Astronomical Data Analysis Software and Systems VIII
ASP Conference Series, Vol. 172, 1999
D. M. Mehringer, R. L. Plante, and D. A. Roberts, eds.

Optimizing Functions Using ASCFIT

David Grumm
Smithsonian Astrophysical Observatory, 60 Garden Street, MS 81, Cambridge, MA 02138 USA

Abstract. As part of the Smithsonian Astrophysical Observatory AXAF Science Center (ASC) Data Analysis Environment, ASCFIT was designed as a generalized fitting engine to enable users to fit models to the data that will be returned by AXAF, the Advanced X-ray Astrophysical Facility (Doe et al. 1996). I present a description of the use of a development version of ASCFIT as a function optimizer, in which the optimization methods available within ASCFIT are applied to seven benchmark functions of varying complexity. These functions are the 2-dimensional Gaussian function previously investigated by ASC, and six functions commonly used in optimization testing. The optimization methods combining the Powell technique with one of several other techniques successfully converged for all but one of the benchmark functions.

1. Introduction

The library of ASCFIT optimization methods was applied to a set of benchmark functions. The optimization methods are in the OPTIM package written by Mark Birkinshaw, and can be categorized as either "single-shot" routines, or "scatter-shot" routines (Birkinshaw 1998). In single-shot routines (Simplex & Powell), the optimization attempts to converge from the initial state to the global minimum in a continuous fashion. In scatter-shot routines (Grid, Monte Carlo, and Simulated Annealing), the entire parameter space is searched to determine a better minimum than the initial state. The functions were optimized by their minimization and vary from continuous functions with a single minimum to discontinuous functions with many local minima. Some of the methods start with a scatter-shot method, and then use a single-shot method to refine the search, as in Monte-Carlo + Powell. The two Simulated Annealing methods differ in that in method 2 all of the parameters are varied simultaneously, while in method 1 they are varied separately. The methods are:

Simplex	Powell
Grid	Grid + Powell
Monte Carlo	Monte Carlo + Powell
Simulated Annealing 1	Simulated Annealing 1 + Powell
Simulated Annealing 2	Simulated Annealing 2 + Powell

2. Functions Optimized by ASCFIT

The benchmark functions $\mathtt{f}_0$ through $\mathtt{f}_5$ are based on those of Ingber & Rosen (1992), who selected them as being typically used for genetic algorithm benchmarking. The 2-dimensional Gaussian function was chosen as it was previously investigated by ASC. These functions are described below.

Function 2dg(x, y) is a two-dimensional Gaussian function:

$$2dg(x,y) = 1 - Ae^{-\frac{(x-x0)^2+(y-y0)^2}{2\sigma^2}},$$
$$x0 = 30.0, y0 = 30.0, A = 1.0, \sigma = 10.0,\ |x| \le 100.0,\ |y| \le 100.0$$

The function $\mathtt{f}_0$ has coefficients s_i, t_i, d_i, and c defined such that the function is a paraboloid with axes parallel to the coordinate axes, and a set of holes that increase in depth near the origin. The function is very difficult to minimize, having over 10^{20} local minima in which the optimization may become trapped:

$$f_0(x_1,...,x_4) = \sum_{i=1}^{4}(t_i sgn(z_i) + z_i)^2 c d_i) \text{ if } |x_i - z_i| < |t_i|$$
$$= \sum_{i=1}^{4} d_i x_i^2 \text{ otherwise},$$
$$z_i = \lfloor |x_i/s_i| + 0.49999 \rfloor sgn(x_i) s_i,$$
$$s_i = 0.2, t_i = 0.05, i = 1, ...4,$$
$$d_i = \{1.0, 1000.0, 10.0, 100.0\},$$
$$c = 0.15,$$
$$-1000.0 \le x_i \le 1000.0, i = 1, ..., 4$$

The function f_1 is a simple sum of squares, with a single minimum:

$$f_1(x_1,x_2,x_3) = \sum_{j=1}^{3} x_j^2;\quad -5.12 \le x_i \le 5.12,\quad i = 1,...,3.$$

The function f_2 is the classical Rosenbrock function:

$$f_2(x_1,x_2) = 100(x_1^2 - x_2)^2 + (1 - x_1)^2;\quad -2.048 \le x_i \le 2.048,\quad i = 1,2.$$

The function f_3 is discontinuous and has a single minimum:

$$f_3(x_1,...,x_5) = 30.0 + \sum_{j=1}^{5} \lfloor x_j \rfloor;\quad -5.12 \le x_i \le 5.12,\quad i = 1,...,5.$$

The function f_4 is a noisy function, as η is a uniform random variable bounded by $[0, 1)$:

$$f_4(x_1,x_2) = \sum_{j=1}^{2} x_j^4 + \eta;\quad -1.28 \le x_i \le 1.28,\quad i = 1,2.$$

Function f_5 has 25 local minima, and is optimized with a value of approximately 0.001996 when $x_1 = x_2 = -32$:

$$f_5(x_1,x_2) = \frac{1}{500 + \sum_{j=1}^{25} \frac{1}{j + \sum_{i=1}^{2}(x_i - a_{ji})^6}}$$
$$a_{j1} = \left\{ \begin{array}{c} -32,-16,0,16,32,-32,-16,0,16,32,-32,-16,0, \\ 16,32,-32,-16,0,16,32,-32,-16,0,16,32 \end{array} \right\},$$
$$a_{j2} = \left\{ \begin{array}{c} -32,-32,-32,-32,-32,-16,-16,-16,-16,-16,0, \\ 0,0,0,0,16,16,16,16,16,32,32,32,32,32 \end{array} \right\},$$
$$-65.536 \le x_i \le 65.536,\quad i = 1,2.$$

3. Results

For each optimization method and function, the initial parameter state was first chosen to be fairly distant from the global minimum, and all ASCFIT default parameters were used (scenario A). The optimization was then run using stricter convergence criteria and the same initial state (scenario B), indicative of the case when one is optimizing a function whose global minimum is not even approximately known (More, Garbow, & Hillstrom 1981). Finally, the optimization was run using the stricter convergence criteria and an initial parameter state chosen to lie much closer to the correct state (scenario C).

Table 1 summarizes the results for scenario B. For each function and technique, the value shown is the normalized final parameter distance, defined as the ratio of the distance between the final parameter state and the global minimum to the maximum distance between any two states within the allowed parameter space. Distances of less than 10^{-6} are shown as 10^{-6}. In general, the results for scenario C (not shown) indicate that, as expected, a more favorable initial state improved the single-shot routines' optimizations of smoothly varying functions. The final column is the execution time for the gaussian, normalized to the time for the simplex method. The result of one technique is plotted in Fig. 1, with the initial parameter distances for scenarios A, B and C.

Table 1. Function optimizations for scenario B.

METHOD	2dg	f_0	f_1	f_2	f_3	f_4	f_5	TIME
Simplex	3E-1	3E-4	2E-5	4E-5	1E-6	3E-1	2E-1	1.0
Powell	4E-6	2E-1	1E-5	2E-5	1E-6	3E-1	3E-1	1.3
Grid	3E-4	1E-6	2E-1	8E-5	6E-2	1E-1	1E-6	76.8
Grid+Powell	4E-6	2E-5	2E-5	1E-4	1E-6	4E-1	1E-4	2435.5
M. C.	6E-4	3E-2	2E-1	1E-3	6E-2	2E-1	1E-6	76.0
M. C.+Powell	1E-6	2E-5	1E-5	7E-5	1E-6	4E-1	1E-6	477.0
S. A. 1	3E-5	2E-4	4E-3	1E-3	1E-6	3E-1	1E-6	35.8
S. A. 1+Powell	1E-6	1E-5	1E-5	6E-5	1E-6	4E-1	1E-6	282.8
S. A. 2	4E-5	2E-4	1E-2	1E-4	3E-2	2E-1	1E-6	1.3
S. A. 2+Powell	2E-5	9E-6	8E-5	1E-6	1E-6	4E-1	1E-6	28.8

4. Discussion

Using optimization scenario C, the global minimum was located for all functions except f_4, in which the routines became trapped in local minima. For this function, the grid and simulated annealing techniques would presumably locate the minimum if strict enough criteria were used; the grid method showed improved optimization for scenario B compared to scenario A. Ingber & Rosen (1992) tested a similar function with 30 variables and obtained comparable results.

The single-shot routines successfully converged for the smoothly varying functions (2dg, f_1, and f_2), particularly when starting from a favorable parameter state. For the discontinuous functions, these routines sometimes became

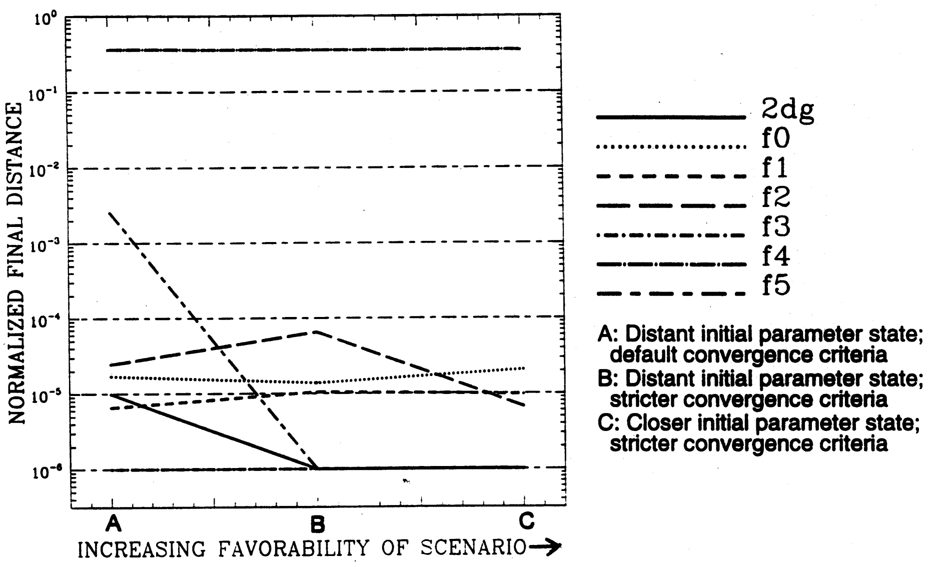

Figure 1. Monte Carlo + Powell.

trapped in local minima. In these cases, the optimizations were not improved substantially by using an initial parameter state closer to the global minimum or by using stricter convergence criteria.

The scatter-shot routines were less successful in optimizing the smooth functions, as they sample the entire parameter space without refining the search near the minimum. However, for both the smooth and the discontinuous functions, the addition of the Powell technique improved the optimization, at the expense of longer execution times (see Table 1). Such methods with refining steps following sampling of the parameter space are to be preferred.

Acknowledgments. This project is supported by NASA contract NAS8-39073. I would like to thank the following individuals for helpful discussions and their review: Mark Birkinshaw, Janet DePonte, Stephen Doe, Diab Jerius, Malin Ljungberg, Dan Nguyen, Aneta Siemiginowska, and Robert Zacher.

References

Birkinshaw, M. 1998, Some Advice on Minimization Routines, (AXAF Science Center Internal Memo)

Doe, S., Ljungberg, M., Siemiginowska, A., & Joye, W. 1996, in ASP Conf. Ser., Vol. 145, Astronomical Data Analysis Software and Systems VII, ed. R. Albrecht, R. N. Hook, & H. A. Bushouse (San Francisco: ASP), 157

Ingber, L. & Rosen, B. 1992, Mathematical and Computer Modelling, 16(11), 87

More, J., Garbow, B. S., Hillstrom, K. E. 1981, ACM Trans on Math. Software, 7, 17