NONLINEAR OPTIMISATION AND THE ASTEROID IDENTIFICATION PROBLEM

M. EUGENIA SANSATURIO
E.T.S.I.I. University of Valladolid, Spain

AND

ANDREA MILANI AND LUISA CATTANEO
Space Mechanics Group. University of Pisa, Italy

Differential correction procedure allows us to improve orbits for which new observations are available; however, it only works provided the original orbit is within the convergence domain of the pseudo–Newton method. Given the strong nonlinearity of the problem, this only occurs when the residuals of the new observations with respect to the old orbit are quite small.

On the contrary, if a single opposition asteroid, observed only on a short arc, is "lost", i.e. not recovered for several years, it can be difficult to identify it with a newly observed one. There are now $\simeq 20,000$ lost asteroids with poorly determined orbits; if the proposed *Spaceguard survey* will be realized, the problem of identifying millions of new discoveries within catalogues of comparable size will be one of the main challenges. We have begun experimenting with algorithms of orbit determination, to perform both positive and negative identification of asteroids lost for many years.

1. Algorithms

Let us recall the definition of the classical algorithm of *differential corrections* for orbit improvement. If the *residuals* are ξ_i, $(i = 1, m)$ and the *weights* are $W = Diag[{\sigma_i}^{-2}]$, the *target function* is

$$Q = \frac{1}{m}\xi^T W \xi = \frac{1}{m}\sum_{i=1}^{m}\frac{{\xi_i}^2}{{\sigma_i}^2}\,.$$

If y are the solve for variables, such as the orbital elements at some epoch time, then the least square solution is $\bar{y}$ such that $Q(\bar{y}) = \min Q(y)$.

S. Ferraz-Mello et al. (eds.), Dynamics, Ephemerides and Astrometry of the Solar System, 193–198.

If the partial derivatives with respect to the solve for variables are available

$$G = -\frac{\partial \xi}{\partial y}\,, \qquad H = -\frac{\partial^2 \xi}{\partial y^2}\,,$$

then we can compute the gradient and the Hessian of the target function

$$\frac{\partial Q}{\partial y} = -\frac{1}{m} 2\xi^T W G\,, \qquad \frac{\partial^2 Q}{\partial y^2} = \frac{1}{m}(2G^T W G - 2\xi^T W H)\,.$$

Truncating the Taylor expansion of the gradient to order one and solving for a value of the change in y corresponding to a stationary point, we obtain a step of *Newton method*

$$\Delta y = \bar{y} - y = -\left[\left.\frac{\partial^2 Q}{\partial y^2}\right|_y\right]^{-1} \left.\frac{\partial Q}{\partial y}\right|_y = (G^T W G - \xi^T W H)^{-1} G^T W \xi\,,$$

where the normal system matrix is $\Gamma^{-1} = G^T W G - \xi^T W H$.

The *pseudo-Newton method* is obtained by replacing Γ^{-1} with an approximation which does not contain the Hessian matrix, that is: $G^T W G$. In both cases, the improvement procedure is stopped by a criterion, such as

$$||\Delta y|| = \sqrt{\frac{\Delta y^T\, \Gamma^{-1}\, \Delta y}{dim(y)}} << 1\,.$$

The differential corrections algorithm is very efficient when the initial estimate of y is close to the solution $\bar{y}$; however, the convergence domain of the Newton method is very small when the target function is strongly nonlinear, as is the case when it contains the solutions of an N-body problem. Moreover, the first derivative G is quite easy to compute by solving numerically the variational equations (the linearized equations for relative motion), but the second derivative H is not often available and using the pseudo-Newton approximation further shrinks the convergence domain.

Differential corrections are therefore a perfectly adequate method for improvement of the orbits already well observed, but they can fail to converge whenever the residuals before the first correction are large; moreover, in this case, the difference between the normal system of the Newton and of the pseudo-Newton method is not small at all. If a single opposition asteroid, observed only on a short arc, is not recovered for several years, then the position uncertainty of the asteroid grows to a very large size until the asteroid is "lost", that is, it can be in widely separated regions of the sky. In this case, if the same asteroid is observed again, the identification with an already known one can be very difficult, because the orbital elements as computed from a short arc can be wrong by a significant amount. Even if the identification is correct, when the two arcs are linked in a single orbit the initial residuals are typically very large and the algorithm can fail.

Two approaches are possible to solve this problem: either we use an optimisation method well suited to strongly nonlinear problems, capable of finding a minimum of the target function even when it is very far from the initial guess, or we define an algorithm to improve the first guess so that the initial residuals are not large (strategy described in the next section).

To make the strongly nonlinear identification problem easier to handle, we adopted a formulation closer to a linear problem. The orbit between the discovery and the supposed recovery of the asteroid is decomposed into a number N of shorter arcs, each with independent initial conditions; the requirement that the asteroid is the same can be translated into constraints on the junction of each couple of consecutive arcs. Thus, the initial value of the residuals is small, but the constraints are not initially satisfied. So, we need to use an algorithm capable of finding the solution of a constrained optimisation problem, with $6N$ variables. To this end, optimisation algorithms are known and have been implemented in a software system codenamed LANCELOT (*Large And Nonlinear Constrained Extended Lagrangian Optimization Techniques*) by Conn, Gould and Toint (1992).

LANCELOT includes two main algorithms: AUGLG (AUGmented LaGrangian) is used to solve the generally constrained problem: $\min\{Q(y)\}$, such that $\{y \in \Re^n,\ l_i \leq y_i \leq u_i,\ 1 \leq i \leq n;\ c_j(y) = 0,\ 1 \leq j \leq m\}$ and the *constraints* are: $c_j(y) = 0\,,\ j = 1, m$. This algorithm minimizes the Augmented Lagrangian function Φ, defined as

$$\Phi(y, \lambda, S, \mu) = Q(y) + \sum_{i=1}^{m} \lambda_i c_i(y) + \frac{1}{2\mu} \sum_{i=1}^{m} s_{ii} c_i(y)^2$$

using SBMIN, then it decreases μ until the constraints are satisfied (with a given accuracy). SBMIN (Simple Bound MINimization) solves the bound-constrained minimum problem: $\min\{Q(y) : y \in \Re^n,\ l_i \leq y_i \leq u_i,\ 1 \leq i \leq n\}$ by unidimensional search methods with confidence region.

In this paper we compare the results obtained in the asteroid identification problem by using the classical differential correction algorithm (DC) and by the use of LANCELOT. The observations and their partial derivatives are computed in both cases by the same subroutines, which solve the full N-body problem with variational equations, with very good accuracy (by Everhart (1985) method, an implicit Runge-Kutta of order 15).

2. Problems and Results

We have used data (observations and orbits) stored at the MPC to test how effective and reliable the classical DC and nonlinear optimisation methods –such as LANCELOT– are to perform positive and negative identifications of single opposition (lost) asteroids. The test cases have been obtained in a simple way: we define a metric for the difference in orbital elements

$$D = \left[\left(\frac{a_1 - a_2}{a_1 + a_2} \right)^2 + (\Delta h)^2 + (\Delta k)^2 + (2\Delta p)^2 + (2\Delta q)^2 + \left(\frac{\lambda_1 - \lambda_2}{10} \right)^2 \right]^{1/2},$$

where a, h, k, p, q are equinoctial elements (Broucke and Cefola, 1972) and λ_1, λ_2 are the mean longitudes reduced to the same epoch by two-body propagation. The use of metrics of this kind is discussed by several authors in different contexts (Muinonen and Bowell, 1993; Milani *et al.*, 1994; Zappalà *et al.*, 1990) and it would be worthwhile to experiment also with other ones. We then computed the distance between all the couples of asteroids in a large catalogue with $\simeq 26,321$ records (essentially all the orbits available from MPC and other sources as of Dec. 1994). We sorted the couples by distance and tested for possible identifications the closest ones. This procedure is easy to automatize and is now used also by the MPC; of course it is not good to detect the most difficult cases, those in which the two sets of elements are far apart because of poor determination over too short arcs.

Problem 1: Positive identifications. Which criteria should be used to confirm that two single opposition asteroids are the same?

If a minimum for the target function is found with $Q < 6$ (weights are such that Q is in $arcsec^2$) and each of the two arcs has at least 3 observations over at least 3 days, the identification is confirmed. As a byproduct of this test of the algorithms and the software, we have found 10 new identifications, later accepted by the MPC, and other 3 already known.

However, in some cases, there are observations to be discarded (normally labeled as such in the MPC files). If this is not done, a larger Q can result even from a good identification. If the arcs either have less than 3 observations, or are less than 3 days long, fake low minima can occur, to the point of generating crazy identifications (two asteroids, in different positions on the same plate, identified with a third one). Searching strategies in surveys must take this into account; if not, they could produce useless data.

If LANCELOT gives the same results as DC, as in most of our tests, there is no point in using a more computationally intensive procedure. However, for the reasons explained above, our tests were not too difficult, because the orbital elements solved from each of the two arcs were close.

Problem 2: Reliability of the identification procedure. Can a low minimum, therefore a positive identification, be missed?

The pure DC algorithm is very sensitive to the initial guess chosen. The convergence domain is very small, thus, the iterative procedure can be divergent, even when the low minimum exists. LANCELOT uses a more robust optimisation algorithm, however it is much more expensive.

An effective solution is to use DC, even in the pseudo-Newton formulation, but with an initial guess for the common orbit containing the two arcs which results in moderately large initial residuals. One such procedure is also used at the MPC, but the algorithm is not documented. We have developed our own algorithm to generate an initial guess for the mean longitude and semimajor axis of such a common orbit. It is obtained by computing a two-body orbit such that the mean longitude coincides with those of the two arcs at their epoch times. If the number of revolutions performed by the common orbit in the time span between the two epochs is known, there is only one semimajor axis satisfying this condition. The number of revolutions is not known a priori, but can be guessed by two-body extrapolation from each of the two short arcs; if the two extrapolations differ by one or more revolutions, the only way is to try several different initial guesses for a, obtained for each possible number of intervening revolutions.

With this initial guess we have obtained by DC all the identifications confirmed by LANCELOT; this would not be so with other simple initial guesses, such as the mean of both sets of elements. We have also tested a set of 25 identifications proposed by the MPC, and found convergence to a low minimum by DC in all cases. We intend to perform a large scale test of such an identification procedure, including much more difficult cases.

Problem 3: Can the target function Q have stationary points other than the absolute minimum?

In theory, there is no mathematical proof that Q cannot have many saddles and even many local minima. In practice, we have found one example 3024PL=93 OO7 where DC, with initial guess at the mean of the two sets of elements, finds a stationary point with $Q = 2547.$, while LANCELOT finds a minimum at $Q = 0.67$; the same minimum can also be obtained by DC with the initial guess computed as explained above. This is likely to be a quite rare case, but it is important to know that a saddle can occur. To test whether a stationary point is a minimum or a saddle we would need to have full information on the second derivatives of Q, that is to know the second derivatives of the solution of the N-body problem; this is possible, but computationally expensive. The use of LANCELOT is other alternative, since a nonlinear optimisation method does not converge towards saddles, but its computational cost is high too. Cases with multiple local minima can also occur when the observations are taken far from opposition.

Problem 4: Negative identifications. What happens when the two arcs do not belong to the same asteroid?

In some cases, a stationary point with $Q > 200$ can be found by DC. However, is it a minimum (Problem 3)? In a large fraction of the cases, LANCELOT can find a minimum with $Q > 200$. In this case, the identification should be refused (unless there is a wrong observation). The advantage of a robust nonlinear optimisation algorithm is obvious in these cases; however, the computational cost is too high to propose a brute force searching method for identifications (such as testing all possible couples with "similar" orbital elements). Moreover, in some cases neither DC nor LANCELOT can find a minimum: the iterations diverge ($e > 1$). In these cases the identification can neither be accepted nor be refused. A method to refuse identifications which is totally reliable has not been found yet.

TABLE 1. Statistics of the results

	DIFF. CORR.	LANCELOT
Positive identifications	28.9 %	28.9 %
Negative identifications	33.3 %	53.3 %
Divergent cases	33.3 %	17.8 %

In a set of 45 tests, we have found the results summarized in Table 1. It is apparent that DC should be used first, followed by an additional investigation of both negative identifications and divergent cases.

Problem 5: Which method will be suitable when the number of observations will increase by a factor of about 100?

If the number of asteroids observed only over a very short arc (1-2 days) is very large, the problem of asteroid identifications may become computationally intractable. If the orbits are good enough, because the observing strategy is such that they are all observed for a longer time, we believe it is possible to develop a fully automated algorithm for positive and negative identification. Such a method must take into account all the possible pitfalls, including multiple minima, saddles, divergent cases which result in neither a positive nor a negative identification, and initial extrapolations wrong by more than one revolution. We still cannot present such an algorithm, but we are working to clarify its mathematical foundations.

Acknowledgements

We thank the MPC, in particular B. Marsden, for their prompt answer to our queries and proposed identifications, Ph. Toint for his assistance in defining a formulation of the problem suitable for LANCELOT. This work has been supported by EEC research network CHRX-CT94-0445.

References

Broucke, R. and Cefola, P: 1972, *Celes. Mech.*, **5**, 303–310.

Conn, A.R., Gould, N.I.M. and Toint, Ph.L.: 1992, *LANCELOT: a Fortran package for large-scale nonlinear optimization*. Springer, Berlin.

Everhart, E.: 1985, in *Dynamics of Comets: their origin and evolution*, A. Carusi and G.B. Valsecchi eds., Reidel, 185–202.

Milani, A., Bowell E., Knežević Z. , Lemaitre A., Morbidelli A. and Muinonen, K.: 1994, in *Asteroids Comets Meteors 1993*, A. Milani, M. Di Martino and A. Cellino eds., Kluwer, 467–470.

Muinonen,K., and Bowell,E.i: 1993, *Icarus*, **104**, 255-279.

Zappalà,V., Cellino,A., Farinella,P. and Kneževič,Z.: 1990, *Astron. J.* **100**, 2030–2046.