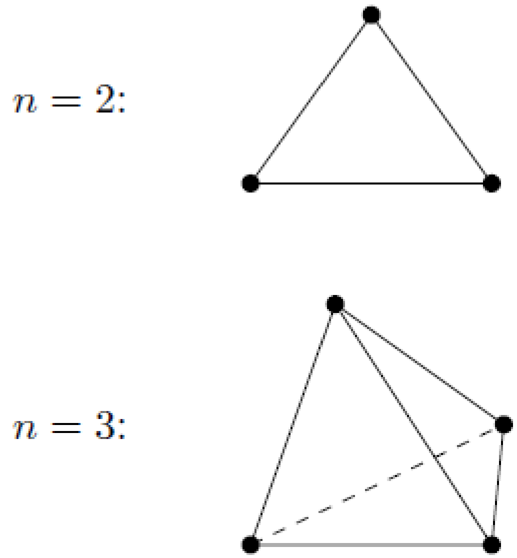
Nelder-Mead simplex algorithm

The simplex method of John Nelder and Roger Mead (*Computer Journal*, 1965) is one of the most widely used derivative-free methods.

[This simplex method is totally different from Dantzig's simplex method for linear programming.]

John Nelder (1924-2010)

Simplex: A geometric figure in R^n that is the convex hull of n+1 vertices.



Idea

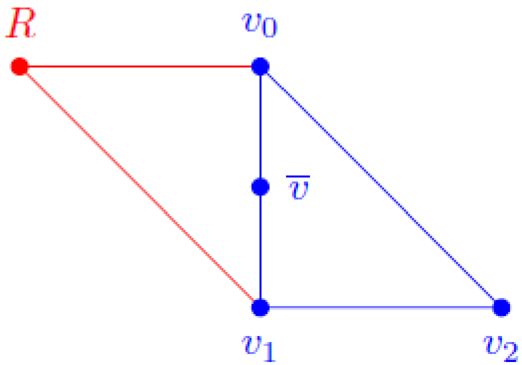
The Nelder-Mead method generates a sequence of simplices which converges (hopefully) to a minimizer. In each simplex, the vertices are sorted:

$$f(\mathbf{v}_0) \leq f(\mathbf{v}_1) \leq \cdots \leq f(\mathbf{v}_n).$$

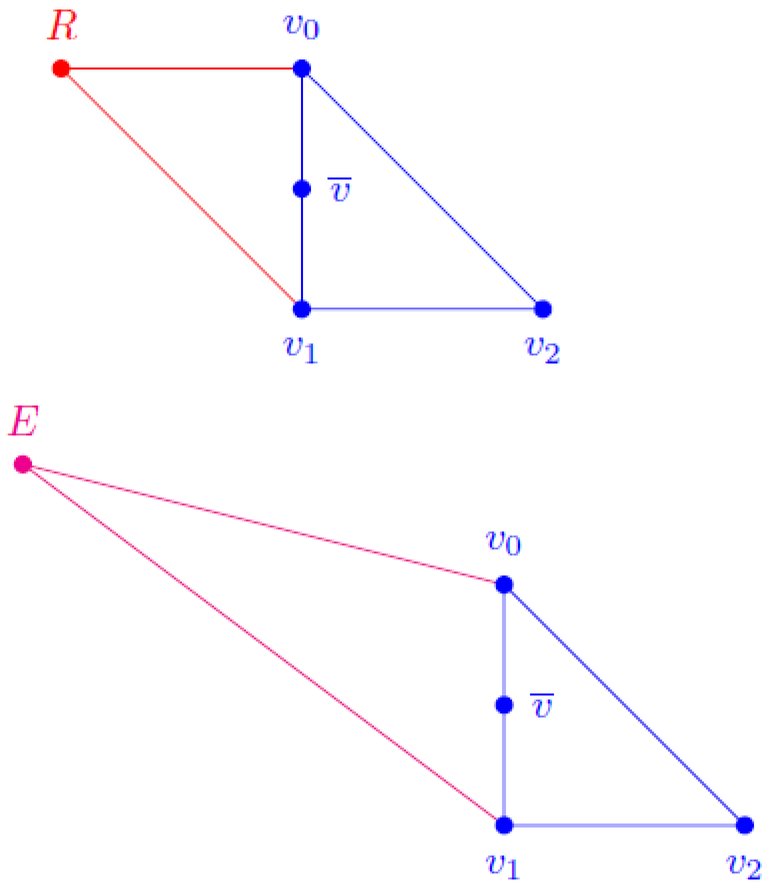
At iteration k, we generate the next simplex using the current one and one of the four operations:

- *Reflection*
- *Expansion*
- *Contraction (inside or outside)*
- *Shrinkage*

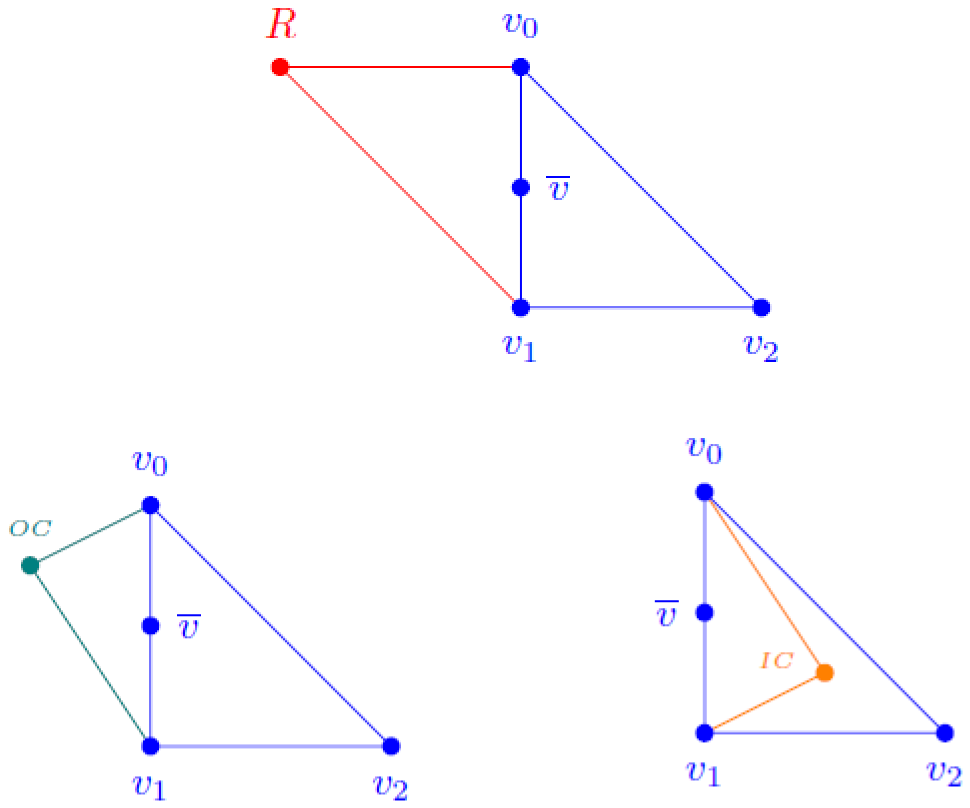
Reflection



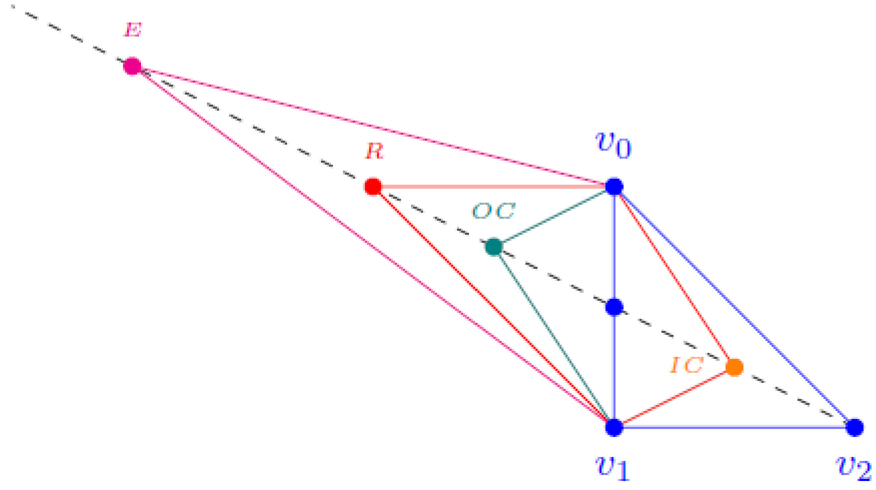
Expansion



Contraction (outside or inside)

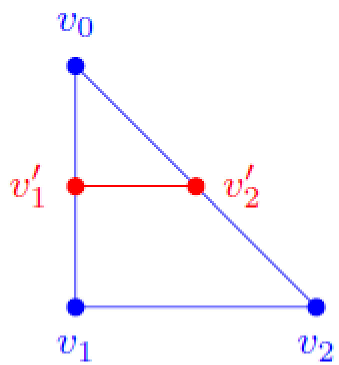


- At each iteration, the Nelder-Mead method replaces the worst vertex with a better vertex (which is one of R, E, OC, or IC.)



- If this fails, then a shrinkage step is used.

Shrinkage



One Iteration of Nelder-Mead Method

1. **Sort.** Evaluate f at the $n+1$ vertices of S and sort the vertices such that $f_0 \leq f_1 \leq \cdots \leq f_n$.
2. **Reflection.** Compute the reflection point $\mathbf{v}_r$ from

$$\mathbf{v}_r = \bar{\mathbf{v}} + \alpha(\bar{\mathbf{v}} - \mathbf{v}_n).$$

   Evaluate $f_r = f(\mathbf{v}_r)$. If $f_0 \leq f_r < f_{n-1}$, replace $\mathbf{v}_n$ with $\mathbf{v}_r$.
3. **Expansion.** If $f_r < f_0$ then compute the expansion point $\mathbf{v}_e$ from

$$\mathbf{v}_e = \bar{\mathbf{v}} + \beta(\mathbf{v}_r - \bar{\mathbf{v}})$$

   and evaluate $f_e = f(\mathbf{v}_e)$. If $f_e < f_r$, replace $\mathbf{v}_n$ with $\mathbf{v}_e$; otherwise replace $\mathbf{v}_n$ with $\mathbf{v}_r$.
4. **Outside Contraction.** If $f_{n-1} \leq f_r < f_n$, compute the outside contraction point

$$\mathbf{v}_{oc} = \bar{\mathbf{v}} + \gamma(\mathbf{v}_r - \bar{\mathbf{v}})$$

   and evaluate $f_{oc} = f(\mathbf{v}_{oc})$. If $f_{oc} \leq f_r$, replace $\mathbf{v}_n$ with $\mathbf{v}_{oc}$. Otherwise, go to step 6.
5. **Inside Contraction.** If $f_r \geq f_n$, compute the inside contraction point x_{ic} from

$$\mathbf{v}_{ic} = \bar{\mathbf{v}} - \gamma(\mathbf{v}_r - \bar{\mathbf{v}})$$

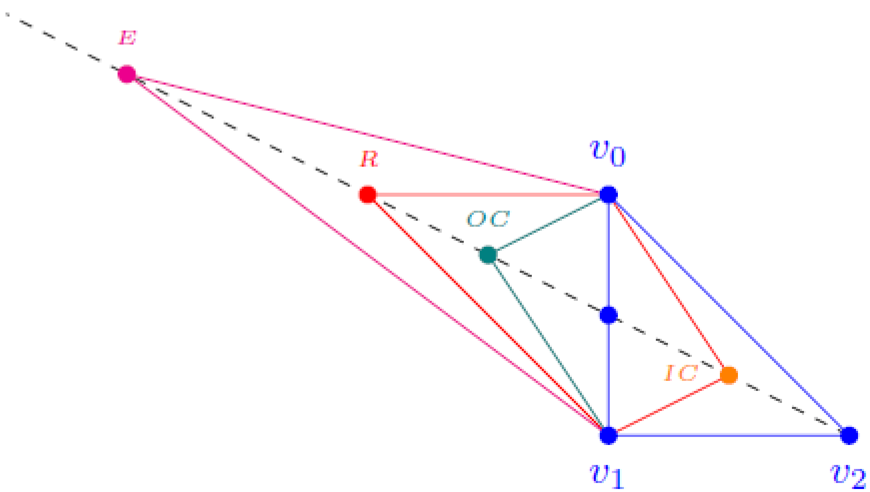
   and evaluate $f_{ic} = f(\mathbf{v}_{ic})$. If $f_{ic} < f_n$, replace $\mathbf{v}_n$ with $\mathbf{v}_{ic}$; otherwise, go to step 6.
6. **Shrinkage.** For $1 \leq i \leq n$, define

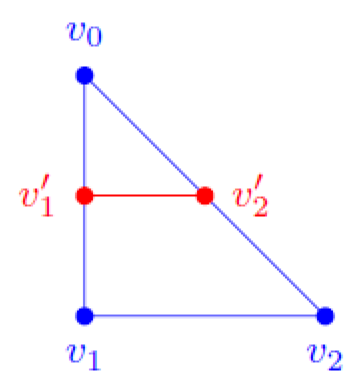
$$\mathbf{v}_i = \mathbf{v}_0 + \delta(\mathbf{v}_i - \mathbf{v}_0).$$

Standard implementation (e.g., in Matlab's FMINSEARCH)

- The parameters are chosen as:

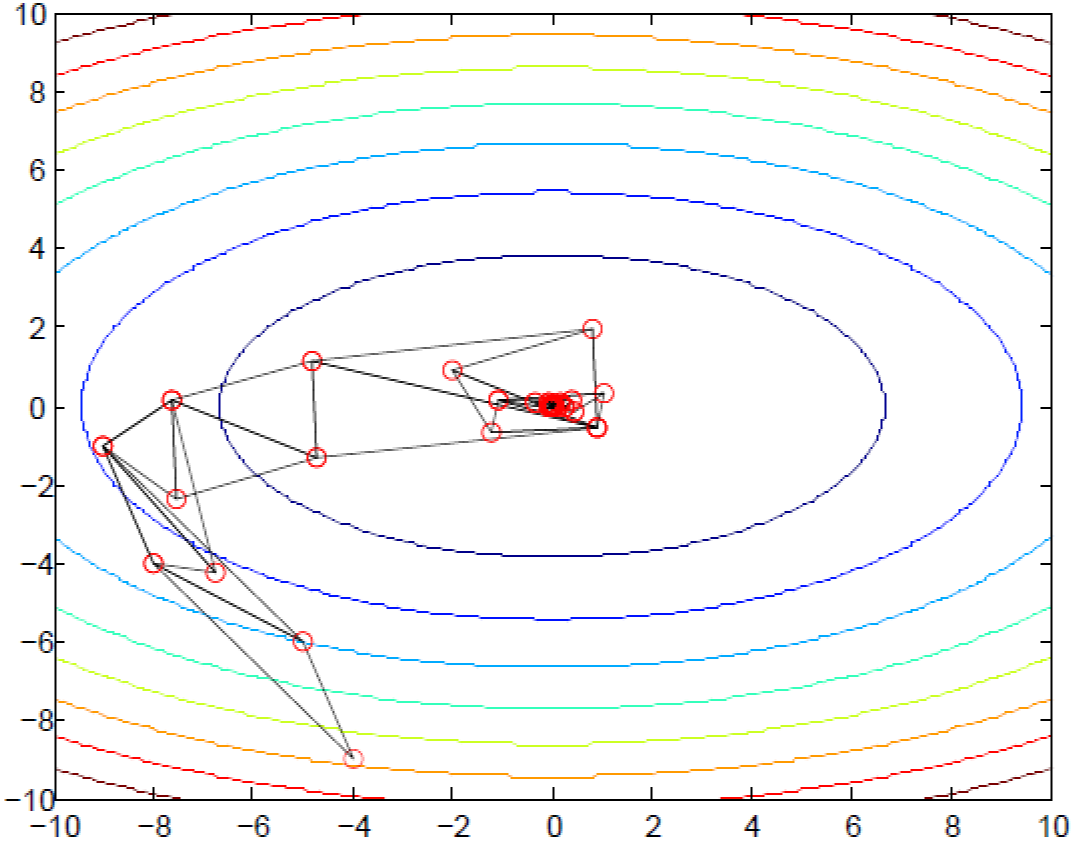
$$\alpha = 1,\ \beta = 2,\ \gamma = \frac{1}{2},\ \delta = \frac{1}{2}$$





Nelder-Mead on function:

$$x_1^2 + 3x_2^2$$



Convergence theory
(Lagarias, Reeds, Wright, and Wright, 1998)

Assuming that the objective function is strictly convex with bounded level sets, they proved

(a) When dimension n=1, the Nelder-Mead method converges to the minimizer.

(b) When dimension n=2, the Nelder-Mead simplex diameters converge to zero.

(Remark: (b) does not mean that the Nelder Mead method converges to the minimizer.)

Convergence theory (continued)

In 2012, Lagarias, Poonen, and Wright proved:

When n=2, if the objective function is strictly convex with bounded level sets, then the Restricted Nelder-Mead Method converges to the minimizer.

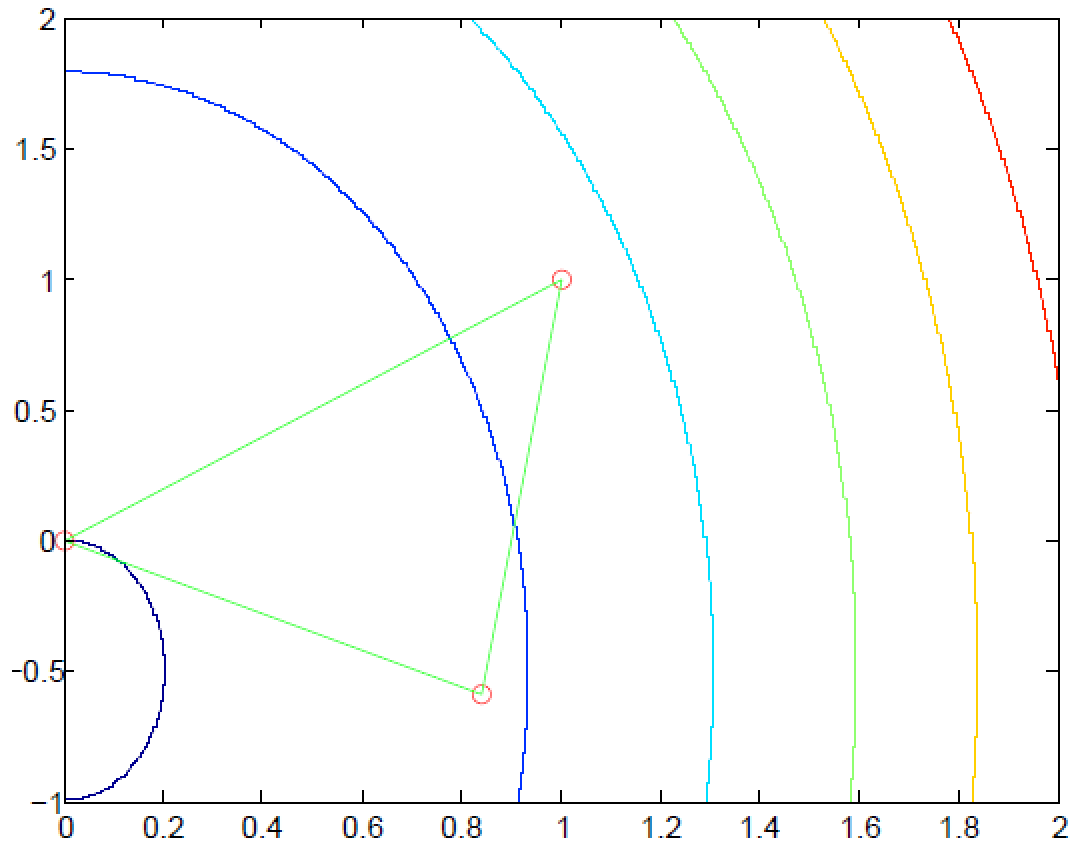
Remark: In the Restricted Nelder-Mead method, no expansion step is allowed.

- We don't even know if the standard Nelder-Mead simplex method always converges to the minimizer when the objective function is $f(x) = x_1^2 + x_2^2$!

- Non-convergent examples have been found.

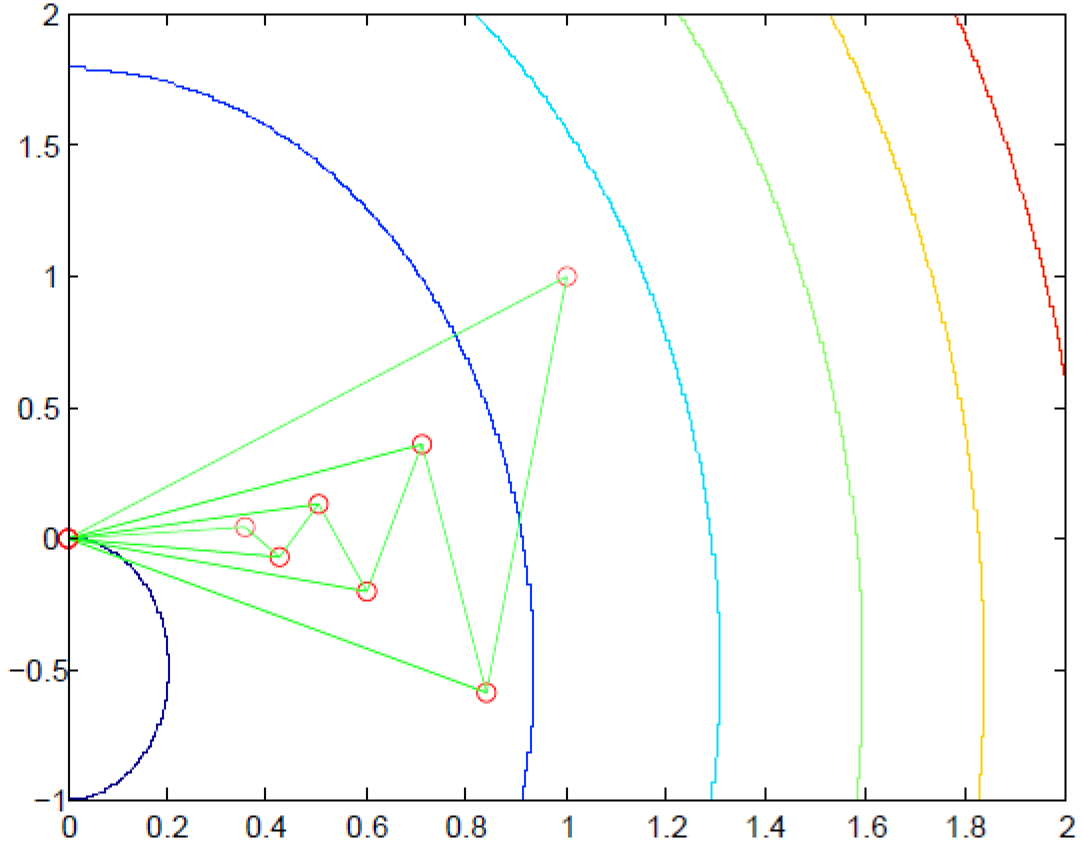
A non-convergent example (McKinnon, 1998)

$f = 360x_1^2 + x_2 + x_2^2$ if $x_1 \le 0$ and $f = 6x_1^2 + x_2 + x_2^2$ if $x_1 \ge 0$.



McKinnon's non-convergent example

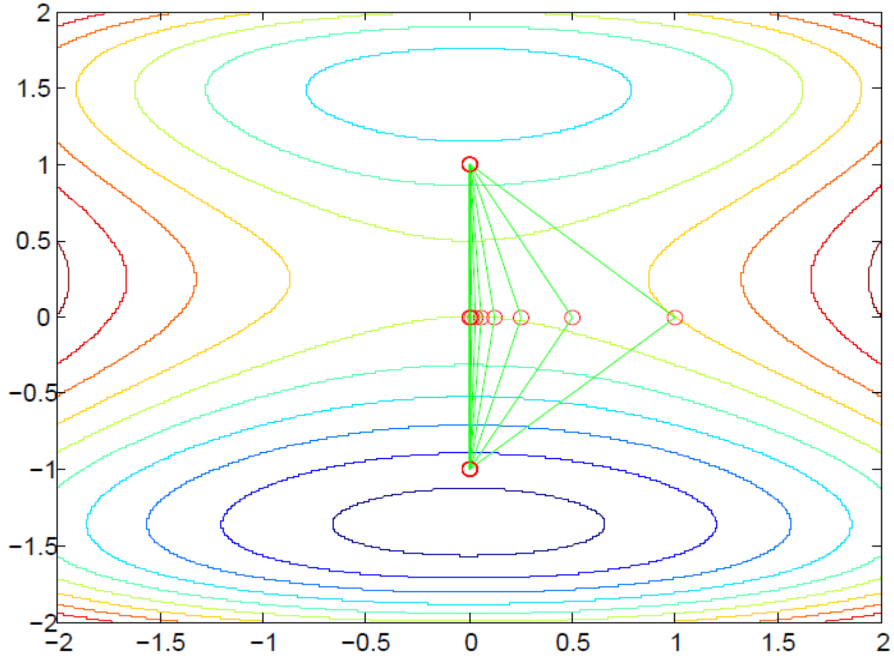
$f = 360x_1^2 + x_2 + x_2^2$ if $x_1 \leq 0$ and $f = 6x_1^2 + x_2 + x_2^2$ if $x_1 \geq 0$.



Another non-convergent example

$$f = x_1^2 + x_2(x_2 + 2)(x_2 - 0.5)(x_2 - 2)$$

- Repeated inside contraction, Nelder-Mead simplices converge to a line segment which does not contain any minimizer. (H, 2000)



Nelder-Mead method in practice

In spite of negative theoretical results, Nelder-Mead method has been one of the most popular derivative-free optimization methods. It is

- widely used in various applications
- implemented in MATLAB (as FMINSEARCH), SCILAB, …
- often claimed as "efficient and reliable" by practitioners
- recommended for solving optimization problems with noisy objective functions
- a citation classic (The Nelder-Mead paper had been cited 15687 times as of March 6, 2013 (Google Scholar).)

Effect of dimensionality

In an interview with Stephen Senn, John Nelder (2003) talked about his method: "Mathematicians hate it because you can't prove convergence; engineers seem to love it because it often works."

However, users of Nelder-Mead and optimization researchers do agree on one thing:

The (standard) Nelder-Mead method can be stagnant or very inefficient when dimension n is high.

Inefficient example

$$f(x) = x_1^2 + x_2^2 + \cdots + x_n^2$$

- In FMINSEARCH (the Matlab implementation of the standard Nelder Mead method), use

$$TolX = 10^{-8}, \quad TolFun = 10^{-8}$$

- Dimension n=32: After 196148 iterations and 231518 function evaluations, the Nelder-Mead method finds an approximate solution with the objective function value: 6.67×10^{-5}

Stagnant example: Extended Rosenbrock banana function

- Objective function:

$$f(x) = \sum_{k=1}^{n/2} 100(x_{2k} - x_{2k-1}^2)^2 + (1 - x_{2k-1})^2$$

Minimizer $= (1,1,\ldots,1)$; minimum value $= 0$.

- When dimension n=18, use FMINSEARCH (Matlab implementation of the standard Nelder Mead method):

  After one million iterations, the objective function value is still 19.8.

Study on effect of dimensionality

(H, Neumann, 2006)

- Focus on $f(x) = x^T x = x_1^2 + x_2^2 + \cdots + x_n^2$.
- The minimizer $\mathbf{0} = (0, 0, \cdots, 0)$ is chosen as a vertex of the initial simplex. It will remain as the best vertex in subsequent simplices.
- According to Lagarias et al. (1998): In this case, Nelder-Mead method uses only reflection, outside contraction, or inside contraction. (No expansion or shrinkage steps.)
- Investigate the effect of dimensionality on the rate of convergence of the Nelder-Mead method in this case.

Rate of convergence

For simplex S_k with vertices $\mathbf{0}, \mathbf{v}_1^{(k)}, \mathbf{v}_2^{(k)}, \cdots, \mathbf{v}_n^{(k)}$, define its oriented length by

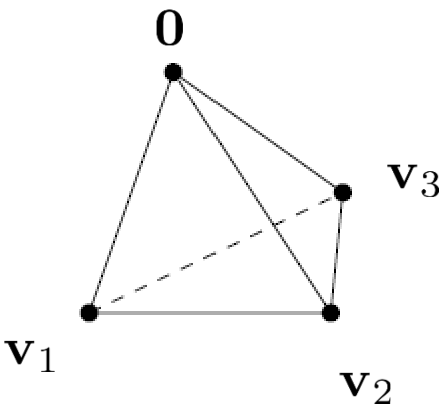
$$\sigma(S_k) := \max_{1 \le j \le n} \|\mathbf{v}_j^{(k)}\|.$$

A sequence of simplices $\{S_k\}$ converges to the minimizer $\mathbf{0} \in R^n$ if $\lim_{k\to\infty} \sigma(S_k) = 0$.
We measure the *rate of convergence* by

$$\rho(S_0, n) := \limsup_{k\to\infty} \left(\prod_{i=0}^{k-1} \frac{\sigma(S_{i+1})}{\sigma(S_i)} \right)^{1/k} \quad \left(= \limsup_{k\to\infty} (\sigma(S_k))^{1/k} \right).$$

The closer the rate $\rho(S_0, n)$ to 1, the slower the convergence.

Oriented length



$$\sigma(S) = \max\{\|\mathbf{v}_1\|, \|\mathbf{v}_2\|, \|\mathbf{v}_3\|\}$$

The maximum of lengths of edges: $\mathbf{0v}_1$, $\mathbf{0v}_2$, and $\mathbf{0v}_3$.

In our analysis, we

- analyze the behavior of the Nelder-Mead method when one type of operation (reflection, inside contraction, or outside contraction) is used repeatedly.
- assume that the worst vertex becomes the second best vertex at each iteration (Note that $\mathbf{0} = (0, 0, \cdots, 0)$ is always the best vertex.)

Therefore, the vertices at step k can be written as

$$\mathbf{0}, \mathbf{v}^{(k+n-1)}, \cdots, \mathbf{v}^{(k+1)}, \mathbf{v}^{(k)}$$

which satisfy

$$f(\mathbf{0}) < f(\mathbf{v}^{(k+n-1)}) < \cdots < f(\mathbf{v}^{(k)}).$$

Recurrence formulas

If one type of operation is performed repeatedly, then the vertices satisfy the recurrence relation:

$$\mathbf{v}^{(k+n)} = a\big(\mathbf{v}^{(k+1)} + \cdots + \mathbf{v}^{(k+n-1)}\big) + b\mathbf{v}^{(k)}$$

where the values of *a* and *b* are:

- Outside contraction: $a = \frac{3}{2n}, \quad b = -\frac{1}{2}$
- Inside contraction: $a = \frac{1}{2n}, \quad b = \frac{1}{2}$
- Reflection: $a = \frac{2}{n}, \quad b = -1$

Characteristic equation

The recurrence equation

$$\mathbf{v}^{(k+n)} = a(\mathbf{v}^{(k+1)} + \cdots + \mathbf{v}^{(k+n-1)}) + b\mathbf{v}^{(k)}$$

is linear with constant coefficients. Its characteristic equation is of the form

$$\mu^n = a(\mu + \cdots + \mu^{n-1}) + b,$$

which is a degree n polynomial. Its roots are called the characteristic values.

<u>**Remark**</u>: When $k \to \infty$, the behavior of $\mathbf{v}^{(k)}$ and therefore the Nelder-Mead simplices depend on the characteristic values with the largest modulus.

Properties of the characteristic values

Theorem (H, Neumann, Xu, 2003):

Let $\mu_{oc}(n)$, $\mu_{ic}(n)$, and $\mu_r(n)$ be a root of the characteristic equations for outside contraction, inside contraction, and reflection respectively. Then we have

$$|\mu_{oc}(n)| < 1, \quad |\mu_{ic}(n)| < 1, \quad |\mu_r(n)| = 1,$$

and

$$\lim_{n\to\infty} |\mu_{oc}(n)| = 1, \quad \lim_{n\to\infty} |\mu_{ic}(n)| = 1.$$

Largest and second largest moduli of characteristic values

To see how the characteristic values with the largest moduli affects the rate of convergence, we computed the largest and second largest moduli of the characteristic values for inside contraction and outside contraction.

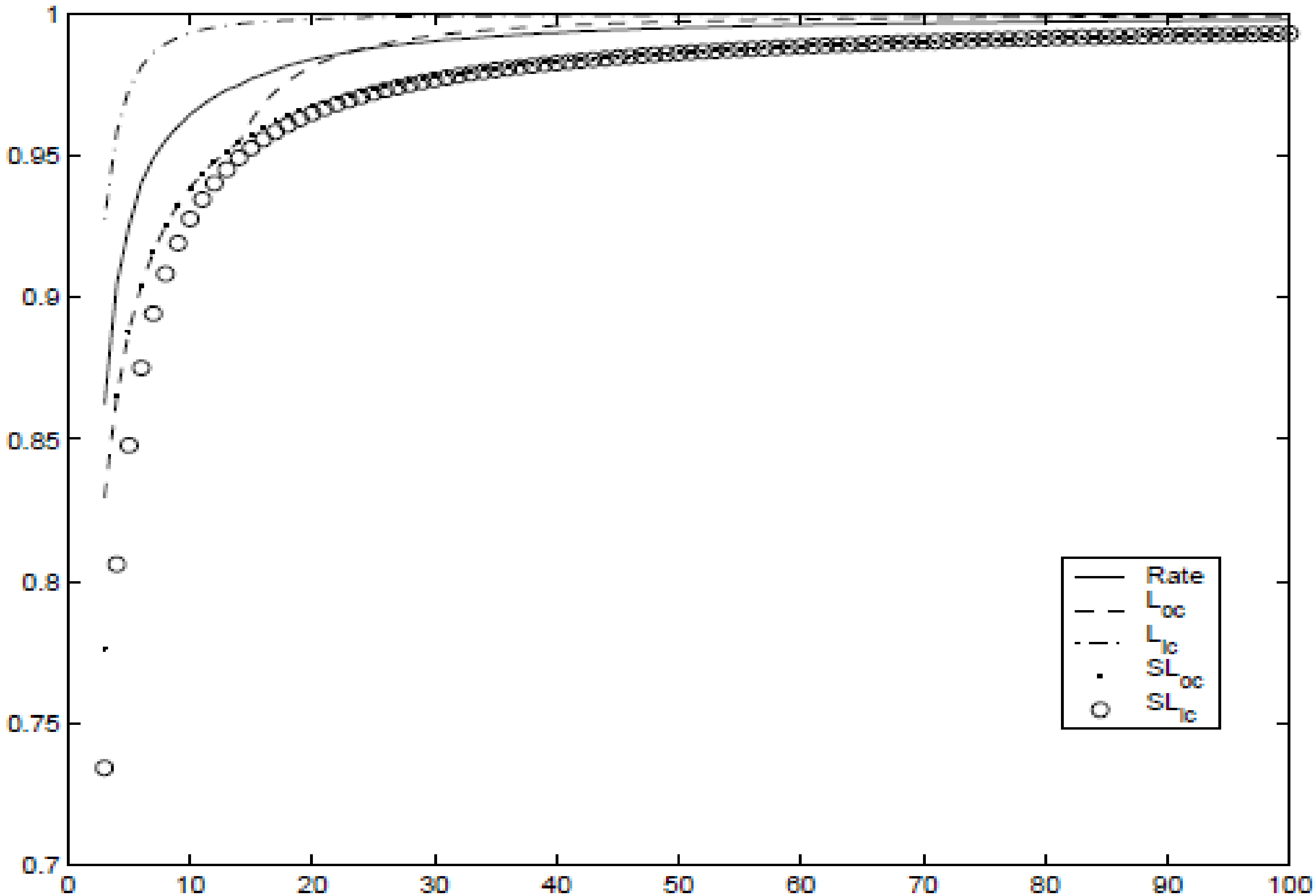
Denote the largest and second largest moduli of characteristic values for *Outside Contraction* by

$$L_{oc}(n) \text{ and } SL_{oc}(n)$$

respectively, and the largest and second largest moduli of characteristic values for *Inside Contraction* by

$$L_{ic}(n) \text{ and } SL_{ic}(n)$$

Rate of Convergence ($3 \leq n \leq 100$)



Conclusion: Effect of dimensionality on the Nelder-Mead method

As n becomes larger and larger, the rate of convergence becomes closer and closer to 1 and therefore, the Nelder-Mead method makes less and less progress per iteration (on average) toward the minimizer. We expect that $\lim_{n\to\infty} \rho(S_0, n) = 1$. It should be noted that when $n = 32$ which is moderately large for a derivative-free optimization method, the rate of convergence is $\rho(S_0, 32) = 0.9912$, already very close to 1.

Some new insights

In the previous analysis, we used the standard quadratic function $f(x) = x_1^2 + x_2^2 + \cdots + x_n^2$ and set its minimizer $(0, 0, \cdots, 0)$ as a vertex of the initial simplex.

In Gao and H (2012), they assume

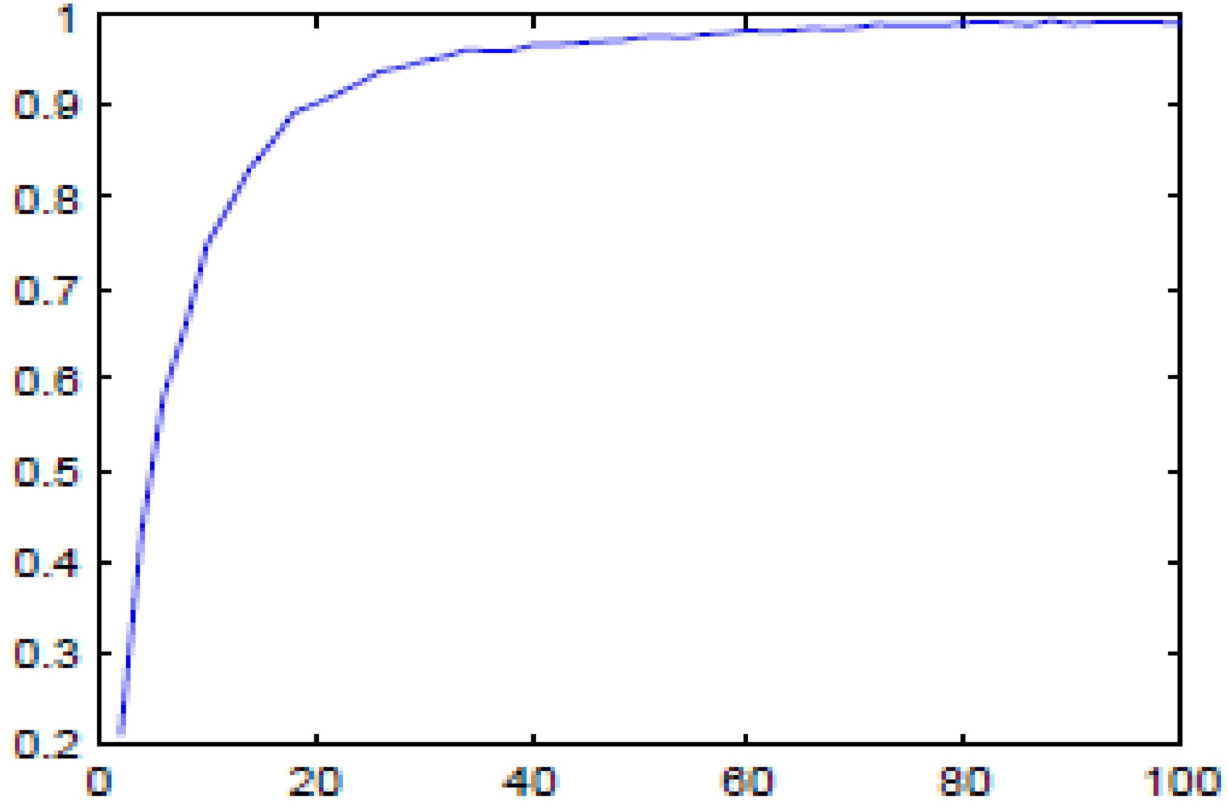
- uniformly convex objective function (e.g., second derivative matrix is positive definite and its smallest eigenvalue is bounded below by a positive constant.).
- general initial simplex.

Under these assumptions, we have

For the standard Nelder-Mead method,

- expansion, inside contraction, outside contraction produce good reduction in the objective function.
- reflection often does not. (and unfortunately, the standard Nelder-Mead method uses too many reflection steps.)
- larger simplex decreases the objective function value more.

Standard Nelder-Mead: Fraction of reflections steps against problem dimension of $\mathbf{x}^T\mathbf{x}$



A new implementation of Nelder-Mead method

To tackle the dimensionality issue of the standard Nelder-Mead method, we propose a new implementation which takes the dimension of the optimization problem into account.

Adaptive parameters

- Recall that the standard implementation such as FMINSEARCH in Matlab uses parameters:

$$\alpha = 1,\ \beta = 2,\ \gamma = \tfrac{1}{2},\ \delta = \tfrac{1}{2}$$

call it: SNMS

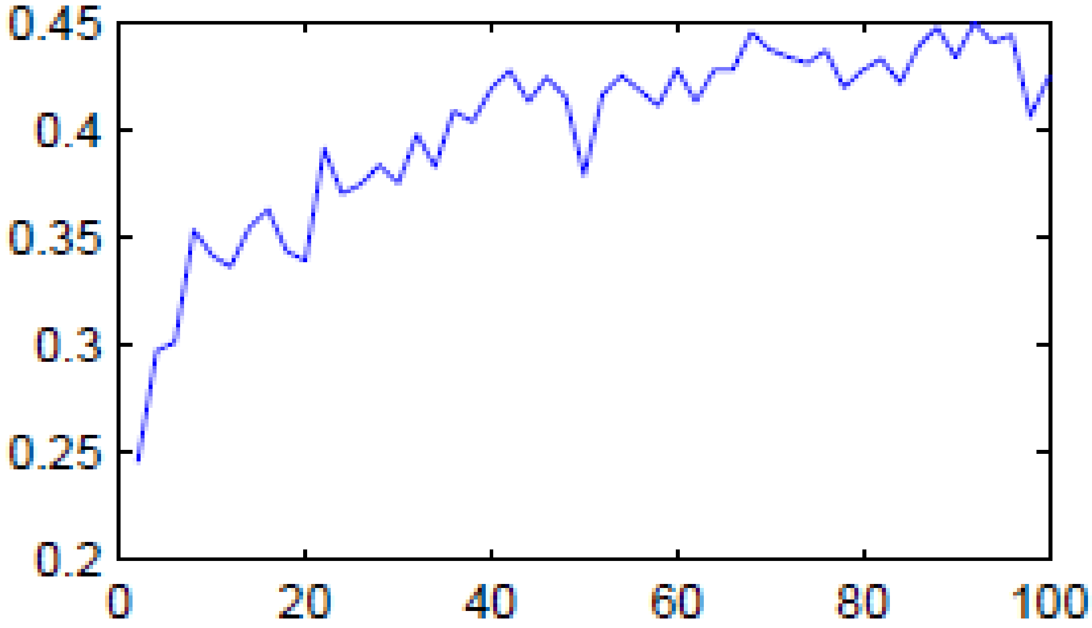
- New implementation uses adaptive parameters:

$$\alpha = 1,\ \beta = 1 + \tfrac{2}{n},\ \gamma = 0.75 - \tfrac{1}{2n},\ \delta = 1 - \tfrac{1}{n}$$

(where n is the dimension of the problem.). Call it: ANMS

- The two implementations are identical when $n=2$, but different when $n \geq 3$.

New implementation: Fraction of reflections steps against problem dimension of $\mathbf{x}^T\mathbf{x}$



Numerical Test on modified Byrd-Nocedal-Zhu quartic function

$$\min_{\mathbf{x}\in\mathbb{R}^n} f(\mathbf{x}) = \mathbf{x}^T D\mathbf{x} + \sigma(\mathbf{x}^T B\mathbf{x})^2,$$

where D is a positive definite matrix of the form

$$D = \mathrm{diag}([1+\epsilon, (1+\epsilon)^2, \cdots, (1+\epsilon)^n])$$

and B the positive definite matrix

$$B = U^T U,\ U = \begin{bmatrix} 1 & \cdot & \cdot & \cdot & 1 \\ & \cdot & & & \cdot \\ & & \cdot & & \cdot \\ & & & \cdot & \cdot \\ & & & & 1 \end{bmatrix},$$

		SNMS		ANMS	
(ϵ, σ)	Dim	nfeval	Final f	nfeval	Final f
$(0.05, 0)$	10	1123	1.1166×10^{-7}	910	9.0552×10^{-9}
$(0.05, 0)$	20	9454	2.7389×10^{-7}	2548	1.8433×10^{-8}
$(0.05, 0)$	30	55603	5.3107×10^{-3}	5067	2.6663×10^{-8}
$(0.05, 0)$	40	99454	1.5977×10^{-2}	8598	3.6816×10^{-8}
$(0.05, 0)$	50	215391	1.6906×10^{-1}	13167	6.7157×10^{-8}
$(0.05, 0)$	60	547475	$1.2685 \times 10^{+1}$	20860	6.8945×10^{-8}
$(0.05, 0.0001)$	10	1787	3.1878×10^{-8}	994	6.0454×10^{-9}
$(0.05, 0.0001)$	20	20824	$1.2984 \times 10^{+1}$	3788	1.5294×10^{-8}
$(0.05, 0.0001)$	30	39557	$1.8108 \times 10^{+2}$	10251	4.0331×10^{-8}
$(0.05, 0.0001)$	40	71602	$4.3797 \times 10^{+2}$	18898	5.7407×10^{-8}
$(0.05, 0.0001)$	50	87660	$8.0726 \times 10^{+2}$	37282	4.7431×10^{-7}
$(0.05, 0.0001)$	60	136991	$1.5369 \times 10^{+3}$	61259	2.0786×10^{-7}

Test on More-Garbow-Hillstrom collection of functions

1. `band`: Broyden banded function;
2. `bv`: Discrete boundary value function;
3. `ie`: Discrete integral equation function;
4. `lin`: Linear function - full rank;
5. `pen1`: Penalty I function;
6. `pen2`: Penalty II function;
7. `rosenbrock`: Extended Rosenbrock function;
8. `singular`: Extended Powell singular function;
9. `trid`: Broyden tridiagonal function;
10. `trig`: Trigonometric function;
11. `vardim`: Variably dimensioned function

		SNMS		ANMS	
Prob	Dim	nfeval	Final f	nfeval	Final f
band	10	1069	2.0581×10^{-6}	741	2.2149×10^{-7}
band	20	5213	8.3121×10^{-5}	1993	5.1925×10^{-7}
band	30	29083	12.1584	3686	6.4423×10^{-7}
band	40	20824	2.4717×10^{-4}	6060	1.0892×10^{-6}
band	50	37802	2.6974×10^{-4}	8357	1.2359×10^{-6}
band	60	38755	2.3999×10^{-4}	10630	1.0002×10^{-6}
bv	10	863	9.5451×10^{-9}	1029	1.0388×10^{-9}
bv	20	5553	7.8216×10^{-6}	7535	3.1789×10^{-10}
bv	30	23150	1.0294×10^{-5}	3860	3.0035×10^{-5}
bv	40	862	1.6788×10^{-5}	1912	1.6110×10^{-5}
bv	50	1087	8.9953×10^{-6}	1905	8.8601×10^{-6}
bv	60	1507	5.3411×10^{-6}	2125	5.3085×10^{-6}
ie	10	1123	5.0253×10^{-9}	774	9.5926×10^{-9}
ie	20	6899	1.2029×10^{-5}	3320	1.0826×10^{-8}
ie	30	43231	0.0015	8711	2.1107×10^{-8}
ie	40	61575	3.7466×10^{-4}	18208	4.0741×10^{-8}
ie	50	155635	0.0031	25961	4.7628×10^{-8}
ie	60	148851	5.1304×10^{-4}	38908	2.2644×10^{-7}

		SNMS		ANMS	
Prob	Dim	nfeval	Final f	nfeval	Final f
lin	10	1974	1.7816×10^{-8}	1020	5.5242×10^{-9}
lin	20	15401	0.0104	3009	1.1136×10^{-8}
lin	30	57260	0.4494	5310	2.1895×10^{-8}
lin	40	83928	0.5527	8025	1.8607×10^{-8}
lin	50	183633	0.0493	11618	1.9984×10^{-8}
lin	60	475121	0.0280	16492	2.7243×10^{-8}
pen1	10	3909	7.5725×10^{-5}	5410	7.0877×10^{-5}
pen1	20	21680	$8.6799 \times 10^{+3}$	14995	1.5778×10^{-4}
pen1	30	64970	$5.1216 \times 10^{+5}$	45852	2.4773×10^{-4}
pen1	40	254995	$2.8566 \times 10^{+5}$	86293	3.3925×10^{-4}
pen1	50	287599	$4.4971 \times 10^{+6}$	198719	4.3179×10^{-4}
pen1	60	654330	$2.6319 \times 10^{+6}$	254263	5.2504×10^{-4}
pen2	10	4017	2.9787×10^{-4}	9741	2.9366×10^{-4}
pen2	20	27241	0.0065	11840	6.3897×10^{-3}
pen2	30	37774	0.0668	16882	0.0668
pen2	40	116916	29.6170	27211	0.5569
pen2	50	204871	4.2997	43444	4.2961
pen2	60	680176	48.1215	55346	32.2627

		SNMS		ANMS	
`Prob`	`Dim`	`nfeval`	`Final f`	`nfeval`	`Final f`
rosenbrock	6	2141	2.1314	1833	1.3705×10^{-9}
rosenbrock	12	6125	14.316	10015	3.3974×10^{-9}
rosenbrock	18	13357	22.000	29854	4.2290×10^{-9}
rosenbrock	24	17156	29.119	50338	4.2591×10^{-9}
rosenbrock	30	19678	50.889	156302	5.4425×10^{-9}
rosenbrock	36	43870	52.201	119135	1.6616×10^{-8}
singular	12	2791	9.5230×10^{-6}	5199	3.9417×10^{-8}
singular	24	15187	3.8012×10^{-4}	11156	4.8767×10^{-6}
singular	32	37754	8.4318×10^{-5}	37925	4.6217×10^{-6}
singular	40	80603	0.0039	38530	9.9115×10^{-6}
singular	52	120947	0.0032	73332	1.8319×10^{-5}
singular	60	233482	0.0024	71258	1.9181×10^{-5}
trid	10	908	6.6529×10^{-7}	740	2.5511×10^{-7}
trid	20	3308	2.7137×10^{-6}	3352	2.9158×10^{-7}
trid	30	7610	1.6093×10^{-5}	11343	3.6927×10^{-7}
trid	40	13888	9.8698×10^{-6}	23173	4.4076×10^{-7}
trid	50	24008	1.7782×10^{-5}	42013	5.0978×10^{-7}
trid	60	34853	2.0451×10^{-5}	64369	7.1834×10^{-7}

		SNMS		ANMS	
Prob	Dim	nfeval	Final f	nfeval	Final f
trig	10	2243	2.7961×10^{-5}	961	2.7952×10^{-5}
trig	20	12519	1.6045×10^{-6}	4194	1.3504×10^{-6}
trig	30	19754	3.5273×10^{-5}	8202	9.9102×10^{-7}
trig	40	23938	1.69412×10^{-5}	17674	1.5598×10^{-6}
trig	50	25328	2.9162×10^{-5}	19426	3.6577×10^{-7}
trig	60	33578	4.8213×10^{-5}	31789	9.6665×10^{-7}
vardim	6	1440	5.3381×10^{-9}	1170	5.9536×10^{-9}
vardim	12	3753	6.6382	4709	8.6227×10^{-9}
vardim	18	6492	8.8146	12815	1.0898×10^{-8}
vardim	24	13844	71.320	35033	1.1237×10^{-8}
vardim	30	19769	85.397	67717	1.5981×10^{-8}
vardim	36	32360	72.101	209340	1.8116×10^{-8}

Summary

(1) The performance of the standard Nelder-Mead simplex method deteriorates as the dimension of the optimization problem increases.

(2) Expansion and contraction (inside or outside) steps help decrease the objective value. Reflection step often does not.

(3) Large simplex helps decrease the objective function value.

(4) Implementation using adaptive parameters is promising for higher dimensional problems.

Future research

- Analyze the convergence properties of Nelder-Mead method for dimension $n \geq 3$.
- Investigate how to set up a suitable initial simplex. (Larger initial simplex can improve performance.)
- Study how to effectively reduce the use of reflection steps.
- Develop new derivative-free algorithms for high dimensional problems.

References

- F. Gao and L. Han, Implementing the Nelder-Mead simplex algorithm with adaptive parameters, Comput. Optim. Appl., 2012, 51: 259-277.
- L. Han, M. Neumann, Effect of dimensionality on the Nelder-Mead simplex method, Optim. Methods Softw., 2006, 21(1): 1-16.
- L. Han, M. Neumann, and J. Xu, On the roots of certain polynomials arising from the analysis of the Nelder-Mead simplex method, Linear Algebra and its applications, 2003, 363: 109-124.
- J. Lagarias, B. Poonen, and M. Wright, Convergence of the restricted Nelder-Mead method in two dimensions, SIAM J Optim., 2012, 22: 501-532.
- J. Lagarias, J. Reeds, M. Wright, and P. Wright, Convergence properties of the Nelder-Mead simplex algorithm in low dimensions, SIAM J Optim., 1998, 9: 112-147.
- K. McKinnon, Convergence of the Nelder-Mead simplex method to a nonstationary point, SIAM J. Optim., 1998, 9: 148-158.
- J. More, B. Garbow, and B. Hillstrom, Testing unconstrained optimization software, ACM Trans. Math. Softw., 1981, 7(1): 17-41.
- J. Nelder and R. Mead, A simplex method for function minimization, Computer Journal, 1965, 7: 308-313.

Thank You