WEIGHTED LEAST ABSOLUTE VALUE STATE ESTIMATION USING INTERIOR POINT METHODS

H. Singh  
Student Member  
Department of Electrical and Computer Engineering  
The University of Wisconsin, Madison WI 53706

F.L. Alvarado  
Fellow  
Department of Electrical and Computer Engineering  
The University of Wisconsin, Madison WI 53706

Abstract

This paper addresses the application of interior point methods to the Weighted Least Absolute Value state estimation problem. Interior point methods are applied to the primal and dual formulations of the problem. The dual formulation involves solving least squares problems identical in structure to those used in conventional Weighted Least Squares state estimation. The dual formulation also provides an initial feasible interior point without any extra effort. Computational issues that are critical to the success of interior point methods are addressed and test results on standard systems are provided. Keywords: WLAV state estimation, interior point methods, least squares solutions, sparsity.

1 Introduction

State estimation was introduced in power systems over two decades ago by Schweppe. Over this period, the predominant approach for solving the problem has been the Weighted Least Squares (WLS) method [29]. Much research in these methods has concentrated on finding better methods for solving large least squares problems, either by the ordinary normal equations method, the augmented matrix method, or the orthogonal factorization method. A comparison of some of these approaches can be found in [17]. Specialized blocking methods were applied to the augmented formulation in [3, 25] and enhancements to orthogonal factorization using Givens rotations were introduced in [31]. The other area of focus in these methods has been in finding computationally viable and effective ways of detecting and eliminating bad data [24]. Another category of methods, which is of interest to us here, is based on non-quadratic criteria. One such criterion is the Weighted Least Absolute Value (WLAV) criterion which is based on the $\ell_1$ norm and has useful bad data rejection properties. WLAV estimators are able to reject bad measurements as long as these are not leverage points [2]. The bad data rejection properties of WLAV estimators are common to all formulations and are not explored in this paper. The WLAV problem was first solved by using a standard linear programming formulation by Irving et al [18]. Kotiuga and Vidyasagar were the first to use specialized methods for $\ell_1$ norms [21]. Two such methods [5, 6] were used by Abur and Çelik [1] and Clements et al [8]. All these methods are based on the simplex method of linear programming in which the solution at every iteration lies at an extreme point of the feasible region. Each iteration of the simplex method involves the solution of two linear systems and a factor update to reflect the change of one column of the basis. This can be done in a fast and stable manner. However, the number of iterations in the simplex method and the effort per iteration grow linearly with system size, resulting in the total effort growth as the square of system size. Simplex implementations of WLAV methods are therefore slower than WLS methods. Another perceived disadvantage of WLAV methods is the need for either an efficient standard simplex code or the implementation of a specialized variant of such a code. In this paper, we consider methods that change the structure of the WLAV solution into that of a conventional WLS solution.

The work by Karmarkar [19] generated a lot of interest as a competitive alternative to the simplex method. There is a plethora of interior point methods available in the literature. Two categories of methods that work well are affine scaling variants [4, 30] and barrier function methods [15, 16]. Both types of methods have been applied to power system problems. Affine scaling methods were used in [26, 28] and barrier function methods were used in [9, 10]. Most of the computational effort in interior point methods involves solving a sequence of least squares problems. The type of method chosen influences the total number of iterations.

In this paper, we use the barrier function method [15] to solve the primal formulation, and an affine scaling method [30] to solve the dual formulation of the WLAV state estimation problem. (The primal and dual formulations were compared using the simplex method in [12].) The problem is set up as an optimization prob-


PRINTED IN USA

0885-8950/94/$04.00 © 1993 IEEE
lem with a linear objective function subject to a set of non-linear constraints. We use the sequential linear programming approach based on a successive linearization of the constraints instead of a direct application of an interior point method to the non-linear problem as was done by Clements, Davis and Frey in [9]. This is computationally extremely useful as the early linearizations do not have to be solved till optimality. The results presented in this paper exploit this feature. The implementation of the method, particularly for the dual problem, is extremely simple if a routine for solving least squares problems is available. The methods presented in this paper exploit the problem structure to obtain an initial interior point which not only results in fewer iterations but is also desirable from a sparsity viewpoint.

An augmented formulation is used to solve the least squares problem for computational as well as numerical reasons. In addition, the augmented formulation permits the easy treatment of unrestricted or "free" variables (a concern in the primal problem formulation).

2 Problem Formulation

The WLAV state estimation problem is usually set up as follows:

\[ \min \sum_{i=1}^{m} w_i |r_i| \quad (1) \]

subject to

\[ z = h(x) + r \quad (2) \]

where:

- \( z \) is the \( m \) dimensional measurement vector
- \( x \) is the \( n \) dimensional state vector (\( n < m \))
- \( r \) is the \( m \) dimensional vector of measurement errors
- \( w \) is the \( m \) dimensional vector of weights
- \( h \) is the \( m \) dimensional vector of non-linear functions

which relate measurements to the system state

The weights are based on the measurement standard deviations. A higher weight signifies a more reliable measurement.

Primal Formulation

Substituting \( \eta - \rho \) for the vector \( r \) where \( \eta \) and \( \rho \) represent the positive and negative parts of \( r \) respectively and linearizing \( h(x) \) about the current \( x^k \) results in:

\[ \min w^T(\eta + \rho) \quad (3) \]

subject to:

\[ \begin{bmatrix} H & I & -I \end{bmatrix} \begin{bmatrix} \Delta x \\ \eta \\ \rho \end{bmatrix} = (\Delta z) \quad (4) \]

\( \eta, \rho \geq 0 \quad (5) \)

These equations represent the linearized constraints at a given operating point, \( x^k \). The problem is solved for \( \Delta x \). The value of \( x \) is then updated, \( x^{k+1} \leftarrow x^k + \Delta x \). The constraints are linearized at the updated operating point, and the procedure is repeated until convergence. We focus our attention on the solution of the linearized problem defined in equations (3-5).

The logarithmic barrier function was first used by Frisch [14] and appears extensively in the book by Fiacco and McCormick [13]. The logarithmic barrier function method consists of applying Newton's method to the following family of problems parameterized by the barrier parameter \( \mu > 0 \).

\[ \min w^T(\eta + \rho) - \mu \sum_{i=1}^{m}(\ln \eta_i + \ln \rho_i) \quad (6) \]

subject to the constraints in equation (4). The solutions to the above family of problems for a sequence of decreasing \( \mu \) approach the solution of the original problem. Various strategies for updating \( \mu \) can be found in the literature and have been used to classify the methods into long step and short step methods [16]. Intermediate problems in the sequence need not be solved to optimality. In our experiments \( \mu \) was initialized to unity and decreased by a factor of ten in every iteration. The variables \( \Delta x \) are unrestricted and do not figure in the barrier function. The Newton search direction [15] is given by the Karush-Kuhn-Tucker first order optimality conditions for the quadratic approximation of (6) subject to retaining feasibility. These can be written as

\[ \begin{bmatrix} 0 & 0 & 0 & H^T \\ 0 & \mu D_{\eta}^{-2} & 0 & I \\ 0 & 0 & \mu D_{\rho}^{-2} & -I \\ H & I & -I & 0 \end{bmatrix} \begin{bmatrix} -p_x \\ -p_\eta \\ -p_{\rho} \\ \pi \end{bmatrix} = \begin{bmatrix} 0 \\ w - \mu D_{\eta}^{-1}e \\ w - \mu D_{\rho}^{-1}e \\ 0 \end{bmatrix} \quad (7) \]

This can be re-written after defining

\[ p_x = -(1/\mu) \ r_x \quad (8) \]

\[ p_\eta = -(D_{\eta}/\mu) \ r_\eta \quad (9) \]

\[ p_{\rho} = -(D_{\rho}/\mu) \ r_{\rho} \quad (10) \]

as

\[ \begin{bmatrix} 0 & 0 & H^T \\ 0 & I & 0 & D_{\eta} \\ 0 & 0 & I & -D_{\rho} \\ H & D_{\eta} & -D_{\rho} & 0 \end{bmatrix} \begin{bmatrix} r_x \\ r_\eta \\ r_{\rho} \\ \pi \end{bmatrix} = \begin{bmatrix} 0 \\ D_{\eta}w - \mu e \\ D_{\rho}w - \mu e \\ 0 \end{bmatrix} \quad (11) \]

where \( \pi \) represents dual variables and \( r_x, r_\eta \) and \( r_{\rho} \) represent the scaled search directions for updating the variables \( \Delta z, \eta \) and \( \rho \) respectively. \( D_{\eta} \) and \( D_{\rho} \) are diagonal matrices whose elements are the current values of \( \eta \) and \( \rho \), i.e. \( D_{\eta} = \text{diag}(\eta) \) and \( D_{\rho} = \text{diag}(\rho) \). \( e \) is a vector of all ones. Assuming that an initial starting point in the interior of the feasible region is known, the solution is updated as follows.

\[ \Delta x^{k+1} = \Delta x^k + \alpha p_x^k \quad (12) \]
\[ y^{k+1} = y^k + \alpha p^k \]
\[ p^{k+1} = p^k + \alpha \rho p^k \]

The value of \( \alpha \) is chosen as 0.999 where \( \alpha_m \) is determined as follows.

\[ \alpha_m = \min \left\{ \left( -\frac{\eta_i}{p_i} : p_i < 0 \right), \left( -\frac{\rho_i}{p_i} : p_i > 0 \right) \right\} \]

This is simply a component wise ratio test of two vectors that selects the minimum ratio as the step size to ensure that feasibility is maintained.

**Dual Formulation**

The dual formulation of the problem defined in (4) was first suggested by Wagner [32].

\[ \max \Delta z^T \pi \]

subject to

\[ H^T \pi = 0 \]
\[ -w \leq \pi \leq w \]

The constraints in this dual problem result in a smaller least squares problem that has the same structure as the WLS state estimation problem. The original primal variables \( \Delta z \) are recovered from the solution of a least squares problem while the search direction for the dual variables \( \pi \) is given by the residuals \( r \) of the same least squares problem. We implemented an affine scaling solution method [30] as outlined below. The bounds on the variables are treated implicitly rather than as explicit constraints. This is reflected in the choice of the scaling matrix \( \tilde{D} \) and the step size. The scaling reflects how close the components of the solution \( (\pi_i) \) are to the bounds \( (w_i \text{ or } -w_i) \) and the step size preserves feasibility with respect to these bounds.

1. Set \( \pi^0 = 0 \) as the initial interior point.
2. Solve the following system for the search direction \( r \) and the original primal variables \( \Delta z \).

\[ \left( \begin{array}{cc} D^{-2} & H \\ H^T & 0 \end{array} \right) \left( \begin{array}{c} \pi \\ \Delta z \end{array} \right) = \left( \begin{array}{c} \Delta z \\ 0 \end{array} \right) \]

3. Choose the step size by finding

\[ \beta = \max \left\{ \left( \frac{r_i}{w_i - \pi_i} \right), \left( \frac{-r_i}{w_i + \pi_i} \right) \right\} \]

4. Update variables according to

\[ x^{k+1} = x^k + \frac{0.97}{\beta} r \]

and repeat until convergence.

This method was used by Meketon for linear least absolute value regression [23]. A barrier function method similar to that used for the primal problem could also be applied to the dual problem. In this case, the following sub-problem is obtained

\[ \max \Delta z^T \pi + \mu \sum_{i=1}^{m} \ln(w_i - \pi_i) + \ln(w_i + \pi_i) \]

subject to the constraints in equation (17).

The least squares problem required to obtain the search direction for this problem is similar in structure to that for the affine scaling method. Based on these observations, the dual formulation seems to offer advantages over the primal one. A comparison of search directions in affine scaling and barrier function interior point methods is given in the appendix.

**3 Initialization**

Interior point methods require a point inside the feasible region for initialization as opposed to the simplex which requires a feasible extreme point at a vertex of the feasible region. For the primal formulation, an initial extreme point is available by inspection, while an initial interior point requires additional effort. For the dual formulation, the converse is true. In general, to obtain an initial feasible interior point, a phase 1 procedure has to be used to solve an augmented problem by introducing an artificial variable. This involves extra computation. Moreover, the extra column introduced into the artificial problem is often dense and has an adverse effect on the sparsity of the least squares problem. However, due to the special structure of the \( \ell_1 \) problem, an initial feasible interior point may be obtained without resorting to a phase 1 procedure. For the primal problem, this is done as follows:

1. Set the \( n \) components of \( \Delta x \) to zero.
2. Set \( e = \Delta z \) where \( e = \eta - \rho \), \( \eta \geq 0 \) and \( \rho \geq 0 \).
3. If \( e_i > 0 \), set \( \eta_i = |e_i| + \frac{1}{3} |e_i| \) and \( \rho_i = \frac{5}{3} |e_i| \).
4. If \( e_i < 0 \), set \( \rho_i = |e_i| + \frac{1}{3} |e_i| \) and \( \eta_i = \frac{5}{3} |e_i| \).
5. If \( e_i = 0 \), then \( \rho_i = \eta_i \neq 0 \).

If \( e_i = 0 \), we assume \( \eta_i = \rho_i = 1 \). This is clearly not possible at the solution point, but it ensures that the initial interior point has all components strictly greater than zero.

The dual formulation uses the origin as an obvious starting point. Thus, while initialization is simpler for the primal formulation when the simplex method is used, it is simpler for the dual formulation when interior point methods are used.
4 Implementation Issues

The least squares problem which forms the major computational burden in any interior point method can be solved using one of many techniques. We restrict our attention to solution by direct methods and compare normal equations with the augmented formulation. The augmented formulation is usually numerically more stable [17]. It involves a larger but usually sparser system of equations. This is especially true for the primal formulation. The number of non-zero elements in the LU factors of the augmented system can be considerably less than the number of non-zero elements in the LU factors of the normal equations. The Harwell routine MA27 was used to solve the augmented problem. This routine automatically forms $1 \times 1$ and $2 \times 2$ blocks in the factorization of symmetric matrices. The structure of the augmented matrix and its LU factors for the primal formulation, using a 5 bus system, is shown in Figure 1. For the dual formulation the sparsity structure of the matrices is not shown as it is identical to that of the conventional weighted least squares method.

There are other reasons for using the augmented formulation in interior point methods. The residuals of the least squares problem (which give the search direction) are available as a part of the solution in the augmented formulation. Moreover, in the primal formulation the augmented system treats unrestricted variables properly. Unrestricted variables are simply assigned a zero diagonal entry in the corresponding matrix in equation (11). This approach cannot be followed in the normal equations case where unrestricted or free variables have to be treated in a less elegant manner by adding a large constant as was done in many early papers on WLAV state estimation. Unrestricted variables are also sometimes handled by incorporating them as the difference of two non-negative variables. Problems with this technique were reported in [22].

WLAV solutions are fundamentally different from WLS solutions. They satisfy a subset of $n$ measurements exactly while the remaining $m - n$ measurements have non-zero residuals. This solution occurs at a vertex of the feasible region. Interior point methods approach the solution asymptotically. It is therefore possible to identify the set of exact measurements before actually reaching the solution. In such a case, computational effort can be reduced by terminating the iterations early and solving an exactly determined system for the solution. Bloomfield and Stieger [7] give a test for identification of LAV solutions which can be used for the WLAV case. The interpolating property of WLAV solutions can be exploited to compute fast approximations to the true solution by solving exactly determined systems [27]. This is based on the assumption that the set of exactly satisfied measurements does not change with the same frequency as the numerical values of the measurements.

5 Test Results

Experiments were conducted using three test systems; a 5 bus system, the IEEE 30 bus system and the IEEE 118 bus system. Details of the measurement configurations used are given in Table 1.
The variation in the objective function with the iterations is shown in Figure 2 for the 30 bus case. Every iteration decreases the objective function by a certain factor. In approaches based on successive linearization, it is not necessary to find the optimal solution in the early linearizations. It is useful to limit the number of iterations in these early linearizations to reduce the computational effort. Due to the initialization procedure required after each linearization, the objective function actually goes up a little as shown in Figure 2. The extra effort undertaken in the first linearization is thus wasteful. Figure 3 shows the objective function variation when the iterations in the early linearizations are limited. This technique helps reduce the total computational effort. However, the choice of such a limit can sometimes be significant in the primal problem. If the limits are imposed prematurely it may result in extra linearizations which may be more expensive than extra iterations within a linearization, due to the re-evaluation of the Jacobian matrix $H$. In our tests, we factored the augmented matrix at every iteration. However, it may be better to use compensation methods that exploit the fact that only selected elements change from one iteration to another, i.e., the elements of the diagonal sub-matrices $D_n$ and $D_p$. The number of iterations as a function of the system size is shown in Figure 4. The number of iterations for the different systems with and without iteration count limits in the early linearizations is given in Table 2.

Table 1: Measurement configurations

<table>
<thead>
<tr>
<th>System</th>
<th>5 bus</th>
<th>30 bus</th>
<th>118 bus</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lines</td>
<td>7</td>
<td>41</td>
<td>179</td>
</tr>
<tr>
<td>Measurements (m)</td>
<td>23</td>
<td>118</td>
<td>472</td>
</tr>
<tr>
<td>Voltages</td>
<td>3</td>
<td>10</td>
<td>44</td>
</tr>
<tr>
<td>Measured Injections</td>
<td>4</td>
<td>24</td>
<td>132</td>
</tr>
<tr>
<td>Zero Injections</td>
<td>2</td>
<td>12</td>
<td>10</td>
</tr>
<tr>
<td>Line Flows</td>
<td>14</td>
<td>72</td>
<td>286</td>
</tr>
</tbody>
</table>

Table 2: Number of iterations to convergence with and without iteration count limits. Dual problem has 2 iterations per linearization.

<table>
<thead>
<tr>
<th>System</th>
<th>Iterations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Primal (no limits)</td>
</tr>
<tr>
<td>5 bus</td>
<td>7+5 $^\dagger$</td>
</tr>
<tr>
<td>30 bus</td>
<td>8+8+5</td>
</tr>
<tr>
<td>118 bus</td>
<td>18+9+8</td>
</tr>
</tbody>
</table>

($^\dagger$) 7+5 signifies 7 and 5 iterations in the first and second linearizations respectively.

An important observation from our tests was that a limit of two iterations per linearization always worked well for the dual problem. This is a major reason for the absence of any growth in the number of total iterations for the dual formulation of the non-linear WLAV problem. A similar observation was made recently in the case of the non-linear quantile regression problem by Koenker and Park [20]. A limiting of iterations in the dual problem is permissible since the variables $\Delta x$ are unrestricted, and hence are always feasible for the primal problem. The convergence criterion used in our tests was based on the infinity norm of the updates to the solution vector. The estimates from the interior point methods were the same as those obtained using the Barrodale-Roberts [5] method. Conventional WLS estimates were obtained for the test systems in 2-3 iterations, each involving the same effort as the dual formulation. This suggests the dual formulation is at least two to three times more computationally intensive than conventional WLS. However, the bad data rejection property of WLAV estimation has to be taken into account while making such a comparison.

The results for the dual formulation are significantly better than the primal formulation used in this paper. They are also better than the results presented in [9] as far as iteration counts are concerned. 

$^1$In [9] the 30 and 118 bus systems were solved in 10 and 26 iterations respectively. The measurement redundancy was slightly lower than 2. The dimension and structure of the system solved in every iteration was the same as that in our dual formulation.
6 Conclusions

Interior point methods can be used to solve the WLAV state estimation problem. The paper compared the primal and dual formulations and illustrated some of the advantages of the dual formulation when interior point methods are used. The special structure of the problem was used to obtain a feasible starting interior point in both formulations. Computational issues were addressed and the advantages of using an augmented formulation were illustrated, including the handling of free variables. The success of the method depends both on the number of iterations and the work per iteration. The dual formulation used in the paper seems to be better than the primal formulation in both aspects.

Acknowledgment

Support from the National Science Foundation under grant ECS-9215271 is gratefully acknowledged. The authors also thank the reviewers for their suggestions.

REFERENCES

The primal affine scaling direction is defined by the Newton direction for the following problem

$$\min_x \left\{ -\sum_{i=1}^n \ln x_i \mid Ax = b, x > 0 \right\} \quad (A.3)$$

This is equivalent to finding the analytic center of the polytope \( \{ x \mid Ax = b, x \geq 0 \} \). The Newton direction \( p_{\text{cent}} \) is given by

$$\begin{pmatrix} \hat{X}^{-2} A^T \\ A \\ 0 \end{pmatrix} \begin{pmatrix} -\mu p_{\text{cent}} \\ \lambda \end{pmatrix} = \begin{pmatrix} -X^{-1}e \\ 0 \end{pmatrix} \quad (A.4)$$

It is shown in [11] that the search direction used in [15] is \( p_{\text{aff}} + \mu p_{\text{cent}} \). We applied this method to the primal formulation in this paper. The search directions in most interior point methods can be written as some linear combination of the affine scaling and centering directions [11]. The reader is referred to [11, 15, 30] for further details regarding treatment of bounds, step size selection and other issues.

Fernando L. Alvarado (F’93) obtained a BS degree from the National University of Engineering in Lima, Peru, a MS degree from Clarkson University, and a Ph.D. from the University of Michigan. He is currently a Professor at the University of Wisconsin in Madison in the Department of Electrical and Computer Engineering. His main interests are in computer applications to power systems and large sparse matrix problems.

Harmohan Singh (S’87) received the BE(Hons) degree from Panjab University, India in 1987 and the MS degree from The University of Alabama, in 1990, both in electrical engineering. He is currently studying for a Ph.D. at the University of Wisconsin-Madison. His research interests include the application of optimization techniques in power systems. He has taught at The University of Alabama and at UW-Madison, where he has been a research assistant since 1990. He worked with the Pacific Gas and Electric Company in the summer of 1992 and is a member of IEEE, SIAM and Eta Kappa Nu.

Appendix

Consider the standard linear programming problem

$$\min \{ c^T x \mid Ax = b, x \geq 0 \} \quad (A.1)$$

The scaled version of this problem is

$$\min \{ \hat{c}^T \hat{x} \mid \hat{A}\hat{x} = \hat{b}, \hat{x} \geq 0 \} \quad (A.2)$$

where \( \hat{c} = c^T X, \hat{A} = \bar{A}X \) and \( \hat{x} = X^{-1}x \). \( X \) is a diagonal matrix with the components of \( x \) along the diagonal, where \( x \) is the current iterate. The scaling ensures that every component of \( \hat{x} \) is at least one unit away from the walls of the positive orthant which are unaffected by the scaling. The primal affine scaling direction, \( p_{\text{aff}} \), is defined as the projected gradient direction for (A.2). This direction is \( \hat{c} \) projected onto the null space of \( \hat{A} \) and rescaled back to the original coordinates by \( X \). It is found by solving a least squares problem. This is the direction used in primal affine scaling methods, one of which [30] was applied in this paper to the dual formulation.