Published in *Proceedings of the IASTED International Conference on Intelligent Systems and Control,* Honolulu, HI, Aug 2004.

MULTIPLE CONCURRENT RECURSIVE LEAST SQUARES IDENTIFICATION

Edward Wilson[a*], David W. Sutter[a], Robert W. Mah[b]
[a] – Intellization – 454 Barkentine Ln, Redwood Shores, CA 94065, USA
[b] – NASA Ames Research Center– MS 269-1, Moffett Field, CA 94035, USA
Ed.Wilson@intellization.com, Dave.Sutter@intellization.com, Robert.W.Mah@nasa.gov

ABSTRACT

A new algorithm, multiple concurrent recursive least squares (MCRLS) is developed for parameter estimation in a system having a set of governing equations describing its behavior that cannot be manipulated into a form allowing (direct) linear regression of the unknown parameters. In this algorithm, the single nonlinear problem is segmented into two or more separate linear problems, thereby enabling the application of existing powerful linear regression algorithms such as recursive least squares (RLS). The individual linear sub-problems contain unknown parameters other than those that are identified; said parameters are initially set at their nominal values, and are subsequently updated by the other concurrently running identification (ID) processes. With all sub-problems sharing their results following each update, the results rival those of more computationally intensive nonlinear optimization algorithms. This algorithm was developed to address, and has been tested on a spacecraft on-line mass-property ID application. Beginning with reasonably accurate initial parameter estimates, the approximation error for this spacecraft mass-property ID example is negligible as compared to errors created by other un-modeled system parameters.

KEY WORDS
identification; linear systems; spacecraft; optimization; on-line; parameter estimation

1. Introduction

1.1 Parameter estimation

In the field of parameter estimation (within the scope of System ID), there is an extensive body of theoretical results, along with explicit expressions for optimal parameter estimates, for the case where the regression function is linearly parameterized in the unknown parameters [1]. Unfortunately, when the unknown parameters do not appear linearly (for example, multiplying one another), the requisite nonlinear regression usually requires nonlinear, iterative methods.

The use of these nonlinear, iterative methods requires a significant increase in both development effort (algorithmic as well as software) and computational complexity (code size, complexity, memory and processing requirements). MCRLS addresses these issues for a broad class of problems, providing an approach to enable nonlinear parameter estimation problems to be solved with linear methods.

The simplification in development and computation enabled by the MCRLS approach is perhaps most important in that it enables implementation of *on-line* parameter estimation for applications that otherwise would not be feasible.

The MCRLS algorithm's principal objective is to enable the application of very powerful, fast, and compact parameter estimation algorithms to a class of problems that otherwise require solution with computationally intensive, and potentially fragile nonlinear approaches. It has been found to provide negligible reduction in estimation accuracy in testing on a realistic, important, and well-formed example application.

1.2 Spacecraft mass-property identification example

Due to the very small disturbance forces and torques present on spacecraft on orbit and in free space, their mass and thruster properties are important from a control and estimation standpoint. Spacecraft mass properties can be calibrated with limited accuracy during ground testing, and change further once on orbit due to: expulsion of fuel mass; reconfiguration (of antennae, solar arrays, etc.); and for servicing robotic spacecraft, potentially variable payloads. Accurate ID of mass properties has been studied extensively, but those methods have not yet been implemented and tested on-line on an actual spacecraft, possibly due to the computational complexity involved.

MCRLS was originally developed to address this spacecraft mass-property ID problem, in support of the X-38 v.201 spacecraft development at NASA Johnson Space Center, and later extended to the general case. As tested on the spacecraft application, MCRLS enables highly accurate mass-property ID and is very simple, compact and fast. It has been implemented on an experimental spacecraft, tested on-board in zero-g aircraft testing, and is presently awaiting launch for space-based validation.

1.3 Related Research

The use of linear least squares regression for the identification of unknown system parameters has been used and studied extensively, as by [1] [2]. However, the requirement that a regression equation be formed with the unknown parameters linearly represented limits its direct applicability to many important problems, including the spacecraft application mentioned here.

Ljung and several other authors have developed approaches for identification of parameters in non-linear systems using methods such as gradient-based optimization, neural networks, Wiener-Hammerstein models, the GMDH approach, etc [1]. However, these methods are significantly more computationally intensive than those for linear problems.

The remaining research items relate specifically to spacecraft mass-property ID. In that application, some of the unknown parameters multiply each other in the governing equations, resulting in a specific type of nonlinearity. These references present, for this particular application, alternative approaches to the MCRLS algorithm, demonstrating the need for such a technology.

Tanygin and Williams developed a least squares (LS) based algorithm to identify mass properties for a spinning vehicle during coasting maneuvers [3].

Bergmann developed an ID approach using a Gaussian second-order filter as presented more generally by Gelb [4] [5]. The second order filter resembles an extended Kalman filter, but has extra terms to explicitly address the second order effects. It makes approximations regarding thruster properties, vehicle gyroscopic dynamics, and is significantly more complex and computationally intensive than MCRLS.

Wilson and Rock developed an ID method based on exponentially weighted RLS using accelerometer and angular rate sensors [6]. Driven by the requirements of real-time on-line implementation as part of a reconfigurable fault-tolerant control system, the problem was reformulated to ID the accelerations resulting from thruster firings. By combining the mass and thruster properties, the regression function was linearized, thereby avoiding the present difficulty.

2. MCRLS

MCRLS estimation is presented first using a very simple example to highlight a typical issue addressed and the MCRLS solution approach, and then in general form.

2.1 Least squares parameter estimation

In LS parameter estimation, first used by Gauss in 1795, unknown parameters are identified using algorithms in which measurement data is fit to the underlying governing equations such that the identified parameter values minimize the squared error (where error is, for example, measurement data minus the ideal measurement data that would occur with zero noise and using the identified parameter values) [7].

The standard form for a linear least squares problem, referred to as "regression form," is given as

$$Ax = b + \varepsilon \quad (1)$$

or, equivalently [2]

$$Ax \cong b \quad (2)$$

where b is a vector of (perfect) measurements, ε is a vector of measurement noise, x contains the parameters to be identified, and matrix A contains known variables and system parameter values (i.e., there is no uncertainty in A). The $\cong$ in the $Ax \cong b$ representation indicates that the left and right sides of the equation would be equal if noise were not present. The LS ID solution, $\hat{x}$, minimizes the sum of the squares of the elements of the error, $A\hat{x} - b$. If the problem at hand can be put into regression form, with noise appearing only in the ε term, $\hat{x}$ can be solved directly (i.e., this is a closed-form solution, rather than an iterative optimization as might be required if the equations can not be put into this form) using one of the following equations:

Unweighted, batch algorithm: $\hat{x} = \left(A^T A\right)^{-1} A^T b$ (3)

Weighted, batch algorithm: $\hat{x} = \left(A^T W A\right)^{-1} A^T W b$ (4)

W is a diagonal weighting matrix. Either of these algorithms can be implemented recursively, and the weighting matrix, *W*, can be chosen to weight the data according to an exponentially decaying function – as is commonly done when implemented recursively [8].

In practice, many times the governing equations do not immediately fit exactly the form $Ax = b + \varepsilon$, with, for example, uncertainty being present in the A matrix and the x not being immediately linearly separated from A and b as required. Often the basic approach to LS ID is to find some governing equations (the physically based equations of motion, for example) that contain the parameter values to be identified and measurement data. Then these equations are manipulated to conform to the $Ax \cong b$ formulation, possibly requiring approximations along the way (dropping higher order terms, for example).

2.2 Simple example and solution

In some problems, such as the spacecraft example, the nonlinear terms are significant and cannot be dropped. A simple example that highlights this situation is presented by the following scalar governing equation:

$$a_1 x_1 + a_2 x_2 + a_{12} x_1 x_2 \cong b \quad (5)$$

where b is the measurement, a_1, a_2, and a_{12} are known values, that vary from measurement to measurement, and x_1 and x_2 are the unknown parameters to be identified (all are scalars). This problem cannot be put into the form of Equation 2, where the A matrix does not contain x. One approach that is feasible for a problem of this simplicity is to let

$$A = [a_1 \quad a_2 \quad a_{12}]; x = [x_1 \quad x_2 \quad x_1 x_2]^T \quad (6)$$

With the equation now in regression form, $\hat{x}$ could be solved readily. However, depending on the noise and biases present, the third element is unlikely to equal the first times the second. The approach of ignoring the third element, would be effectively throwing away information.

MCRLS would address this problem by accepting the fact that it cannot fit directly into linear regression form, and breaking it into two parts that will fit directly. $\hat{x}_1$ is to be ID'ed assuming that $\hat{x}_2$ is perfectly known, and vice versa. Equation 5 is re-written as:

$$\begin{aligned}(a_1 + a_{12}\hat{x}_2)x_1 &= b - a_2\hat{x}_2 \\ (a_2 + a_{12}\hat{x}_1)x_2 &= b - a_1\hat{x}_1\end{aligned} \tag{7}$$

or, in matrix form,

$$\begin{bmatrix} a_1 + a_{12}\hat{x}_2 & 0 \\ 0 & a_2 + a_{12}\hat{x}_1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} b - a_2\hat{x}_2 \\ b - a_1\hat{x}_1 \end{bmatrix} \tag{8}$$

The first equation in Equation 7 is set up to ID x_1, and treats x_2 as if it were a known quantity, substituting in the best estimate of x_2, $\hat{x}_2$. $\hat{x}_1$ can now be solved directly. The second equation does the converse. RLS ID processes for $\hat{x}_1$ and $\hat{x}_2$ are initialized with appropriate parameter and covariance values, and updated estimates are shared with the other ID prior to each RLS update.

One measure of the ID error introduced by this approximation can be made by running a simulation where in one case the secondary parameters (e.g., $\hat{x}_2$ in the top equation) are set to their true values, and in a second case they are set by the other ID process, and to compare the accuracies of the ID results. In many real applications, including the spacecraft example, this approximation error is insignificant as compared to system disturbances and other unknown parameters in the system. There are always some additional unmodeled effects, and as long as the MCRLS approximation error is significantly smaller than the error introduced by ignoring these effects, it should be considered a viable option.

2.3 General algorithm

This approach can be generalized by dividing the vector of unknown parameters, x, is into an arbitrarily large number of groups, each containing an arbitrarily large number of parameters. The preferred approach is to use the minimal number of groups to preserve the individual exact linear solutions to the extent possible; however, the other extreme is possible, in which each parameter has its own single-parameter RLS ID process.

$$\begin{aligned}&A_1(\hat{x}_2,\cdots,\hat{x}_n,\cdots)x_1 = b_1(\hat{x}_2,\cdots,\hat{x}_n,\cdots) \\ &A_2(\hat{x}_1,\hat{x}_3,\cdots,\hat{x}_n,\cdots)x_2 = b_2(\hat{x}_1,\hat{x}_3,\cdots,\hat{x}_n,\cdots) \\ &\vdots \\ &A_n(\hat{x}_1,\cdots,\hat{x}_n,\cdots)x_n = b_n(\hat{x}_1,\cdots,\hat{x}_n,\cdots)\end{aligned} \tag{9}$$

Equation 9 describes an arbitrary number of arbitrarily sized groups, where the subscripts indicate the number for the group of parameters, and the groups can be arbitrarily sized.

The first line in this equation is used as the regression equation in a RLS ID solution of the first vector of unknown parameters, $\hat{x}_1$. This continues for each group up to n.

As covered in the related research, the accuracy of RLS ID depends on the selection of good values for the initial parameter estimates and the estimated covariance for the error in those estimates. It is also important to weight each subsequent measurement appropriately, first according to the estimated covariance of the measurement error, and second to exponentially weight the data so that newer measurements are considered more important than older ones (this is not strictly required, but is commonly done to allow the ID to track true changes in system parameters, and is a trivial adjustment to the RLS ID algorithm to implement). When using MCRLS, careful selection of these parameters is even more important, since the IDs are dependent on each other.

If the uncertainties in the initial estimates are set too high, and the measurements are noisy, there is a chance that one of the IDs will diverge initially. The dependence of the other IDs on this makes it especially important. So it may be helpful to perform an outlier check on the measurements and prevent this from happening. If the estimate error covariances are set properly and the measurements are free of outliers, this is not a concern.

2.4 Additional benefits

In addition to the principal benefit of enabling the approximate solution of a nonlinear problem with linear methods, several other benefits are present:

It is not always clear as to which parameters in a set of equations should be treated as known, and which should be estimated. One significant benefit of MCRLS is that parameters can be added to or removed from the list of unknowns without needing to reformulate the entire problem. For example, in the spacecraft application, the related research examples all treat thruster parameters as perfectly known, and the algorithms cannot be updated to accommodate a change in this assumption. With MCRLS, the additional estimation of thruster strengths, directions, and locations can be added incrementally, as required.

As a practical consideration, it is common for the unknown parameters in a system to have widely varying degrees of uncertainty. For example, in the spacecraft example, due to the accuracy of ground-test equipment, center of mass is known with better relative accuracy than the diagonal terms in the inertia matrix, which are known better than the off-diagonal terms. Thruster locations are known very well, thruster directions are less certain, and for some systems, the thruster magnitude is even less certain. The capability of MCRLS ID to accommodate these problem characteristics is an important one.

In many applications, certain measurements are more directly related to some parameters than others, and MCRLS facilitates the accommodation of this. For example, in the spacecraft example, when thrusters are fired to produce a pure torque, the resulting rotational motion as measured by the gyros is independent of the center of mass. Any update to the center of mass ID using this data is based on noise rather than physics and should be avoided. With the ID already segmented, MCRLS facilitates implementation of such physically based decisions, although the design of this logic may be difficult.

Another benefit that is especially valuable for systems requiring on-line implementation on embedded processors where code size may be limited is that the code is very small. The RLS implementation consists of a small number of lines of code, due to the simplicity of the RLS algorithm, and it can be shared by the individual IDs, leading to a significant savings in code size, and the time and cost for development, testing, and sustaining engineering.

2.5 Extensions

Although the use of RLS to perform the segmented individual IDs has been used, and is clearly well suited to the algorithm, other algorithms would be able to fit directly into this framework.

Although well-suited for real-time, on-line implementation, MCRLS ID could be applied off-line. Repeated cycling of the data in an off-line application has the effect of reducing the influence of the initial estimates and reducing the interdependence of the individual IDs.

There may be cases where one or more of the unknown parameters changes in a known manner. For example, in the spacecraft application the total (vehicle + fuel + payload) mass properties are unknown, but the fuel location and mass are known. The remaining fuel mass may be calculated accurately as a function of time using a procedure known as burn-time integration (BTI). In this case, the MCRLS ID is designed to ID the difference between true and nominal total mass properties, while the nominal mass properties are updated using BTI. Even if the BTI is not perfectly accurate the ID will partially account for that. So the known part of the change is incorporated exactly and immediately, and the unknown deviation of the parameters from nominal continues to be ID'ed. Although this is a very useful augmentation to the MCRLS ID, it can be implemented independently.

2.6 Convergence

Although no convergence proof has yet been developed, empirical evidence from testing suggests that with sufficiently accurate initial estimates, ID accuracy will be comparable to that obtainable if true rather than estimated parameters were used in each ID process (e.g., if the x_1 RLS ID process used x_2 rather than $\hat{x}_2$).

2.7 Optimality

Whereas the problem formulation and solutions for the baseline LS problem (Equations 1-4) are provably optimal, the MCRLS results are clearly not optimal in the least squares sense. Even with careful estimate-error-covariance-matrix initialization and measurement weighting, MCRLS ID can only approach the mathematical optimality of a Kalman Filter or the results of a nonlinear optimization.

However, in virtually all real applications the strict requirements of the LS formulation are not met, and any algorithmic sub-optimality is only important as compared to other sub-optimal effects (un-modeled disturbances or dynamics, unknown variations in "known" parameters, dropped higher order terms, sensor biases, etc.).

At some point, additional algorithmic complexity is not warranted due to the deviations from the idealized problem statement. It is with this rationale that uncertainty in the A matrix is conventionally accepted and ignored for system ID (e.g., use of past measurements, or uncertain model parameters may appear) [1].

2.8 Viability

The viability of MCRLS is clearly not universal. The following two approaches test whether the MCRLS is viable for a given application. Test 1 determines if the MCRLS approximation error is negligible in the presence of other un-modeled effects. Test 2 determines how much better an optimal nonlinear ID would perform. The development of theoretical (as opposed to these empirical examples) tests may be possible.

Test 1: A simulation of the system is developed and used to generate a set of data to be used in the ID. In the baseline case, the various RLS IDs use estimated values for the unknown parameters that are treated as known (i.e., normal MCRLS). Then in the second case, each individual RLS ID instead uses the true values as used in the simulation for these parameters. This eliminates any potential coupling and approximation error – since the individual RLS IDs are ideal in their linear forms. The test is then to see how much greater the ID errors are for the first case as compared to the second. For many applications, including the spacecraft example to follow, the ID error introduced by other un-modeled effects (in the simulation, but not in the ID) will be greater than the MCRLS approximation error.

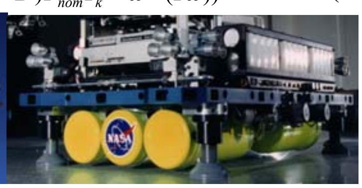
Test 2: The nonlinear governing equations are run with parameters set to the MCRLS-ID'ed values, and system inputs to calculate the expected measurements. The sum-of-squares deviation of these measurements from the actual (or simulated) measurements is the error metric (prediction error). With the MCRLS-ID'ed parameters as a starting point, off-line, a gradient-based search (or other nonlinear ID that is not concerned with run-time efficiency) is carried out to find the optimal parameters, quantifying the potential improvement in the cost function.

3. Spacecraft Example

The MCRLS algorithm was developed to address a spacecraft system ID problem that is presented in detail in [9][10], and summarized here.

Accurate knowledge of spacecraft center of mass and inertia parameters is important for advanced control, estimation, and fault detection and isolation (FDI) applications that use these values. In addition to the difficulty faced in measuring these accurately prior to launch, these values may change following launch due to: fuel depletion; structural reconfiguration; and variable payload. The goal was to ID center of mass (CM) and inertia by monitoring the vehicle motions (as measured by gyros) in response to thruster firings. The challenge (as similarly encountered by the other related research examples) is that the governing rotational equation of motion, given in Equation 10, contains CM and inertia in a way that they cannot be linearly separated.

$$\dot{\omega} = I^{-1}((L \times D)F_{nom}T_k - \omega \times (I\omega)) \qquad (10)$$



The issue present even in this greatly simplified representation of the problem is that the inertia parameters (or in this case, inverse-inertia, I^{-1}) multiply the thruster location parameters, L, which are direct functions of the center of mass. The problem is made significantly more complex if it is also desired to ID thruster properties, total mass, and other system properties. Fortunately, the MCRLS algorithm scales to accommodate each additional set of parameters without impacting the overall complexity. This has been applied successfully in simulation to several spacecraft shown here, tested in laboratory hardware and in 0-g aircraft simulators, and is awaiting launch to the ISS for space-based validation.

4. Conclusions

An efficient algorithm, MCRLS, has been presented for parameter estimation in a system having a set of governing equations describing its behavior that cannot be put into regression form with the unknown parameters linearly represented. In this algorithm, the vector of unknown parameters is segmented into two or more groups where each individual group of parameters may be isolated linearly by manipulation of said equations.

Multiple concurrent and independent RLS identifications of each said group run, treating other unknown parameters appearing in their regression equation as if they were known perfectly, and with their values provided by RLS estimation from the other groups.

The primary advantage is that it enables the use of fast, compact, efficient linear algorithms to solve problems that would otherwise require nonlinear solution approaches. MCRLS is presented with application to ID of mass and thruster properties for a thruster-controlled spacecraft. Other benefits, extensions, optimality considerations, and methods for testing its applicability are discussed.

Acknowledgements

Research funded by NASA Headquarters, HQ AA, PWC 349-00, and by the NASA CICT/Intelligent Systems Program 302-10-10.

References

[1] L. Ljung, *System identification: theory for the user* (Upper Saddle River, NJ: Prentice Hall, 1999).

[2] C. L. Lawson and R. J. Hanson, *Solving least squares problems, Series in automatic computation* (Englewood Cliffs, NJ: Prentice Hall, 1974).

[3] S. Tanygin and T. Williams, Mass property estimation using coasting maneuvers, *Journal of Guidance, Control, and Dynamics, 20*(Jul-Aug) 1997, 625-632.

[4] A. Gelb, et al, *Applied optimal estimation* (Cambridge, MA: MIT Press, 1974).

[5] E.V. Bergmann, B.K. Walker, and D.R. Levy, Mass property estimation for control of asymmetrical satellites, *Journal of Guidance, Control, and Dynamics, 10*(Sept.-Oct.) 1987, 483-491.

[6] E. Wilson and S.M. Rock, Reconfigurable control of a free-flying space robot using neural networks, *Proc. 1995 American Control Conference*, Seattle WA, Jun 1995.

[7] K.F. Gauss, *Theoria motus corporum celestium, English translation: Theory of the motion of the heavenly bodies* (Hamburg: 1809; New York: Dover, 2004).

[8] G. Franklin, J.D. Powell, and M.L. Workman, *Digital control of dynamic systems, second edition* (Menlo Park, CA: Addison Wesley, 1990).

[9] E. Wilson, C. Lages, and R.W. Mah, On-line, gyro-based, mass-property identification for thruster-controlled spacecraft, *Proc. 2002 IEEE International Midwest Symposium on Circuits and Systems*, Tulsa, OK, Aug 2002.

[10] E. Wilson, D.W. Sutter, and R.W. Mah, MCRLS for on-line spacecraft system identification, *Proc. IASTED International Conference on Intelligent Systems and Control*, Honolulu, HI, Aug 2004.