# LP-BASED STRATEGY FOR MODELING AND OPTIMIZATION USING INTERVAL METHODS

Youdong Lin and Mark A. Stadtherr*
University of Notre Dame
Notre Dame, IN 46556

*Abstract*

The interval-Newton approach provides the power to solve nonlinear equation solving and global optimization problems with complete mathematical and computational certainty. The primary drawback to this approach is that computation time requirements may become quite high. In this paper, a strategy for using linear programming (LP) techniques to improve computational efficiency is considered. In particular, an LP strategy is used to determine exact (within round out) bounds on the solution set of the linear interval equation system that must be solved in the context of the interval-Newton method. The strategy is tested using global optimization problems arising in parameter estimation based on the error-in-variables approach.

*Keywords*

Interval analysis; Global optimization; Nonlinear equation system solving; Parameter estimation

## Introduction

At the core of many process operations problems is the need to solve a nonlinear process model, or to optimize a nonlinear function subject to constraints. In either case, there may arise issues with the reliability of the problem-solving method used. For example, if there are multiple solutions to the model, have all been located? If there are multiple local optima, has the global solution been found? Interval mathematics can provide the engineer with the tools needed to resolve these issues with mathematical and computational certainty, thus providing a degree of problem-solving reliability not available when using standard methods. In recent years, it has been shown that strategies based on an interval-Newton approach can be used to reliably solve a wide variety of nonlinear equation solving and optimization problems in chemical process engineering (e.g., Gau and Stadtherr, 2000), providing a mathematical and computational guarantee either that all solutions have been located in an equation solving problem or that the global optimum has been found in an optimization problem.

The primary drawback to this approach is the potentially high computational cost. One way to improve the efficiency of the interval-Newton approach is to more tightly bound the solution set of the linear interval equation system that is at the core of this approach. In this paper, we consider a strategy for using linear programming (LP) techniques to determine exact (within round out) bounds on the solution set of this linear interval system. By providing tight interval bounds on the solution set, the goal is to more quickly identify intervals that contain a unique solution, and intervals that contain no solution, thus leading to an overall improvement in computational efficiency.

## Interval Analysis

Several good introductions to interval computations are available (e.g., Kearfott, 1996). Of particular interest here is the interval-Newton method. Given an $n \times n$ nonlinear equation system $\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{0}$ with a finite number of real roots in some initial interval, this technique provides

* Author to whom all correspondence should be addressed. Fax: (574) 631-8366; E-mail: markst@nd.edu

the capability to enclose all the roots of the system that lie within the given initial interval. For the unconstrained optimization problem $\min_{\boldsymbol{x}} \phi(\boldsymbol{x})$, a common approach is to seek stationary points, that is, to solve the nonlinear system $\boldsymbol{f}(\boldsymbol{x}) \equiv \nabla\phi(\boldsymbol{x}) = \boldsymbol{0}$. The global optimum will be one of roots of this nonlinear equation system, but there may be other roots as well, representing local optima and saddle points. To identify the global optimum, it is critical that none of the roots be missed, and such a guarantee can be provided by the interval-Newton approach. If it is a constrained optimization problem, then the interval-Newton method can be applied to solve the KKT or Fritz-John conditions.

Given some initial interval $\boldsymbol{X}^{(0)}$, the interval-Newton solution algorithm is applied to a sequence of subintervals. For a subinterval $\boldsymbol{X}^{(k)}$ in the sequence, the first step is the function range test. An interval extension $\boldsymbol{F}(\boldsymbol{X}^{(k)})$ of the function $\boldsymbol{f}(\boldsymbol{x})$ is calculated, which provides upper and lower bounds on the range of values of $\boldsymbol{f}(\boldsymbol{x})$ in $\boldsymbol{X}^{(k)}$. If there is any component of the interval extension $\boldsymbol{F}(\boldsymbol{X}^{(k)})$ that does not include zero, then the interval can be discarded, since no solution of $\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{0}$ can exist in this interval. The next subinterval in the sequence may then be considered. Otherwise, testing of $\boldsymbol{X}^{(k)}$ continues.

For a global minimization problem, the next step is the objective range test. The interval extension $\Phi(\boldsymbol{X}^{(k)})$, containing the range of $\phi(\boldsymbol{x})$ over $\boldsymbol{X}^{(k)}$ is computed. If the lower bound of $\Phi(\boldsymbol{X}^{(k)})$ is greater than a known upper bound on the global minimum, then $\boldsymbol{X}^{(k)}$ can be discarded since it cannot contain the global minimum and need not be further tested. In cases that all the stationary points are desired rather than just the global minimum, this test step can be turned off.

The next step is the interval-Newton test. The linear interval equation system

$$F'(\boldsymbol{X}^{(k)})(\boldsymbol{N}^{(k)} - \boldsymbol{x}^{(k)}) = -\boldsymbol{f}(\boldsymbol{x}^{(k)}) \qquad (1)$$

is solved for a new interval $\boldsymbol{N}^{(k)}$, where $F'(\boldsymbol{X}^{(k)})$ is an interval extension of the Jacobian of $\boldsymbol{f}(\boldsymbol{x})$, and $\boldsymbol{x}^{(k)}$ is an arbitrary point in $\boldsymbol{X}^{(k)}$. It has been shown that any root contained in $\boldsymbol{X}^{(k)}$ is also contained in the image $\boldsymbol{N}^{(k)}$. This implies that when $\boldsymbol{X}^{(k)} \cap \boldsymbol{N}^{(k)}$ is empty, then no root exists in $\boldsymbol{X}^{(k)}$, and also suggests the iteration scheme $\boldsymbol{X}^{(k+1)} = \boldsymbol{X}^{(k)} \cap \boldsymbol{N}^{(k)}$. In addition, if $\boldsymbol{N}^{(k)} \subset \boldsymbol{X}^{(k)}$, there is a *unique* root contained in $\boldsymbol{X}^{(k)}$ and thus in $\boldsymbol{N}^{(k)}$. Thus, after computation of $\boldsymbol{N}^{(k)}$, there are three possibilities: 1. $\boldsymbol{X}^{(k)} \cap \boldsymbol{N}^{(k)} = \varnothing$, meaning there is no root in the current interval $\boldsymbol{X}^{(k)}$ and it can be discarded; 2. $\boldsymbol{N}^{(k)} \subset \boldsymbol{X}^{(k)}$, meaning that there is exactly one root in the current interval $\boldsymbol{X}^{(k)}$. 3. Neither of the above, meaning that no conclusion can be drawn. In the last case, if $\boldsymbol{X}^{(k)} \cap \boldsymbol{N}^{(k)}$ is sufficiently smaller than $\boldsymbol{X}^{(k)}$, then the interval-Newton test can be reapplied to the resulting intersection. Otherwise, the intersection is bisected, and the resulting two subintervals are added to the sequence of subintervals to be tested. This approach is referred to as an interval-Newton/generalized-bisection (IN/GB) method. At termination, when the subintervals in the sequence have all been tested, either all the real roots of $\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{0}$ have been tightly enclosed or it is determined that no roots exist.

Clearly, the solution of the linear interval system given by Eq. (1) is essential to this approach. To see the issues involved in solving such a system, consider the general linear interval system $A\boldsymbol{x} = \boldsymbol{B}$, where the matrix $A$ and the right hand side vector $\boldsymbol{B}$ are interval-valued. The solution set $S$ of this system is defined by $S = \{\boldsymbol{x} \mid \tilde{A}\boldsymbol{x} = \boldsymbol{b}, \tilde{A} \in A, \boldsymbol{b} \in \boldsymbol{B}\}$. However, in general this set is not an interval and may have a very complex polygonal geometry. Thus to "solve" the linear interval system, one instead seeks an interval $\boldsymbol{X}$ containing $S$. Computing the interval hull (the tightest interval containing $S$) is NP-hard (Rohn and Kreinovich, 1995), but there are several methods for determining an interval $\boldsymbol{X}$ that contains but overestimates $S$. Various interval-Newton methods differ in how they solve Eq. (1) for $\boldsymbol{N}^{(k)}$ and thus in the tightness with which the solution set is enclosed. By obtaining bounds that are as tight as possible, the overall performance of the interval-Newton approach can be improved since with a smaller $\boldsymbol{N}^{(k)}$ it is more likely that either $\boldsymbol{X}^{(k)} \cap \boldsymbol{N}^{(k)} = \varnothing$ or $\boldsymbol{N}^{(k)} \subset \boldsymbol{X}^{(k)}$ will be satisfied. Thus, intervals that contain no solution or that contain a unique solution may be more quickly identified, and the number of bisections needed may be reduced.

Frequently, the $\boldsymbol{N}^{(k)}$ is computed component-wise using an interval Gauss-Seidel approach, preconditioned with an inverse-midpoint matrix. Though the inverse-midpoint preconditioner is a good general-purpose preconditioner, it is not always the most effective approach (Kearfott, 1996). Recently, a hybrid preconditioning scheme (HP/RP) has been proposed to obtain more efficient computational performance (Gau and Stadtherr, 2002). It combines a simple pivoting preconditioner with the standard inverse-midpoint scheme. However, it still cannot yield the tightest enclosure of the solution set, which, as noted above, is in general an NP-hard problem. In next section, a linear programming strategy will be applied to solve the linear interval system, Eq. (1), arising in the context of interval-Newton methods. Using this approach, exact component-wise bounds on the solution set can be calculated, while avoiding exponential time complexity. A similar LP strategy has also been proposed for use in the context of certain types of constraint satisfaction problems (e.g., Jaulin et al., 2001).

## LP Strategy for Linear Interval System

Consider again the linear interval system $A\boldsymbol{x} = \boldsymbol{B}$. Oettli & Prager (1964) show that the solution set $S$ is determined by the constraints

$$\left|\hat{A}\boldsymbol{x}-\hat{\boldsymbol{B}}\right| \le \Delta A|\boldsymbol{x}|+\Delta \boldsymbol{B} \qquad (2)$$

where $\hat{A}$ is the component-wise midpoint matrix of the interval matrix $A$, $\Delta A$ is the component-wise half-width (radius) matrix of $A$, and similarly $\hat{\boldsymbol{B}}$ and $\Delta\boldsymbol{B}$ are the midpoint and radius of $\boldsymbol{B}$. Equation (2) is not directly useful for computing bounds on the solution set because of the absolute value operation on the right-hand side. In general, the solution may lie in all $2^n$ orthants for an $n$-dimensional problem. In each orthant, each component of $\boldsymbol{x}$ keeps a constant sign, and thus the absolute value can be dropped. For a given orthant, define the vector $\boldsymbol{\alpha}$ by

$$\alpha_j = \begin{cases} 1 & x_j \ge 0 \\ -1 & x_j \le 0 \end{cases} \quad j = 1,2,\ldots,n\,. \qquad (3)$$

Now let $D_\alpha$ be a diagonal matrix whose entries are the components of $\boldsymbol{\alpha}$, thus $|\boldsymbol{x}| = D_\alpha \boldsymbol{x}$ and $\boldsymbol{x} = D_\alpha|\boldsymbol{x}|$. Equation (2) becomes:

$$\left|\hat{A}\boldsymbol{x}-\hat{\boldsymbol{B}}\right| \le \Delta A D_\alpha \boldsymbol{x}+\Delta \boldsymbol{B} \qquad (4)$$

That is:

$$\begin{pmatrix} \hat{A}-\Delta A D_\alpha \\ -\hat{A}-\Delta A D_\alpha \end{pmatrix}\boldsymbol{x} \le \begin{pmatrix} \boldsymbol{B}^U \\ -\boldsymbol{B}^L \end{pmatrix} \qquad (5)$$

where the superscripts $U$ and $L$ indicate upper and lower interval bounds. To then determine the tightest interval enclosing the solution set, the set of optimization problems max $x_j$ and min $x_j$, $j$=1,2,…,$n$, each with the 2$n$ linear inequality constraints given by Eq. (5) can be solved for each orthant using LP techniques. In general, there are $2^n$ orthants and so the solution time complexity will be exponential.

In the context of the interval-Newton method, however, the exponential time complexity can be avoided by manipulating the selection of the real point in Eq. (1). Here $\boldsymbol{x}^{(k)}$ is an arbitrary point in $\boldsymbol{X}^{(k)}$ typically taken to be the midpoint. If $\boldsymbol{x}^{(k)}$ is chosen to be a corner of $\boldsymbol{X}^{(k)}$ instead, then, since it is only the intersection of $\boldsymbol{X}^{(k)}$ and $\boldsymbol{N}^{(k)}$ that is of interest, we can seek the solution ($\boldsymbol{N}^{(k)} - \boldsymbol{x}^{(k)}$) of Eq. (1) in just one orthant. Therefore, to solve the linear interval system, only 2$n$ LP subproblems with 2$n$ constraints need to be solved. Furthermore, the LP subproblems have properties that can be exploited. First, all the 2$n$ subproblems share the same constraints; that is, the same feasible region. Thus, an initial feasible basis for the LP subproblems needs to be found only once. Second, the objective function of each subproblem consists of just one variable. This makes the problem much simpler since it is not necessary, as it is in the general case, to calculate the gain in objective value when choosing variables to enter and exit the basis.

LISS_LP (Linear Interval System Solver by Linear Programming) is a procedure that we have developed based on the above scheme. This procedure can be used to replace the inverse-midpoint preconditioned Gauss-Seidel method for solving Eq. (1). It can be combined with the row basis pivoting preconditioner and optimal real point scheme of Gau & Stadtherr (2002) to achieve best performance. The pivoting preconditioner can also help LISS_LP to select one of the corners of $\boldsymbol{X}^{(k)}$, the choice of which may have a significant impact on the overall performance of LISS_LP. Since the LP subproblems are solved using floating point arithmetic in the current implementation of LISS_LP, which may cause rounding error concerns, a practical error bound estimator on the solution of the linear system was adopted to ensure the reliability of the solution.

**Results and Discussion**

In this section, we present the results of numerical experiments to test the effectiveness of LISS_LP in replacing the inverse-midpoint preconditioned interval Gauss-Seidel approach. To do this we compare LISS_LP to the HP/RP method of Gau and Stadtherr (2002). Three error-in-variables (EIV) parameter estimation test problems were used. The problem sources and other details can be found in Gau and Stadtherr (2000). Tests were on a Sun Blade 1000 model 1600 workstation.

*Problem 1: Parameter Estimation in VLE Modeling*

This is a parameter estimation problem using the EIV approach to estimate parameters in the van Laar equation for activity coefficients. These two parameters are estimated from binary VLE data for the binary system of methanol and 1,2-dichloroethane. The experimental data consist of five experimental data points for four measured state variables, namely pressure, temperature, and liquid- and vapor-phase mole fraction of methanol. The problem is formulated as an unconstrained global optimization problem with 12 variables, as explained by Gau and Stadtherr (2000).

This global optimization problem was solved successfully, with computational performance results shown in Table 1, where the number of interval-Newton (I-N) tests performed, and the CPU time in seconds are given. When LISS_LP is applied, the number of I-N tests is substantially reduced, indicating the effectiveness of LISS_LP in reducing the number of intervals that must be tested in comparison to when the inverse-midpoint preconditioned Gauss-Seidel approach is used. Essentially, by reducing the size of $\boldsymbol{N}^{(k)}$, LISS_LP is able to more quickly identify and discard intervals that contain no solution.

*Table 1. Performance on Problem 1*

| | I-N tests | CPU time (s) |
|---|---|---|
| HP/RP | 303,589 | 664.4 |
| LISS_LP | 156,182 | 496.7 |

*Problem 2: Parameter Estimation in Reactor Modeling*

This is a parameter estimation problem using the EIV approach to estimate kinetic parameters for an irreversible, first-order reaction A $\rightarrow$ B, using data from an adiabatic CSTR, with 10 data points for 5 measured state variables. The problem is formulated as an unconstrained global optimization problem with 22 variables, as explained by Gau and Stadtherr (2000).

This global optimization problem was successfully solved, with the results for computational performance shown in Table 2. Again there was a large reduction in the number of I-N tests required. However, in this problem and in Problem 1, the percent reduction in overall CPU time is less than the percent reduction in I-N tests. This occurs due to the overhead in solving the LP subproblems.

*Table 2. Performance on Problem 2*

| | I-N tests | CPU time (s) |
|---|---|---|
| HP/RP | 9,505 | 24.0 |
| LISS_LP | 2,004 | 19.2 |

*Problem 3: Parameter Estimation in Batch Reaction Kinetics Modeling*

This problem is a parameter estimation problem using the EIV approach to estimate the kinetic parameters for an isothermal batch reactor. There are two sets of data for this problem. The first data set consists of 15 data points and the optimization problem has 32 independent variables. The second data set consists of 7 data points and 16 independent variables. For the second data set, some of the data points are suspicious, and, to investigate, Gau and Stadtherr (2000) turned off the objective range test to obtain all of the stationary points, not just the global optimum. We will do likewise here.

The optimization problems were solved successfully and the performance results are shown in Table 3 and Table 4. In first data set, although the number of interval-Newton tests has been reduced by a factor of about four when using LISS_LP, the CPU time is almost the same for both methods. In this and the other problems, the relatively high overhead in solving the LP subproblems is due to a large extent to the fact that dense linear algebra is used currently in LISS_LP, and it is quite inefficient, especially as the problem size becomes larger. In the second data set, two stationary points are found, including the global optimum. Again there is a significant reduction in I-N tests, but the CPU time is not proportionately reduced due to the overhead in solving LP subproblems.

*Table 3. Performance on Problem 3: first data set*

| | I-N tests | CPU time (s) |
|---|---|---|
| HP/RP | 144,833 | 976.2 |
| LISS_LP | 30,560 | 992.6 |

*Table 4. Performance on Problem 3: second data set*

| | I-N tests | CPU time (s) | No. Stationary points |
|---|---|---|---|
| HP/RP | 69,421 | 91.8 | 2 |
| LISS_LP | 18,715 | 68.0 | 2 |

**Concluding Remarks**

We have described here a LP-based method to solve the linear interval system arising in the context of the interval-Newton approach for nonlinear equation solving and global optimization. The method can obtain tighter bounds on the solution set than the standard inverse-midpoint preconditioned Gauss-Seidel method, and thus leads to a large reduction in the number of subintervals that must be tested during the interval-Newton procedure. However, as seen especially on the largest problem considered here (first data set of Problem 3), the overhead required to solve the LP subproblems may lead to relatively smaller, or even no, improvements in overall computation time. Since this is due to a large extent to the use of inefficient dense linear algebra routines in the current implementation of LISS_LP, it is clear that an implementation that exploits problem sparsity is needed.

**Acknowledgments**

This work has been supported in part by the donors of The Petroleum Research Fund, administered by the ACS, under Grant 35979-AC9, and by the Indiana 21st Century Research & Technology Fund.

**References**

Gau, C.-Y. and Stadtherr, M. A. (2000). Reliable Nonlinear Parameter Estimation Using Interval Analysis: Error-in-Variable Approach. *Comput. Chem. Eng., 24*, 631-638.

Gau, C.-Y. and Stadtherr, M. A. (2002). New Interval Methodologies for Reliable Chemical Process Modeling. *Comput. Chem. Eng., 26*, 827-840.

Jaulin, L., Kieffer, M., Didrit, O., and Walter, É. (2001). *Applied Interval Analysis*. Springer-Verlag, London.

Kearfott, R. B. (1996). *Rigorous Global Search: Continuous Problems*. Kluwer Academic Publishers, Dordrecht.

Otelli, W. and Prager, W. (1964). Compatibility of Approximate Solution of Linear Equation with Given Error Bounds for Coefficients and Right-Hand Sides. *Numrische Mathematik, 6*, 405-408.

Rohn, J. and Kreinovich, V. (1995). Computing exact componentwise bounds on solution of linear systems with interval data is NP-hard. *SIAM J. Matrix. Anal., 16*, 415-420.