APPLICATION OF OPTIMAL MULTIPLIER METHOD IN WEIGHTED LEAST-SQUARES STATE ESTIMATION PART I: THEORY

Jianping Meng Christopher L. DeMarco

Department of Electrical and Computer Engineering

University of Wisconsin-Madison

Madison, Wisconsin 53706

Abstract: Standard algorithms for state estimation may be viewed as quasi-Newton's methods applied to the first order optimality conditions of a least squares minimization problem. Previous work in the literature has documented the (somewhat surprising) fact that when a full Newton's method is applied to the same formulation, convergence properties are far worse than the quasi-Newton's method, until the iterates reach an EXTREMELY small neighborhood of the solution. Motivated by these results, and by availability of efficient algorithms to compute higher order derivatives necessary in an exact Newton formulation, this paper proposes several Newton's method variants to improve state estimator convergence. In the companion paper [3] benchmarks for the IEEE 118 and 300 bus test systems are provided, with comparisons against classical normal equation method, Hatchel's method, and QR algorithms. In these benchmark examples, the new algorithms developed show more reliable convergence for ill-conditioned cases, while making minimal sacrifices in computational efficiency for well-conditioned cases.

Keywords: Weighted Least-Squares(WLS) State Estimation, Optimal Multiplier, Exact Newton's Method, Quasi-Newton's Method

I. Introduction

There exists a long standing literature on problems of equilibrium state estimation in electric power systems. While variants exist, many methods begin from an optimization problem associated with minimizing a weighted sum of squares measurement residuals, under the assumption that the "ideal" measurement values are a known function of the desired system state. This optimization problem is often augmented with equality constraints for functions of state which are assumed to be known exactly; e.g. network buses having identically zero load. In typical formulations, the measurement functions are smooth, and often infinitely many times continuously differentiable. Hence, without previous knowledge of the long literature in the subject, a plausible approach to such a constrained optimization problem might be to form the Kuhn-Tucker necessary conditions for optimality, and apply a Newton's method to solution of the equations thus formed. Starting from this premise, the classical approach in the state estimation literature may be interpreted as a quasi-Newton, or "dishonest Newton", method, in which a computational convenient approximation to the Jacobian of the Kuhn-Tucker conditions is employed, rather than the exact Jacobian. In the terminology of the state estimation literature, the approximate Jacobian is the "gain" matrix for the problem.

This observation is hardly novel. For the case in which exact equality constraints are not appended to the problem, the particular gain matrix employed is quite natural. In particular, the standard gain matrix yields an iteration step that coincides with the minimization of the quadratic cost function formed by replacing the residual function by its first order Taylor expansion about the previous iterate. However, it is a natural question to consider what improvements in convergence properties might be gained by employing information from an exact Jacobian evaluation, and with it, an exact Newton iteration. Standard results ensure that for a problem with an isolated to the Kuhn-Tucker equations, one should see quadratic convergence to this point from initial conditions in a neighborhood of the point. In contrast, one would typically expect the quasi-Newton's method to display only linear convergence rates from a neighborhood of the solution, and perhaps a smaller region of attraction to the solution.

The motivation for the paper here stems primarily from the work of [1] and the Ph.D. thesis [11], which provide surprising numerical results when examining the use of full Newton's methods in least squares state estimation. In particular, [1] and [11] describe a range of numerical experiments in which use of the exact Newton's method, from typical initial conditions (flat start, or slight variants thereof), consistently diverge, while the standard quasi-Newton's method converges reliably. Moreover, [1] attempts to gain the advantage of quadratic convergence rates from the exact Newton, by beginning from quasi-Newton steps early in the iteration process, and "switching" to an exact Newton latter in the process. These experiments indicated that the iterates needed to be within an extremely small neighborhood of the solution in order for this benefit to be realized. To the state estimation neophyte, this would seem particularly surprising, given the robust convergence properties that exact Newton methods display for many (but certainly not all) power flow examples.

The work in [1] provides one interpretation of the difficulties of the exact Newton method, by examining the spectral radius of the iteration mapping for exact versus quasi-Newton, showing that the quasi-Newton had more desirable spectral radius properties until the iterates moved very close to the final accumulation point. However, in a sense, this is simply a more quantitative indicator of the poor convergence properties already observed qualitatively. While we have wrestled with the development of a structural explanation of the weak performance of the exact Newton, we must honestly report that we have no further insights beyond those of [1].

However, motivated by [1], this paper examines more pragmatic issues of algorithm design, aimed at improving convergence of least squares state estimators: (i) efficient calculation of the exact Jacobian of the Kuhn-Tucker conditions, in both polar and rectangular coordinate representations of the residual function; (ii) improved means of alternating or interpolating between exact Newton and quasi-Newton steps; (iii) use of optimal multipliers (computable at very low cost in a rectangular coordinate formulation) to improve convergence of steps in exact Newton or quasi-Newton directions. The problem proves a challenging one, as our experience in a range of randomly generated measurement sets for the IEEE 118 and 300 bus systems shows a high degree of variation in performance between various methods, case by case. Moreover, in the specific test cases we constructed, some highly regarded existing methods for improving convergence (QR, Hatchel's "sparse tableau" formulation) performed surprisingly poorly. However, by combining the concepts (i) through (iii) above, we will propose a Newton method variant that displays reliably improved convergence for a wide range ill-conditioned cases, while sacrificing relatively little in computational efficiency in well conditioned cases.

II. Notation

A key element of this paper's development will be treatment of the least squares estimation problem in both polar and rectangular coordinates. This allows examination of closed form optimal step size computations that are available in the rectangular form. It is therefore useful to define notation for both complex vector quantities, as well as distinct notation to separately represent corresponding real and imaginary parts, and corresponding magnitude and phase.

To do so, let the study network be composed n buses, and l branches. We assume a simply connected network, with a single angle reference bus. The number of generator buses is denoted n_g. The number of measurements is denoted m.

The following notation will be used throughout this paper, and the companion paper [3]. Upper case bold face Roman characters denote matrices; underlined lower case characters represent column vector quantities.

$\mathbf{A}$:primitive incidence matrix, $\mathbf{A} \in \Re^{n \times 2l}$
$\mathbf{Y}_p$:primitive admittance matrix, $\mathbf{Y}_p \in \Re^{2l \times 2l}$
$\mathbf{Y}_b$:bus admittance matrix, $\mathbf{Y}_b \in \Re^{n \times n}$
$\underline{s}$:bus injection complex power vector, $\underline{s} \in \mathbf{C}^n$
$\underline{e}$:bus voltage real part vector, $\underline{e} \in \Re^n$
$\underline{f}$:bus voltage imaginary part vector, $\underline{f} \in \Re^n$
$\underline{f}_r$:reduced bus voltage imaginary part vector, $\underline{f}_r \in \Re^{n-1}$
$\underline{\vec{v}}$:bus voltage complex vector, $\underline{\vec{v}} \in \mathbf{C}^n$
$\underline{\nu}$:bus voltage magnitude square vector, $\underline{\nu} \in \Re^n$
$\underline{i}$:bus injection current vector, $\underline{i} \in \mathbf{C}^n$
$\underline{x}$:reduced state variable vector, $\underline{x} = [\underline{e} \;\; \underline{f}_r]^T$, $\underline{x} \in \Re^{2n-1}$
$\underline{h}_1(\underline{x})$:active line flow(at both ends) vector
$\underline{h}_2(\underline{x})$:reactive line flow(at both ends) vector
$\underline{h}_3(\underline{x})$:active bus injection vector
$\underline{h}_4(\underline{x})$:reactive bus injection vector
$\underline{h}_5(\underline{x})$:bus voltage magnitude square vector
$\underline{h}(\underline{x})$:measurement function vector, $\underline{h} = [\underline{h}_1 \;\; \underline{h}_2 \;\; \underline{h}_3 \;\; \underline{h}_4 \;\; \underline{h}_5]^T$
$\mathbf{H}(\underline{x})$:measurement Jacobian matrix, $\mathbf{H} \in \Re^{m \times (2n-1)}$
$\underline{\hat{z}}$:measurement vector, $\underline{\hat{z}} \in \Re^m$
$\underline{\epsilon}$:measurement noise vector with normal distribution, $\underline{\epsilon} \in \Re^m$
$\mathbf{R}$: the diagonal noise covariance matrix, $\mathbf{R} \in \Re^{m \times m}$

The inverse of $\mathbf{R}$ will be referred to as the measurement weighting matrix throughout this paper, and is assumed to be diagonal. If one wishes to view the least squares estimate as a maximum likelihood estimator with additive measurement errors having a Gaussian distribution, this would correspond to an assumption that measurement errors are uncorrelated.

III. Review of Basic WLS State Estimation Algorithm

The mathematical model of power system estimation used by the classic least-squares algorithm is given by:

$$\underline{\hat{z}} = \underline{h}(\underline{x}) + \underline{\epsilon} \tag{1}$$

As noted above, under the assumption of additive, Gaussian measurement error, the maximum likelihood estimate $\underline{x}^*$ is found by solving the simple unconstrained minimization problem:

$$minimize\, J(\underline{x}) = \frac{1}{2}\left[\underline{\hat{z}} - \underline{h}(\underline{x})\right]^T \mathbf{R}^{-1}\left[\underline{\hat{z}} - \underline{h}(\underline{x})\right] \tag{2}$$

As a necessary condition for optimality, any solution will satisfy the first order Kuhn-Tucker condition

$$\mathbf{H}(\underline{x})^T \mathbf{R}^{-1}\left[\underline{\hat{z}} - \underline{h}(\underline{x})\right] = \underline{0} \tag{3}$$

If one applies a Newton-Raphson method, one seeks an estimate $\underline{x}^*$ which corresponds to an accumulation point of the iteration

$$\underline{x}^{k+1} = \underline{x}^k - [\nabla^2 J(\underline{x}^k)]^{-1} \nabla J(\underline{x}^k) \tag{4}$$

where $\nabla J(\underline{x})$ is column vector of the first order derivative of the cost function w.r.t. the state variable, and $\nabla^2 J(\underline{x})$ is the second order derivative of the cost function w.r.t. the state variables. Employing the specific form of $J(x)$ yields

$$\nabla J(\underline{x}) = \frac{\partial J(\underline{x})}{\partial \underline{x}} = -\mathbf{H}^T(\underline{x})\mathbf{R}^{-1}\Delta\underline{z} \tag{5}$$

$$\nabla^2 J(\underline{x}) = \frac{\partial \nabla J(\underline{x})}{\partial \underline{x}} = \mathbf{G}(\underline{x}) - \frac{\partial \mathbf{H}^T(\underline{x})}{\partial \underline{x}}\mathbf{R}^{-1}\Delta\underline{z} \tag{6}$$

where $\mathbf{G}(\underline{x}) = \mathbf{H}^T(\underline{x})\mathbf{R}^{-1}\mathbf{H}(\underline{x})$ and $\Delta\underline{z} = \underline{\hat{z}} - \underline{h}(\underline{x})$ is the residual.

If one considers the linearized state estimation process, and ignores the second order derivative term within $\nabla^2 J(\underline{x})$, the traditional normal equation method is obtained

$$\underline{x}_{k+1} = \underline{x}_k + \mathbf{G}^{-1}(\underline{x}^k)\mathbf{H}^T(\underline{x}^k)\mathbf{R}^{-1}\Delta\underline{z}^k \tag{7}$$

Since the iteration process above uses only an approximation to $\nabla^2 J(\underline{x})$, one can consider this a "quasi Newton's method".

Here the matrix defining the linear equation set is $\mathbf{G}(\underline{x})$; this is often referred to as the "gain matrix" in state estimation literature. For our development, we will refer to this quantity as the "approximate gain" matrix and denote it as $\mathbf{G}_a(\underline{x})$. If one does not ignore the second order derivative information represented by $\bigtriangledown^2 J(\underline{x})$, the following iteration process is obtained

$$\underline{x}_{k+1} = \underline{x}_k + \mathbf{G}_e(\underline{x}^k)^{-1}\mathbf{H}^T(\underline{x}^k)\mathbf{R}^{-1}\Delta\underline{z}^k \tag{8}$$

This is the exact Newton's method. Here the gain matrix is $\mathbf{G}_e(\underline{x}) = \mathbf{G}(\underline{x}) - \frac{\partial \mathbf{H}^T(\underline{x})}{\partial \underline{x}}\mathbf{R}^{-1}\Delta\underline{z}$. We will refer to it as the "exact gain" matrix.

The normal equation method has proven popular because of a number of convenient numerical properties it offers. It preserves sparsity in a network having relatively few branches incident on each node. Moreover, in the linear equation solution the gain matrix is symmetric and positive definite, allowing relatively simple algorithms to be efficiently employed in the factorization process. For many practical cases, the normal equation method offers very fast convergence at low computational cost per iteration. However, if the system displays any of a variety of problems that lead to ill-conditioning, the normal equation method will give rise to numerical problems. One common ill-conditioning problem arises from exact zero injections. In a power system, there are typically a number of buses from which no external current is drawn of injected, so that the complex power injection is physically constrained to be exactly zero. While no measurement device is present at such a bus, it is useful to impose the zero power injection constraint. Such a zero injection is typically referred to as a virtual measurement, and can be incorporated in the measurement set along with physically measured injections. Such virtual measurements may impose equality constraints appended to the previously developed Kuhn-Tucker conditions for optimality, or they may be approximately incorporated in the original objective function, with large weights (a penalty function approach to equality constraints) When a large number of these virtual measurements are incorporated in the objective function, with large weights, ill-conditioning of the gain matrix can result [10]. In such cases normal equation method often displays convergence problems.

To a naive observer, approaching the state estimation problem from a background of iterative power flow calculations, it might seem that the in such a case, use of an exact Newton method (8) might be likely to improve convergence properties, at the cost of more computation per iteration to evaluate to the second order derivative term. However, numerical experience reported in the literature shows quite the opposite conclusion [1] and [11].

As documented in [1] (and re-enforced by our own numerical experience), in the early iterations, use of exact Newton's method significantly degrades the convergence properties. In particular, starting from flat voltage, use of the exact Newton often results in extremely large steps, and as a practical matter, immediate divergence ([11] offers a simple two bus illustrative example in which this property is illustrated dramatically). Certainly, there is no contradiction to general analytic properties of the Newton's method in these observations, as the Newton's method is guaranteed to be a contraction mapping with quadratic convergence properties only in some sufficiently small neighborhood of a solution point. The expectation of quadratic convergence for exact Newton, versus linear convergence rate for quasi-Newton (i.e., traditional normal form with the standard approximate gain matrix) is observed in the numerical examples of [1], for very small neighborhoods of solutions.

Hence, [1] finishes its exposition with the pragmatic suggestion that a state estimation iteration should start with the standard normal equation method in its first two to three iterations, and then switch to exact Newton's method for the remaining steps to convergence. Accepting this heuristic as the existing state of the art for exploiting the second order information of the exact gain matrix, our exposition "picks up" from this point.

IV. Measurement Functions in Rectangular Coordinates and Their Taylor Expansion

As noted previously, one of the attractive features of rectangular coordinate representations is that once a step direction is determined for bus voltage phasor quantities, the power mismatch equations are simple quadratic expressions with respect to step size. Hence, step size selection criterion that are themselves a simple function of mismatch often has minima that can be trivially computed. This property extends in an obvious way to the least squares state estimation problem. Given any method of selecting a step direction, one can trivially compute the step size that minimizes the least squares objective function along that ray. This motivates our choice of a rectangular formulation here.

To illustrate the properties of the rectangular formulation, we describe the measurement functions in a simple closed form fashion, consistent with evaluation MATLAB [12].

$$\underline{h}_1(\underline{x}) = real\left[\mathbf{A}' * \underline{\vec{v}}. * (\mathbf{Y}_p^* * \mathbf{A}' * \underline{\vec{v}}^*)\right] \tag{9}$$
$$\underline{h}_2(\underline{x}) = imag\left[\mathbf{A}' * \underline{\vec{v}}. * (\mathbf{Y}_p^* * \mathbf{A}' * \underline{\vec{v}}^*)\right] \tag{10}$$
$$\underline{h}_3(\underline{x}) = real\left[\underline{\vec{v}}. * (\mathbf{Y}^* * \underline{\vec{v}}^*)\right] \tag{11}$$
$$\underline{h}_4(\underline{x}) = imag\left[\underline{\vec{v}}. * (\mathbf{Y}^* * \underline{\vec{v}}^*)\right] \tag{12}$$
$$\underline{h}_5(\underline{x}) = \underline{e}^2 + \underline{f}^2 \tag{13}$$

where:

$$\underline{\vec{v}} = \underline{e} + j\underline{f}$$

It is obvious that the above measurement functions are quadratic functions of the state variables $\underline{e}, \underline{f}_r$. Note the imaginary part of the reference bus voltage magnitude is always 0.

As noted previously, the measurement functions are quadratic. In exact analogy to developments for the power mismatch equations for power flow [6], the Taylor has three terms, which may be conveniently expressed as

$$\underline{h}(\underline{x}) = \underline{h}(\underline{x}_e) + \mathbf{H}\Delta\underline{x} + \underline{h}(\Delta\underline{x}) \tag{14}$$

where: $\underline{x}_e$: estimate of $\underline{x}$, $\Delta\underline{x}$: error(correction vector)

For the a derivation of the measurement Jacobian matrix, please refer to the companion paper [3]. As noted in the introduction, one of the motivations for the work presented here is the observation that the structure of the Jacobian matrix as presented in [3] allows very efficient sparse matrix computation of the second order derivative terms.

V. WLS State Estimation Problems by Nonlinear Programming Approach

The objective of WLS algorithm for power system state estimation is to minimize the following cost function.

$$F(\underline{x}) = \frac{1}{2}\left[\hat{\underline{z}} - \underline{h}(\underline{x})\right]^T \mathbf{R}^{-1}\left[\hat{\underline{z}} - \underline{h}(\underline{x})\right] \tag{15}$$

If there are zero injection buses, we will adopt the penalty function approach to their treatment, and associate large weights to these virtual measurements. Let $\mathbf{W} = \mathbf{R}^{\frac{1}{2}}$, then

$$F(\underline{x}) = \frac{1}{2}\left[\mathbf{W}(\hat{\underline{z}} - \underline{h}(\underline{x}))\right]^T\left[\mathbf{W}(\hat{\underline{z}} - \underline{h}(\underline{x}))\right] \tag{16}$$

In the presentation to follow, we will offer a number of methods that differ in their choice of step direction computation. However, independent of the step direction, the proposed methods all share a general structure, as outlined below.

Step 1. From a known iterate value, $\underline{x}_e^k$, construct the cost function F^k as a second order Taylor expansion about $\underline{x}_e^k$.

$$F^k = \frac{1}{2}\left[\mathbf{W}(\hat{\underline{z}} - \underline{h}(\underline{x}_e^k + \mu^k\Delta x^k)\right]^T\left[\mathbf{W}(\hat{\underline{z}} - \underline{h}(\underline{x}_e^k + \mu^k\Delta x^k)\right] \tag{17}$$

where k is the iteration step

Step 2. Determine a step direction, $\Delta\underline{x}^k$.

Step 3. Compute the scalar multiplier μ^k so that minimizes the cost function along the selected step direction.

Step 5. Update estimate $\underline{x}_e^{k+1} = \underline{x}_e^k + \mu^k\Delta\underline{x}^k$.

Step 6. If norm of $\underline{x}_e^{k+1} = \underline{x}_e^k$ is sufficiently small, terminate; else set k = k + 1, go to the step 2.

As noted in works such as [6] and [2], power flow problems can be treated in the same least squares minimization framework, selecting a weighting matrix as the identity matrix, and functions h just as power mismatch expressions. If the cost function is driven to zero, the power flow solution is found; for injection values for which no power flow solution exists, [2] employees an iteration structure similar to that above to find a point that satisfies necessary conditions for minimizing the Euclidean norm of the mismatch. It is interesting to note that the work of [2] used an exact Newton's method to set step direction, and observed excellent convergence properties in all its examples. In the state estimation problems, we have already noted published results that suggest the experience will be very different. However, the work here will share the key idea of [2], and original work of [6], in selecting the optimal μ^k to accelerate the convergence. Standard numerical texts [4] offer a wide range of options in approximate optimal multiplier μ selection; the rectangular coordinate formulation affords us the luxury of easily computing an exact optimal multiplier.

VI. Derivation of The Optimal Multiplier Method

The derivation below closely follows that of [6]. From the previous section, we have already expressed the measurement functions as quadratic functions of the state variables. From the Taylor series expansion we have the following overdetermined equation.

$$\mathbf{W}\left[\hat{\underline{z}} - \underline{h}(\underline{x}_e) - \mathbf{H}\mu\Delta\underline{x} - \underline{h}(\mu\Delta\underline{x})\right] = \underline{0} \tag{18}$$

By multiplying the scalar μ we can adjust the correction vector $\Delta\underline{x}$. It follows that

$$\mathbf{W}\left[\hat{\underline{z}} - \underline{h}(\underline{x}_e) - \mu\mathbf{H}\Delta\underline{x} - \mu^2\underline{h}(\Delta\underline{x})\right] = \underline{0} \tag{19}$$

For simplicity we define the vectors $\underline{a}, \underline{b}, \underline{c}$ as follows

$$\begin{aligned}\underline{a} &= \mathbf{W}\left[\hat{\underline{z}} - \underline{h}(\underline{x}_e)\right] && (20)\\ \underline{b} &= -\mathbf{W}\mathbf{H}\Delta\underline{x} && (21)\\ \underline{c} &= -\mathbf{W}\underline{h}(\Delta\underline{x}) && (22)\end{aligned}$$

Then (19) can be simply written as below

$$\underline{a} + \mu\underline{b} + \mu^2\underline{c} = 0 \tag{23}$$

In order to determine the value of the optimal μ in a least-squared sense, the following cost function should be minimized.

$$F = \frac{1}{2}(\underline{a} + \mu\underline{b} + \mu^2\underline{c})^T(\underline{a} + \mu\underline{b} + \mu^2\underline{c}) \tag{24}$$

Candidate minimizers for the above equation can be obtained by solving:

$$\frac{\partial F}{\partial\mu} = 0. \tag{25}$$

That is

$$g_3\mu^3 + g_2\mu^2 + g_1\mu + g_0 = 0 \tag{26}$$

where

$$\begin{aligned}g_0 &= \underline{a}^T\underline{b} && (27)\\ g_1 &= \underline{b}^T\underline{b} + 2\underline{a}^T\underline{c} && (28)\\ g_2 &= 3\underline{b}^T\underline{c} && (29)\\ g_3 &= 2\underline{c}^T\underline{c} && (30)\end{aligned}$$

Solution of such a cubic equation can expressed in closed form via Cardano's formula, or at very low cost numerically via standard root finding routines.

VII. Application of Optimal Multiplier Method and Exact Newton Method to WLS State Estimation Problems

Standard measurement quantities in a state estimator are line flows, bus injections and bus voltage magnitudes. To facilitate our rectangular coordinate formulation, and in particular to maintain the advantage of quadratic measurement functions, we treat voltage magnitude measurements as providing information regarding the square of the voltage magnitude. If one were adopting a strict maximum likelihood estimator formulation, and strictly required that Gaussian errors entered as additive noise in the measurement of voltage magnitude, the transition to the square of the measurement would be problematic. However, a pragmatic approach recognizes that the additive Gaussian error assumption is far from realized in practice, and while physical equipment used to collect a voltage magnitude from a three phase system varies, techniques that construct a sum of squares of individual phase quantities are common. Hence, we may argue that treating the unlying measurement quantity as a square of voltage magnitude is at least reasonable, without undo concern for the distribution of the measurement error.

For the WLS state estimation problem we need to solve the equation (15). If the maximum component of the difference between the two consecutive state variable vector is less than the predefined tolerance, the converged result is obtained. We now propose three WLS state estimation algorithms based on optimal multiplier method.

Algorithm 1:
The computation procedure is as follows.
Step 1. Use the flat voltage, i.e. $\underline{e}$ is a all ones vector and $\underline{f}_r$ is a zeros vector, as the initial guess to start the algorithm. Note the imaginary part of the reference bus voltage magnitude is always zero.
Step 2. Solve the following overdetermined equation using normal equation method

$$\mathbf{W}\left[\underline{\hat{z}} - h(\underline{x}_e^k)\right] = \underline{0} \tag{31}$$

that is we need to solve

$$\mathbf{H}(\underline{x})^T \mathbf{R}^{-1}\left[\underline{\hat{z}} - \underline{h}(\underline{x})\right] = \underline{0} \tag{32}$$

to get $\Delta\underline{x}^k$ in the least-squares sense

$$\Delta\underline{x}^k = \mathbf{Ga}_k \backslash \left[\mathbf{H}_k^T \mathbf{R}^{-1}(\underline{\hat{z}} - \underline{h}(\underline{x}_e^k))\right] \tag{33}$$

Step. 3 Compute the following coefficients

$$\begin{aligned} \underline{a}^k &= \mathbf{W}[\underline{\hat{z}} - \underline{h}(\underline{x}_e^k)] && (34)\\ \underline{b}^k &= -\mathbf{W}\mathbf{H}^k \Delta\underline{x}^k = -\underline{a}^k && (35)\\ \underline{c}^k &= -\mathbf{W}\underline{h}(\Delta\underline{x}^k) && (36) \end{aligned}$$

to solve the cubic equation (26) to get the optimal multiplier μ^k.
Step 4. Update the estimate vector $\underline{x}^{k+1} = \underline{x}^k + \mu^k \Delta\underline{x}$, measurement Jacobian and the measurement residual $\underline{\hat{z}} - \underline{h}(\underline{x}^{k+1})$.
Step 5. If the solution converges, stop the iteration; otherwise go back to step 2. Note: in theory the new mismatch can also be computed by

$$\begin{aligned} &\mathbf{W}\left[\underline{\hat{z}} - \underline{h}(\underline{x}_e^{k+1})\right]\\ &= \mathbf{W}\left[\underline{\hat{z}} - \underline{h}(\underline{x}_e^{k+1} + \mu\Delta\underline{x}^k)\right]\\ &= \mathbf{W}\left[\underline{\hat{z}} - [\underline{h}(\underline{x}_e^k) + \mu\mathbf{H}^k\Delta\underline{x}^k + \mu^2\underline{h}(\Delta\underline{x}^k)]\right]\\ &= \mathbf{W}\left[\underline{\hat{z}} - h(\underline{x}_e^k) - \mu\mathbf{H}^k\Delta\underline{x}^k - \mu^2\underline{h}(\Delta\underline{x}^k)\right]\\ &= \underline{a}^k + \mu\underline{b}^k + \mu^2\underline{c}^k \end{aligned} \tag{37}$$

But in practice it's better to recompute the measurement function to reduce computation error.

Algorithm 2:
The computation procedure is as follows.
Step 1. Use the flat voltage as the initial guess to start the algorithm.
Step 2. Solve the following overdetermined equation using normal equation method

$$\mathbf{W}\left[\underline{\hat{z}} - h(\underline{x}_e^k)\right] = \underline{0} \tag{38}$$

that is we need to solve

$$\mathbf{H}(\underline{x})^T \mathbf{R}^{-1}\left[\underline{\hat{z}} - \underline{h}(\underline{x})\right] = \underline{0} \tag{39}$$

to get $\Delta\underline{x}^k$ in the least-squares sense

$$\Delta\underline{x}^k = \mathbf{Ga}_k \backslash \left[\mathbf{H}_k^T \mathbf{R}^{-1}(\underline{\hat{z}} - \underline{h}(\underline{x}_e^k))\right] \tag{40}$$

Step. 3 Compute the following coefficients

$$\begin{aligned} \underline{a}^k &= \mathbf{W}[\underline{\hat{z}} - \underline{h}(\underline{x}_e^k)] && (41)\\ \underline{b}^k &= -\mathbf{W}\mathbf{H}^k \Delta\underline{x}^k = -\underline{a}^k && (42)\\ \underline{c}^k &= -\mathbf{W}\underline{h}(\Delta\underline{x}^k) && (43) \end{aligned}$$

to solve the cubic equation (26) to get the optimal multiplier μ^k
Step 4. Update the estimate vector $\underline{x}^{k+1} = \underline{x}^k + \mu^k \Delta\underline{x}$, measurement Jacobian and the measurement residual $\underline{\hat{z}} - \underline{h}(\underline{x}^{k+1})$.
Step 5. If the solution converges stop the iteration, otherwise go back to step 2.
Step 6. Do above iteration 3 times.
Step 7. Now switch to exact Newton's method to get

$$\Delta\underline{x}^k = \mathbf{Ge}_k \backslash \left[\mathbf{H}_k^T \mathbf{R}^{-1}(\underline{\hat{z}} - \underline{h}(\underline{x}_e^k))\right] \tag{44}$$

Step 8. Compute the following coefficients

$$\begin{aligned} \underline{a}^k &= \mathbf{W}[\underline{\hat{z}} - \underline{h}(\underline{x}_e^k)] && (45)\\ \underline{b}^k &= -\mathbf{W}\mathbf{H}^k \Delta\underline{x}^k && (46)\\ \underline{c}^k &= -\mathbf{W}\underline{h}(\Delta\underline{x}^k) && (47) \end{aligned}$$

to solve the cubic equation (26) to get the optimal multiplier μ^k
Step 9. Update the estimate vector $\underline{x}^{k+1} = \underline{x}^k + \mu^k \Delta\underline{x}$ and measurement residual $\underline{\hat{z}} - \underline{h}(\underline{x}^{k+1})$.
Step 10. If the solution converges, stop the iteration; otherwise set k = k + 1 and go back to step 7.

Algorithm 3:
The computation procedure is as follows.
Step 1. Use the flat voltage as the initial guess to start the algorithm.
Step 2. Solve the following overdetermined equation using quasi newton method and exact Newton method

$$\mathbf{W}\left[\underline{\hat{z}} - h(\underline{x}_e^k)\right] = \underline{0} \tag{48}$$

that is we need to solve

$$\mathbf{H}(\underline{x})^T \mathbf{R}^{-1}\left[\underline{\hat{z}} - \underline{h}(\underline{x})\right] = \underline{0} \tag{49}$$

to get $\Delta\underline{x}_1^k$ and $\Delta\underline{x}_2^k$ in the least-squares sense

$$\Delta\underline{x}_1^k = \mathbf{Ga}_k \backslash \left[\mathbf{H}_k^T \mathbf{R}^{-1}(\hat{\underline{z}} - \underline{h}(\underline{x}_e^k))\right] \tag{50}$$

$$\Delta\underline{x}_2^k = \mathbf{Ge}_k \backslash \left[\mathbf{H}_k^T \mathbf{R}^{-1}(\hat{\underline{z}} - \underline{h}(\underline{x}_e^k))\right] \tag{51}$$

Step. 3 Compute the following coefficients

$$\begin{aligned}
\underline{a}_1^k &= \mathbf{W}[\hat{\underline{z}} - \underline{h}(\underline{x}_e^k)] && (52)\\
\underline{b}_1^k &= -\mathbf{W}\mathbf{H}^k \Delta\underline{x}_1^k = -\underline{a}_1^k && (53)\\
\underline{c}_1^k &= -\mathbf{W}\underline{h}(\Delta\underline{x}_1^k) && (54)\\
\underline{a}_2^k &= \mathbf{W}[\hat{\underline{z}} - \underline{h}(\underline{x}_e^k)] && (55)\\
\underline{b}_2^k &= -\mathbf{W}\mathbf{H}^k \Delta\underline{x}_2^k = -\underline{a}_2^k && (56)\\
\underline{c}_2^k &= -\mathbf{W}\underline{h}(\Delta\underline{x}_2^k) && (57)
\end{aligned}$$

to solve the cubic equation (26) to get the optimal multiplier μ^k and γ^k.

Step 4. Compare the two cost functions

$$J_1 = \left[\hat{\underline{z}} - \underline{h}(\underline{x}^k + \mu\Delta\underline{x}_1^k)\right]^T \mathbf{R}^{-1}\left[\hat{\underline{z}} - \underline{h}(\underline{x}^k + \mu\Delta\underline{x}_1^k)\right] \tag{58}$$

$$J_2 = \left[\hat{\underline{z}} - \underline{h}(\underline{x}^k + \gamma\Delta\underline{x}_2^k)\right]^T \mathbf{R}^{-1}\left[\hat{\underline{z}} - \underline{h}(\underline{x}^k + \gamma\Delta\underline{x}_2^k)\right] \tag{59}$$

If J_1 is less equal than J_2, $x^{k+1} = x^k + \mu\Delta\underline{x}_1^k$, otherwise $x^{k+1} = x^k + \gamma\Delta\underline{x}_1^k$

Step 5. Update the measurement residual $\hat{\underline{z}} - \underline{h}(\underline{x}^{k+1})$ and measurement Jacobian.

Step 6. If the solution converges, stop the iteration; otherwise go back to step 2.

VIII. Closing Remarks

This paper analyzes the existing drawback of the traditional normal equation method in power system state estimation problems. If the input measurement values are ill conditioned, the normal equation method will probably display poor numerical stability. Base on the optimal multiplier method which was successfully used in load flow computation to handle ill conditioned cases and second order exact Newton's method which has fast convergence speed when it reaches the small neighborhood of the solution, three Newton type algorithms combining the above two ideas are proposed to improve the state estimator convergence for ill conditioned cases. Detail formulation for these algorithms variants is given. Part II [3] of the companion paper will pick up at this point and discuss the simulation of the proposed algorithms in more detail.

References

[1] Robert A.M. van Amerongen, "On Convergence Analysis and Convergence Enhancement of Power System Least-Squares State Estimators", *IEEE Transactions on Power Systems*, Vol. 10, No. 4, November 1995.

[2] Thomas J. Overbye, "A Power Flow Measure for Unsolvable Cases", *IEEE Transactions on Power Systems*, Vol. 9, No. 3, August, 1994

[3] Jianping Meng, Christopher L. DeMarco "Application of Optimal Multiplier Method in Weighted Least-Squares State Estimation, Part II: Simulation", *The 31st North American Power Symposium*, San Luis Obispo, California, October, 1999.

[4] J.M. Ortega and W.C. Rheinboldt, *Iterative Solution of Nonlinear Equations in Several Variables*, Academic Press, New York and London, 1970.

[5] B. Stott, O. Alsac, "Fast Decoupled Load Flow", *IEEE Transactions on Power Apparatus and Systems*, Vol. PAS-93, pp. 859-869, May/June 1974.

[6] S. Iwamoto, Y. Tamura, "A Load Flow Calculation Method for Ill-conditioned Power Systems", *IEEE Transactions on Power Apparatus and Systems*, Vol. PAS-100, No.4, April 1981.

[7] M. Allam, "Exact Static State Estimation Algorithm Retaining Non-linearity", *Proc. PSCC VII*, pp. 1007-1013, July, 1981.

[8] N.D. Rao, L. Roy, "A Cartesian Coordinate Algorithm for Power System State Estimation",*IEEE Transactions on Power Apparatus and Systems*, Vol. PAS-102, No.5, May 1983.

[9] J.W. Gu, K.A. Clements, G.R. Krumpholz, P.W. Davis, "The Solution of Ill-conditioned Power System State Estimation Problems via The Method of Peters and Wilkinson", *IEEE Transactions on Power Apparatus and Systems*, Vol. PAS-102, No.10, October 1983.

[10] A. Monticelli, C.A.F. Murari, F.F. Wu, "A Hybrid State Estimator: Solving Normal Equations by Orthogonal Transformations", *IEEE Transactions on Power Apparatus and Systems*, Vol. PAS-104, No.12, December 1985.

[11] Charles William Hansen, *Model Enhancements for State Estimation in Electric Power Systems*, Ph.D. Thesis, School of Electrical and Computer Engineering, Georgia Institute of Technology, October, 1993.

[12] *Using MATLAB Version 5*, The Math Works Inc., 1996.

Biographies

Jianping Meng obtained his BS degree and two MS degrees all in electrical engineering from Tsinghua University, North China Institute of Electric Power and University of Wisconsin-Madison in 1991, 1994 and 1997 respectively. He is currently a Ph.D. candidate with the department of electrical and computer engineering at University of Wisconsin-Madison.

Christopher L. DeMarco obtained the BS degree from Masachussets Institute of Technology in 1980 and the MS and Ph.D. degrees from University of California-Berkeley in 1982 and 1985, all in electrical engineering. He is currently a professor with the department of electrical and computer engineering at University of Wisconsin-Madison.