State Estimation in Power Networks IV

Convergence Results

Ljung, Lennart; Lindahl, Sture

1974

Document Version:
Publisher's PDF, also known as Version of record

Link to publication

Citation for published version (APA):
Ljung, L., & Lindahl, S. (1974). *State Estimation in Power Networks IV: Convergence Results*. (Technical Reports TFRT-7041). Department of Automatic Control, Lund Institute of Technology (LTH).

Total number of authors:
2

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

TFRT-7041

STATE ESTIMATION IN POWER NETWORKS IV
CONVERGENCE RESULTS

LENNART LJUNG
STURE LINDAHL

Report 7420 Oktober 1974
Lund Institut of Technology
Division of Automatik Control

STATE ESTIMATION IN POWER NETWORKS IV
CONVERGENCE RESULTS

Lennart Ljung and Sture Lindahl

STATE ESTIMATION IN POWER NETWORKS IV
CONVERGENCE RESULTS

Lennart Ljung and Sture Lindahl

ABSTRACT

The convergence of a proposed method for state estimation in power systems is analysed for a case with constant state vector. In particular, a set into which estimates obtained by two versions of the SCI method converge are determined.

TABLE OF CONTENTS

1. Introduction.

The safe and economic operation of power systems requires the knowledge of certain variables in the power system. To avoid the overloading of transmission lines it is necessary to know the line flows in real time. To do an economic load dispatch calculation it is often required that the power demand at every bus in the system is known.

The power system state estimator is designed to use redundant measurements and a model of the power system in order to estimate the present state. Often the bus voltages are taken as state variables. The measurements can be bus voltage magnitudes, line current magnitudes, active and reactive line power flows, total bus current magnitude, and finally active and reactive bus power injections.

Several methods for power system state estimation have been proposed. A literature survey is given in [1].

To our knowledge there are no convergence results of the proposed methods reported in the literature. To avoid too much simulation and off-line checkout on real data it is valuable if some convergence results could be obtained. In this report we study the convergence properties for two versions of the SCI-method, proposed by R. E. Larson et. al. [6, 7, 8, 9]. The analysis is performed under the assumption that the state is constant but unknown.

2. Problem, Model and Methods.

Let $x(n)$ denote the p-dimensional state of the power system and let $y(n)$ be the r-dimensional observation vector (measurements) taken at time n. The observations are non-linear functions of the state and it is assumed that they are corrupted by additive noise:

$$y(n) = g(x(n)) + e(n) \tag{1}$$

where $\{e(n)\}$ is a sequence of random variables with zero mean and

$$E\, e(n)\, e(n)^T = R$$

Usually it is assumed that $\{e(n)\}$ is a sequence of gaussian, independent variables, but in this report we assume only that

$$E\, e(n)\, e(m) \leq \frac{C}{1 + |m-n|^{\alpha}} \quad , \alpha > 0$$

which seems to be more realistic.

It is assumed that g is twice continously differentiable and we denote

$$G(x_o) = \left.\frac{\partial g(x)}{\partial x}\right|_{x=x_o} \qquad \text{(r/p-matrix)}$$

Problem

Given the equations for the dynamics of the state, and the measurements $\{y(n), y(n-1), \ldots, y(0)\}$, determine an estimate $\hat{x}(n)$ of the state $x(n)$ at time n.

Model

An often used model for the dynamics of the state is

$$x(n+1) = x(n) + v(n) \tag{3}$$

where $\{v(n)\}$ is a sequence of independent, random vectors with $E\, v(n) = 0$ and

$$E\, v(n)\, v(n)^T = Q(n)$$

Method I (F. C. Schweppe, [2, 3, 4, 5]).

This method uses only one measurement vector and does not use any model of state dynamics.

$$\hat{x}_{i+1} = \hat{x}_i + K_i[y(n) - g(\hat{x}_i)] \qquad i = 1,\ldots,r \tag{5a}$$

$$K_i = P_i G^T(\hat{x}_i)R^{-1} \tag{5b}$$

$$P_i^{-1} = G^T(\hat{x}_i)R^{-1}\, G(\hat{x}_i) \tag{5c}$$

Method II (A non-sequential version of the SCI-method).

$$\hat{x}(n+1) = \hat{x}(n) + K(n)[y(n) - g(\hat{x}(n))] \tag{6a}$$

$$K(n) = P(n)G^T(\hat{x}(n))R^{-1} \tag{6b}$$

$$P^{-1}(n+1) = (P(n) + \mathrm{diag}(Q(n)))^{-1} + \mathrm{diag}[G^T(\hat{x}(n))R^{-1}G(\hat{x}(n))] \tag{6c}$$

$$P(0) = C_1 \cdot I \tag{6d}$$

Notice that P(n) is always diagonal. This choice is made in order to decrease the computational effort. Since the off diagonal elements are "stripped", the method has sometimes been called "stripped Kalman filtering".

Method III (The sequential version of the SCI-method, R. E. Larson et. al. [6, 7, 8, 9]).

$$\hat{x}(n+1)_j = \hat{x}(n+1)_{j-1} + K(n+1)_j[y(n)_j - g(\hat{x}(n+1)_{j-1})_j] \quad (7a)$$

$$K(n+1)_j = P(n+1)_j \; G^T(\hat{x}(n+1)_{j-1})_j \; R_{jj}^{-1} \quad (7b)$$

$$P^{-1}(n+1)_j = P(n+1)_{j-1}^{-1}$$
$$+ \operatorname{diag}[G^T(\hat{x}(n+1)_{j-1})_j \; R_{jj}^{-1} \; G(\hat{x}(n+1)_{j-1})_j] \quad (7c)$$

$$P(n+1)_0 = P(n)_r + \operatorname{diag}(Q(n)) \quad (7d)$$

$$P(0)_0 = \operatorname{diag}(P_{10}, P_{20}, \ldots, P_{r0}) = C_1 \cdot I \quad (7e)$$

$$\hat{x}(n+1)_0 = \hat{x}(n)_r$$

The difference to method II is that a new estimate is calculated after each component of the measurement vector is obtained

where

$\hat{x}(n)_j$	the estimate at line n based on j of the measurements at line n
$y(n)$	the j:th component of the measurements at line n
$g(x)_j$	the j:th component of g(x)
$G(x)_j$	the j:th row of G(x).
R_{jj}^{-1}	the jj:th component of R^{-1}

3. Convergence Results.

In this section the convergence properties of the algorithms (6) and (7) will be analysed. Since the estimates should track the time varying true state and the measurements are noisy, there is no possibility that $\hat{x}(n) - x(n)$ tends to zero as n increases. Therefore, the convergence analysis will deal with the idealized case

$$x(n+1) = x(n) = x^*$$

i.e. $Q = 0$ and we have a constant but unknown state vector. Such analysis will have relevance also in the case where the state is varying slowly in time ($|Q(n)|$ small). If the algorithm is not capable of converging to the true value in the case $Q(n) = 0$, then it will, in general, have poor tracking properties in the time varying case (3). The analysis is based on the theory for recursive stochastic algorithms given by Ljung in [10].

Theorem 1.

The estimate $x(n)$ generated by algorithm (6) with $Q = 0$ converges with probability one to the set

$$D_c = \{x | G^T(x)R^{-1}[g(x) - g(x^*)] = 0\} \cup \{\infty\}$$

as n tends to infinity.

Proof

Introduce $S(n) = n\,P(n)$. Then (6) can be written

$$\hat{x}(n+1) = \hat{x}(n) + \frac{1}{n} S(n)G^T(\hat{x}(n))R^{-1}[g(x^*) - g(\hat{x}(n)) + e(n)] \tag{8}$$

$$S^{-1}(n+1) = S^{-1}(n) + \frac{1}{n+1} \operatorname{diag}[G(\hat{x}(n))^T R^{-1} G(\hat{x}(n)) - S^{-1}(n)]$$

Choose a compact subspace D_B of R^p and a compact subspace $\tilde{D}_B$ of R^{p+p^2}. Introduce

$$f(x) = G^T(x)R^{-1}(g(x^*) - g(x))$$

$$V(x, e(n)) = G^T(x)R^{-1} e(n)$$

$$H(x) = \mathrm{diag}[G^T(x)R^{-1} G(x)]$$

and

$$Q(x, e(n)) = f(x) + V(x, e(n))$$

From (8)

$$\hat{x}(n+1) = \hat{x}(n) + \frac{1}{n} S(n) Q(\hat{x}(n), e(n))$$

$$S^{-1}(n+1) = S^{-1}(n) + \frac{1}{n}[H(\hat{x}(n))-S^{-1}(n)]$$

Let

$$z_n = (\hat{x}(n), \mathrm{col}\ S^{-1}(n))$$

Then

$$z_{n+1} = z_n + \frac{1}{n} \tilde{Q}_n(z_n, e(n))$$

where the p first elements of $\tilde{Q}$ are

$$[\tilde{Q}_n(z_n, e(n))]^{1,\dots p} = S(n) Q(\hat{x}(n), e(n))$$

and

$$[\tilde{Q}_n(z_n, e(n)]^{p+1,\dots,p+p^2} = \mathrm{col}[H(\hat{x}(n))-S^{-1}(n)]\cdot\frac{n}{n+1}$$

Since g(x) is twice continous differentiable

$$|G'(x)| < K \qquad x \in D_B$$

Hence $\tilde{Q}(z, e(n))$ is Lipschitz continuous in $\tilde{D}_B$ with Lipschitz constant $K|e(n)|$. Theorem 1 in Ljung [10] can thus be applied. The noise conditions are clearly satisfied, since

$$\zeta_{n+1} = \zeta_n + \frac{1}{n}[\tilde{Q}_n(z^o, e(n)) - \zeta_n]$$

where

$$z^o = (x^o, \mathrm{col}[S^o)^{-1}])$$

and

$$\rho_{n+1} = \rho_n + \frac{1}{n}[K|e_n| - \rho_n]$$

converge w.p.1 to

$$\lim_{n\to\infty} E\, \tilde{Q}_n(z^o, e(n)) = \begin{bmatrix} S^o\ f(x^o) \\ \mathrm{col}[H(x^o) - (S^o)^{-1}] \end{bmatrix}$$

and

$E\, K|e_n|$, respectively

according to the results of Section 4 in Ljung [10] and (2).

The convergence of (6) now relies upon the stability of the ordinary differential equation (ODE).

$$\dot{z} = \lim_{n\to\infty} E\, \tilde{Q}_n(z, e(n))$$

or

$$\dot{x} = S\, f(x)$$

$$\dot{S}^{-1} = H(x) - S^{-1}$$

Choose as Lyapunov function

$$V(x, S^{-1}) = [g(x) - g(x^*)]^T R^{-1}[g(x) - g(x^*)]$$

Then

$$\dot{V}(x, S^{-1}) = 2[g(x) - g(x^*)]^T R^{-1} G(x)\dot{x} =$$

$$= 2[g(x) - g(x^*)]^T R^{-1} G(x)S\, G^T(x)\, R^{-1}[g(x) - g(x^*)]$$

But, since S is positive definite by necessity, $\dot{V}(x, S^{-1}) \leq 0$ and $\dot{V}(x, S^{-1}) = 0 \Rightarrow x \in D_c$. It now follows from the corrolary of Theorem 1 in Ljung [10], that $\hat{x}(n) \rightarrow D_c$ w.p.1 as $n\rightarrow\infty$

□

Remark.

The conclusion of the theorem can be sharpened. It is stated that x(n) will converge either to infinity or to a stationary point of the ODE (10). In fact, as shown by Gustavsson, Ljung and Söderström in [11], $\hat{x}(n)$ can converge only to stable stationary points of (10). Hence, the unstable points can be excluded from the set D_c. To do so linearize (10a) around $x_o \in D_c$:

$$\Delta\dot{x} = [H(x_o)]^{-1} \frac{d}{dx} f(x)\Big|_{x=x_o} \Delta x \qquad (12)$$

where $\Delta x = x - x_o$. To find $\frac{d}{dx} f(x)$ introduce the notation

$r_{k\ell}$ = the $k\ell$ element of R^{-1}

and the matrix

$(a_{ij}) = A(x)$

where

$$a_{ij} = \sum_{k,\ell=1}^{r} \frac{\partial g^{(k)}(x)}{\partial x_j \, \partial x_i} r_{k,\ell}(g^{(\ell)}(x) - g^{(\ell)}(x^*)) \tag{13}$$

Then, after some calculation, it is found that

$$\frac{d}{dx} f(x) = -H(x) - A$$

Consequently, the stability properties of (8) are determined by the matrix

$$B(x) = -I - [H(x)]^{-1} A(x)$$

Introduce the set

$$D_s = \{x \mid B(x) \text{ has all eigenvalues in the left hand plane}\}.$$

Then the conclusions of Theorem 1 can be sharpened to

$$\hat{x}(n) \to D_c \cap D_s \quad \text{w.p.1} \quad \text{as } n \to \infty. \tag{14}$$

□

The sequential method (7) can be treated in a simular way. It follows from (7) that with the notations introduced in the proof above

$$z_{n+1}^{(1)} = z_n^{(1)} = \frac{1}{n} \bar{Q}(z_r^{(1)}, \ldots, z_n^{(r)}; e(n))$$

It is easily seen that

$$\bar{Q}(z_n^{(1)}, \ldots, z_n^{(1)}; e(n)) = \tilde{Q}(z_n^{(1)}; e(n))$$

and that

$$|\bar{Q}(z_n^{(1)}, \ldots, z_n^{(r)}; e(n)) - \tilde{Q}(z_n^{(1)}, e(n))| \leq$$

$$\leq K \max_{1<i\leq r} |z_n^{(i)} - z_n^{(1)}| \, |e(n)|$$

But if the estimates $z_n^{(i)}$ do not tend to infinity then

$$\max_{1<i\leq r} |z_n^{(i)} - z_n^{(1)}| \to 0$$

and consequently the sequence $\{z_n\}$ produced by (6) and the subsequences $\{z_n^{(1)}\}$ produced by (7) will have the same asymptotic properties. Hence the convergence result, Theorem 1, is valid also for the sequential method (7).

4. Concluding Remarks.

Theorem 1 states to which points the state estimate will converge. It is naturally desirable that $D_c = \{x^*\}$. To obtain this, $G(x)[g(x) - g(x^*)] = 0$ must imply $x = x^*$. This is a condition on the measurement vector. By selecting a suitable and sufficiently large set of measurements at each timepoint n, this condition can be satisfied. If we use the input data to a conventional load flow, we know that rank $G(x) = p - 1$. By adding the measurements of the slack bus voltage the rank can be increased to p.

The point $\{\infty\}$ in the set D_c poses no practical problem if $g(x)g(x)^T$ tends to infinity as $|x|$ tends to infinity. In any implementation the estimates x_n will not be allowed to wander off to infinity. One straightforward and simple-minded way to exclude the point $\{\infty\}$ from D_c is to restart the algorithm in $\bar{x}$ if $|x_n|$ is unrealistically large, where $\bar{x}$ denotes the solution of a load flow.

5. References.

[1] van Overbeek, A. J. M., "State Estimation in Power Networks I - A Literature Survey", Report 7331(C), November 1973, Lund Institute of Technology, Division of Automatic Control.

[2] Schweppe, F. C., "Power System Static-State Estimation, Part I: Exact Model", IEEE Transactions on Power Apparatus and Systems, Vol. PAS-89, No. 1, January 1970, pp 120-125.

[3] Schweppe, F. C. and Rom, D. B., "Power System Static-State Estimation, Part II: Approximate Model", IEEE Transactions on Power Apparatus and Systems, Vol. PAS-89, No. 1, January 1970, pp 125-130.

[4] Schweppe, F. C., "Power System Static-State Estimation, Part III: Implementation", IEEE Transactions on Power Apparatus and Systems, Vol. PAS-89, No. 1, January 1970, pp 130-135.

[5] Masiello, R. D. and Schweppe, F. C., "A Tracking Static-State Estimator", IEEE Transactions on Power Apparatus and Systems, Vol. PAS-90, No. 3, May/June 1971, pp 1025-1033.

[6] Larson, R. E., Tinney, W. F. and Peschon, J., "State Estimation in Power Systems, Part I" Theory and Feasibility", IEEE Transactions on Power Apparatus and Systems, Vol. PAS-89, No. 3, March 1970, pp 345-352.

[7] Larson, R. E., Tinney, E. F., Hajdu, L. P. and Piercy, D.S., "State Estimation in Power Systems, Part II: Implementation and Applications", IEEE Transactions on Power Apparatus and Systems, Vol. PAS-89, No. 3, March 1970, pp 353-363.

[8] Debs, A. E. and Larson, R. E., "A Dynamic Estimator for Tracking the State of a Power System", IEEE Transactions on Power Apparatus and Systems, Vol. PAS-89, No. 7, September/October 1970, pp 1670-1676.

[9] Debs, A. E., Larson, R. E. and Hajdu, L. P., "On-Line Sequential State Estimation for Power Systems", Proceedings of the Fourth Power Systems Computation Conference, Vol. 3, Department of Electrical and Electronic Engineering, Queen Mary College, University of London, London, 1973, paper 3.3/3.

[10] Ljung, L., "Convergence of Recursive Stochastic Algorithms", Preprints of the Stochastic Control Symposium, IFAC, Budapest, 1974.

[11] Gustavsson, I., Ljung, L. and Söderström, T., "A comparative Study of recursive Identification Methods", Forthcoming report, Lund Institute of Technology, Division of Automatic Control, Lund, Sweden.