

ECEN 615

Methods of Electric Power Systems Analysis

Lecture 17: Least Squares, State Estimation

Prof. Tom Overbye

Dept. of Electrical and Computer Engineering

Texas A&M University

overbye@tamu.edu



TEXAS A&M
UNIVERSITY

Announcements



- Read Chapter 9 from the book
- Homework 4 is due on Thursday October 31.

Least Squares Solution



- We write $(\mathbf{a}^i)^T$ for the row i of \mathbf{A} and \mathbf{a}^i is a column vector
- Here, $m \geq n$ and the solution we are seeking is that which minimizes $\|\mathbf{Ax} - \mathbf{b}\|_p$, where p denotes some norm
- Since usually an overdetermined system has no exact solution, the best we can do is determine an \mathbf{x} that minimizes the desired norm.

Choice of p



- We discuss the choice of p in terms of a specific example
- Consider the equation $\mathbf{Ax} = \mathbf{b}$ with

$$\mathbf{A} = \begin{bmatrix} \mathbf{1} \\ \mathbf{1} \\ \mathbf{1} \end{bmatrix} \quad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} \quad \text{with } b_1 \geq b_2 \geq b_3 \geq 0$$

(hence three equations and one unknown)

- We consider three possible choices for p :

Choice of p



(i) $p = 1$

$\|\mathbf{Ax} - \mathbf{b}\|_1$ is minimized by $x^* = b_2$

(ii) $p = 2$

$\|\mathbf{Ax} - \mathbf{b}\|_2$ is minimized by $x^* = \frac{b_1 + b_2 + b_3}{3}$

(iii) $p = \infty$

$\|\mathbf{Ax} - \mathbf{b}\|_\infty$ is minimized by $x^* = \frac{b_1 + b_3}{2}$

The Least Squares Problem



- In general, $\|\mathbf{Ax} - \mathbf{b}\|_p$ is not differentiable for $p = 1$ or $p = \infty$
- The choice of $p = 2$ (Euclidean norm) has become well established given its least-squares fit interpretation
- The problem $\min_{\mathbf{x} \in \mathbb{R}^n} \|\mathbf{Ax} - \mathbf{b}\|_2$ is tractable for 2 major reasons
 - First, the function is differentiable

$$\phi(\mathbf{x}) = \frac{1}{2} \|\mathbf{Ax} - \mathbf{b}\|_2^2 = \frac{1}{2} \sum_{i=1}^m \left[(\mathbf{a}^i)^T \mathbf{x} - b_i \right]^2$$

The Least Squares Problem, cont.



- Second, the Euclidean norm is preserved under orthogonal transformations:

$$\|(\mathbf{Q}^T \mathbf{A})\mathbf{x} - \mathbf{Q}^T \mathbf{b}\|_2 = \|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2$$

with \mathbf{Q} an arbitrary orthogonal matrix; that is, \mathbf{Q} satisfies

$$\mathbf{Q}\mathbf{Q}^T = \mathbf{Q}^T\mathbf{Q} = \mathbf{I} \quad \mathbf{Q} \in \mathbb{R}^{n \times n}$$

The Least Squares Problem, cont.



- We introduce next the basic underlying assumption: \mathbf{A} is full rank, i.e., the columns of \mathbf{A} constitute a set of linearly independent vectors
- This assumption implies that the rank of \mathbf{A} is n because $n \leq m$ since we are dealing with an overdetermined system
- Fact: The least squares solution \mathbf{x}^* satisfies

$$\mathbf{A}^T \mathbf{A} \mathbf{x}^* = \mathbf{A}^T \mathbf{b}$$

Proof of Fact



- Since by definition the least squares solution \mathbf{x}^* minimizes $\phi(\bullet)$ at the optimum, the derivative of this function zero:

$$\phi(\mathbf{x}) = \frac{1}{2} \|\mathbf{Ax} - \mathbf{b}\|_2^2 = \frac{1}{2} (\mathbf{x}^T \mathbf{A}^T \mathbf{Ax} - \mathbf{x}^T \mathbf{A}^T \mathbf{b} - \mathbf{b}^T \mathbf{A} \mathbf{x} + \mathbf{b}^T \mathbf{b})$$

$$\mathbf{0} = \left. \frac{\partial \phi(\mathbf{x})}{\partial \mathbf{x}} \right|_{\mathbf{x}^*} = \left. \frac{\partial}{\partial \mathbf{x}} \left\{ \frac{1}{2} (\mathbf{x}^T \mathbf{A}^T \mathbf{Ax} - \mathbf{x}^T \mathbf{A}^T \mathbf{b} - \mathbf{b}^T \mathbf{A} \mathbf{x} + \mathbf{b}^T \mathbf{b}) \right\} \right|_{\mathbf{x}^*}$$

$$= \left. \frac{\partial}{\partial \mathbf{x}} \left\{ \frac{1}{2} (\mathbf{x}^T \mathbf{A}^T \mathbf{Ax} - 2\mathbf{x}^T \mathbf{A}^T \mathbf{b} + \mathbf{b}^T \mathbf{b}) \right\} \right|_{\mathbf{x}^*}$$

$$= \mathbf{A}^T \mathbf{Ax}^* - \mathbf{A}^T \mathbf{b}$$

Implications



- This underlying assumption implies that \mathbf{A} is full rank $\Leftrightarrow \exists \mathbf{x} \neq \mathbf{0} \ni \mathbf{A}\mathbf{x} \neq \mathbf{0}$
- Therefore, the fact that $\mathbf{A}^T\mathbf{A}$ is positive definite (*p.d.*) follows from considering any $\mathbf{x} \neq \mathbf{0}$ and evaluating

$$\mathbf{x}^T \mathbf{A}^T \mathbf{A} \mathbf{x} = \|\mathbf{A} \mathbf{x}\|_2^2 > 0,$$

which is the definition of a *p.d.* matrix

- We use the shorthand $\mathbf{A}^T\mathbf{A} > \mathbf{0}$ for $\mathbf{A}^T\mathbf{A}$ being a symmetric, positive definite matrix

Implications



- The underlying assumption that \mathbf{A} is full rank and therefore $\mathbf{A}^T\mathbf{A}$ is *p.d.* implies that there exists a unique least squares solution
- Note: we use the inverse in a conceptual, rather than a computational, sense

$$\mathbf{x}^* = (\mathbf{A}^T\mathbf{A})^{-1} \mathbf{A}^T\mathbf{b}$$

- The below formulation is known as the normal equations, with the solution conceptually straightforward

$$(\mathbf{A}^T\mathbf{A})\mathbf{x} = \mathbf{A}^T\mathbf{b}$$

Example: Curve Fitting



- Say we wish to fit five points to a polynomial curve of the form

$$f(t, \mathbf{x}) = x_1 + x_2 t + x_3 t^2$$

- This can be written as

$$\mathbf{Ax} = \mathbf{y} = \begin{bmatrix} 1 & t_1 & t_1^2 \\ 1 & t_2 & t_2^2 \\ 1 & t_3 & t_3^2 \\ 1 & t_4 & t_4^2 \\ 1 & t_5 & t_5^2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \end{bmatrix}$$

Example: Curve Fitting



- Say the points are $\mathbf{t} = [0, 1, 2, 3, 4]$ and $\mathbf{y} = [0, 2, 4, 5, 4]$.

Then

$$\mathbf{Ax} = \mathbf{y} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \\ 1 & 4 & 16 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \\ 4 \\ 5 \\ 4 \end{bmatrix}$$

$$\mathbf{x}^* = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{b} = \begin{bmatrix} 0.886 & 0.257 & -0.086 & -0.143 & 0.086 \\ -0.771 & 0.186 & 0.571 & 0.386 & -0.371 \\ 0.143 & -0.071 & -0.143 & -0.071 & 0.143 \end{bmatrix} \begin{bmatrix} 0 \\ 2 \\ 4 \\ 5 \\ 4 \end{bmatrix}$$

$$\mathbf{x}^* = \begin{bmatrix} -0.2 \\ 3.1 \\ -0.5 \end{bmatrix}$$

Implications



- An important implication of positive definiteness is that we can factor $\mathbf{A}^T\mathbf{A}$ since $\mathbf{A}^T\mathbf{A} > \mathbf{0}$

$$\mathbf{A}^T\mathbf{A} = \mathbf{U}^T\mathbf{D}\mathbf{U} = \mathbf{U}^T\mathbf{D}^{1/2}\mathbf{D}^{1/2}\mathbf{U} = \mathbf{G}^T\mathbf{G}$$

- The expression $\mathbf{A}^T\mathbf{A} = \mathbf{G}^T\mathbf{G}$ is called the Cholesky factorization of the symmetric positive definite matrix $\mathbf{A}^T\mathbf{A}$

A Least Squares Solution Algorithm



Step 1: Compute the lower triangular part of $\mathbf{A}^T \mathbf{A}$

Step 2: Obtain the Cholesky Factorization $\mathbf{A}^T \mathbf{A} = \mathbf{G}^T \mathbf{G}$

Step 3: Compute $\mathbf{A}^T \mathbf{b} = \hat{\mathbf{b}}$

Step 4: Solve for \mathbf{y} using forward substitution in

$$\mathbf{G}^T \mathbf{y} = \hat{\mathbf{b}}$$

and for \mathbf{x} using backward substitution in

$$\mathbf{G} \mathbf{x} = \mathbf{y}$$

Note, our standard LU factorization approach would work; we can just solve it twice as fast by taking advantage of it being a symmetric matrix

Practical Considerations



- The two key problems that arise in practice with the triangularization procedure are:
 - First, while \mathbf{A} maybe sparse, $\mathbf{A}^T\mathbf{A}$ is much less sparse and consequently requires more computing resources for the solution
 - In particular, with $\mathbf{A}^T\mathbf{A}$ second neighbors are now connected! Large networks are still sparse, just not as sparse
 - Second, $\mathbf{A}^T\mathbf{A}$ may actually be numerically less well-conditioned than \mathbf{A}

Loss of Sparsity Example



- Assume the \mathbf{B} matrix for a network is

$$\mathbf{B} = \begin{bmatrix} -1 & 1 & 0 & 0 \\ 1 & -2 & 1 & 0 \\ 0 & 1 & -2 & 1 \\ 0 & 0 & 1 & -1 \end{bmatrix}$$

- Then $\mathbf{B}^T\mathbf{B}$ is $\mathbf{B}^T\mathbf{B} = \begin{bmatrix} 2 & -3 & 1 & 0 \\ -3 & 6 & -4 & 1 \\ 1 & -4 & 6 & -3 \\ 0 & 1 & -3 & 2 \end{bmatrix}$

- Second neighbors are now connected!

Numerical Conditioning



- To understand the point on numerical ill-conditioning, we need to introduce terminology
- We define the norm of a matrix $\mathbf{B} \in \mathbb{R}^{m \times n}$ to be

$$\|\mathbf{B}\| = \max_{\mathbf{x} \neq \mathbf{0}} \left\{ \frac{\|\mathbf{B}\mathbf{x}\|}{\|\mathbf{x}\|} \right\}$$

= *maximum stretching of the matrix* \mathbf{B}

- This is the maximum singular value of \mathbf{B}

Numerical Conditioning Example



- Say we have the matrix

$$\mathbf{B} = \begin{bmatrix} 10 & 0 \\ 0 & 0.1 \end{bmatrix}$$

- What value of \mathbf{x} with a norm of 1 that maximizes $\|\mathbf{B}\mathbf{x}\|$?
- What value of \mathbf{x} with a norm of 1 that minimizes $\|\mathbf{B}\mathbf{x}\|$?

$$\|\mathbf{B}\| = \max_{\mathbf{x} \neq \mathbf{0}} \left\{ \frac{\|\mathbf{B}\mathbf{x}\|}{\|\mathbf{x}\|} \right\}$$

= *maximum stretching of the matrix B*

Numerical Conditioning



$$= \max_i \left\{ \sqrt{\lambda_i}, \lambda_i \text{ is an eigenvalue of } \underline{\mathbf{B}}^T \underline{\mathbf{B}} \right\},$$

i.e., λ_i is a root of the polynomial

$$p(\lambda) = \det[\mathbf{B}^T \mathbf{B} - \lambda \mathbf{I}]$$

Keep in mind the eigenvalues of a p.d. matrix are positive

- In other words, the ℓ_2 norm of \mathbf{B} is the square root of the largest eigenvalue of $\mathbf{B}^T \mathbf{B}$

Numerical Conditioning



- The conditioning number of a matrix \mathbf{B} is defined as

$$\kappa(\mathbf{B}) = \left\| \mathbf{B} \right\| \left\| \mathbf{B}^{-1} \right\| = \frac{\left| \sigma_{max}(\mathbf{B}) \right|}{\left| \sigma_{min}(\mathbf{B}) \right|} \left. \begin{array}{l} \text{the max / min stretching} \\ \text{ratio of the matrix } \mathbf{B} \end{array} \right\}$$

- A well-conditioned matrix has a small value of $\kappa(\mathbf{B})$, close to 1; the larger the value of $\kappa(\mathbf{B})$, the more pronounced is the ill-conditioning

Power System State Estimation



- Overall goal is to come up with a power flow model for the present "state" of the power system based on the actual system measurements
- SE assumes the topology and parameters of the transmission network are mostly known
- Measurements come from SCADA, and increasingly, from PMUs

Power System State Estimation



- Problem can be formulated in a nonlinear, weighted least squares form as

$$\min J(\mathbf{x}) = \sum_{i=1}^m \frac{[z_i - f_i(\mathbf{x})]^2}{\sigma_i^2}$$

where $J(\mathbf{x})$ is the scalar cost function, \mathbf{x} are the state variables (primarily bus voltage magnitudes and angles), z_i are the m measurements, $\mathbf{f}(\mathbf{x})$ relates the states to the measurements and σ_i is the assumed standard deviation for each measurement

Assumed Error



- Hence the goal is to decrease the error between the measurements and the assumed model states \mathbf{x}
- The σ_i term weighs the various measurements, recognizing that they can have vastly different assumed errors

$$\min J(\mathbf{x}) = \sum_{i=1}^m \frac{[z_i - f_i(\mathbf{x})]^2}{\sigma_i^2}$$

- Measurement error is assumed Gaussian (whether it is or not is another question); outliers (bad measurements) are often removed

State Estimation for Linear Functions



- First we'll consider the linear problem. That is where

$$\mathbf{z}^{meas} - \mathbf{f}(\mathbf{x}) = \mathbf{z}^{meas} - \mathbf{H}\mathbf{x}$$

- Let \mathbf{R} be defined as the diagonal matrix of the variances (square of the standard deviations) for each of the measurements

$$\mathbf{R} = \begin{bmatrix} \sigma_1^2 & 0 & \cdots & 0 \\ 0 & \sigma_2^2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \sigma_m^2 \end{bmatrix}$$

State Estimation for Linear Functions



- We then differentiate $J(\mathbf{x})$ w.r.t. \mathbf{x} to determine the value of \mathbf{x} that minimizes this function

$$J(\mathbf{x}) = \left[\mathbf{z}^{meas} - \mathbf{H}\mathbf{x} \right]^T \mathbf{R}^{-1} \left[\mathbf{z}^{meas} - \mathbf{H}\mathbf{x} \right]$$

$$\nabla J(\mathbf{x}) = -2\mathbf{H}^T \mathbf{R}^{-1} \mathbf{z}^{meas} + 2\mathbf{H}^T \mathbf{R}^{-1} \mathbf{H}\mathbf{x}$$

At the minimum we have $\nabla J(\mathbf{x}) = \mathbf{0}$. So solving for \mathbf{x} gives

$$\mathbf{x} = \left[\mathbf{H}^T \mathbf{R}^{-1} \mathbf{H} \right]^{-1} \mathbf{H}^T \mathbf{R}^{-1} \mathbf{z}^{meas}$$

Simple DC System Example



- Say we have a two bus power system that we are solving using the dc approximation. Say the line's per unit reactance is $j0.1$. Say we have power measurements at both ends of the line. For simplicity assume $\mathbf{R}=\mathbf{I}$. We would then like to estimate the bus angles. Then

$$z_1 = P_{12} = \frac{\theta_1 - \theta_2}{0.1} = 2.2, \quad z_2 = -2.0 = P_{21} = \frac{\theta_2 - \theta_1}{0.1}$$

$$\mathbf{x} = \begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix}, \mathbf{H} = \begin{bmatrix} 10 & -10 \\ -10 & 10 \end{bmatrix}, \mathbf{H}^T \mathbf{H} = \begin{bmatrix} 200 & -200 \\ -200 & 200 \end{bmatrix}$$

We have a problem since $\mathbf{H}^T \mathbf{H}$ is singular. This is because of lack of an angle reference.

Simple DC System Example, cont.



- Say we directly measure θ_1 (with a PMU) to be zero; set this as the third measurement. Then

$$z_1 = P_{12} = \frac{\theta_1 - \theta_2}{0.1} = 2.2, \quad z_2 = -2.0 = P_{21} = \frac{\theta_2 - \theta_1}{0.1}, \quad z_3 = 0$$

$$\mathbf{x} = \begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix}, \quad \mathbf{z} = \begin{bmatrix} 2.2 \\ -2 \\ 0 \end{bmatrix}, \quad \mathbf{H} = \begin{bmatrix} 10 & -10 \\ -10 & 10 \\ 1 & 0 \end{bmatrix}, \quad \mathbf{H}^T \mathbf{H} = \begin{bmatrix} 201 & -200 \\ -200 & 200 \end{bmatrix}$$

$$\mathbf{x} = \left[\mathbf{H}^T \mathbf{R}^{-1} \mathbf{H} \right]^{-1} \mathbf{H}^T \mathbf{R}^{-1} \mathbf{z}^{meas}$$

$$\mathbf{x} = \begin{bmatrix} 201 & -200 \\ -200 & 200 \end{bmatrix}^{-1} \begin{bmatrix} 10 & -10 & 1 \\ -10 & 10 & 0 \end{bmatrix} \begin{bmatrix} 2.2 \\ -2 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ -0.21 \end{bmatrix}$$

Note that the angles are in radians

Nonlinear Formulation



- A regular ac power system is nonlinear, so we need to use an iterative solution approach. This is similar to the Newton power flow. Here assume m measurements and n state variables (usually bus voltage magnitudes and angles) Then the Jacobian is the \mathbf{H} matrix

$$\mathbf{H}(\mathbf{x}) = \frac{\partial \mathbf{f}(\mathbf{x})}{\partial \mathbf{x}} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{bmatrix}$$

Measurement Example



- Assume we measure the real and reactive power flowing into one end of a transmission line; then the z_i - $f_i(\mathbf{x})$ functions for these two are

$$P_{ij}^{meas} = \left[-V_i^2 G_{ij} + V_i V_j \left(G_{ij} \cos(\theta_i - \theta_j) + B_{ij} \sin(\theta_i - \theta_j) \right) \right]$$

$$Q_{ij}^{meas} = \left[V_i^2 \left(B_{ij} + \frac{B_{cap}}{2} \right) + V_i V_j \left(G_{ij} \sin(\theta_i - \theta_j) - B_{ij} \cos(\theta_i - \theta_j) \right) \right]$$

- Two measurements for four unknowns
- Other measurements, such as the flow at the other end, and voltage magnitudes, add redundancy

SE Iterative Solution Algorithm



- We then make an initial guess of \mathbf{x} , $\mathbf{x}^{(0)}$ and iterate, calculating $\Delta\mathbf{x}$ each iteration

$$\Delta\mathbf{x} = \left[\mathbf{H}^T \mathbf{R}^{-1} \mathbf{H} \right]^{-1} \mathbf{H}^T \mathbf{R}^{-1} \begin{bmatrix} z_1 - f_1(\mathbf{x}) \\ \vdots \\ z_m - f_m(\mathbf{x}) \end{bmatrix}$$

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \Delta\mathbf{x}$$

Keep in mind that \mathbf{H} is no longer constant, but varies as \mathbf{x} changes. often ill-conditioned

This is exactly the least squares form developed earlier with $\mathbf{H}^T \mathbf{R}^{-1} \mathbf{H}$ an n by n matrix. This could be solved with Gaussian elimination, but this isn't preferred because the problem is often ill-conditioned

Nonlinear SE Solution Algorithm, Book Figure 9.11

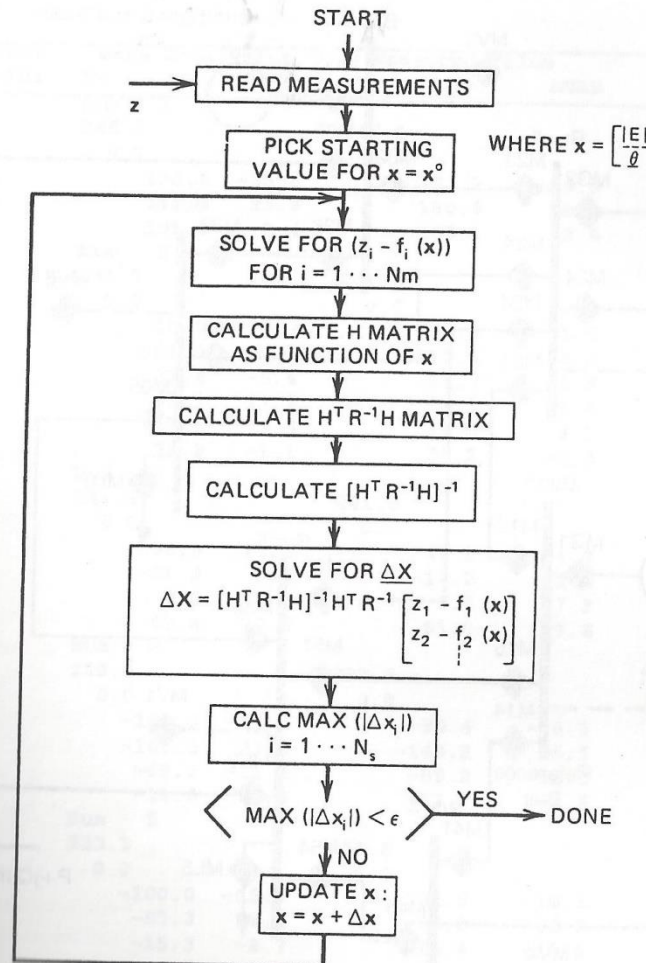


FIGURE 9.11 State estimation solution algorithm.

Example: Two Bus Case



- Assume a two bus case with a generator supplying a load through a single line with $x=0.1$ pu. Assume measurements of the p/q flow on both ends of the line (into line positive), and the voltage magnitude at both the generator and the load end. So $B_{12} = B_{21}=10.0$

$$P_{ij}^{meas} = \left[V_i V_j \left(B_{ij} \sin(\theta_i - \theta_j) \right) \right]$$

$$Q_{ij}^{meas} = \left[V_i^2 B_{ij} + V_i V_j \left(-B_{ij} \cos(\theta_i - \theta_j) \right) \right]$$

$$V_i^{meas} - V_i = 0$$

We need to assume a reference angle unless we directly measuring phase

Example: Two Bus Case



• Let $\mathbf{z}^{meas} = \begin{bmatrix} P_{12} \\ Q_{12} \\ P_{21} \\ Q_{21} \\ V_1 \\ V_2 \end{bmatrix} = \begin{bmatrix} 2.02 \\ 1.5 \\ -1.98 \\ -1 \\ 1.01 \\ 0.87 \end{bmatrix}$ $x^0 = \begin{bmatrix} V_1 \\ \theta_2 \\ V_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \sigma_i = 0.01$

We assume an angle reference of $\theta_1=0$

$$H(\mathbf{x}) = \begin{bmatrix} V_2 10 \sin(-\theta_2) & -V_1 V_2 10 \cos(-\theta_2) & V_1 10 \sin(-\theta_2) \\ 20V_1 - V_2 10 \cos(-\theta_2) & -V_1 V_2 10 \sin(-\theta_2) & -V_1 10 \cos(-\theta_2) \\ V_2 10 \sin(\theta_2) & V_1 V_2 10 \cos(\theta_2) & V_1 10 \sin(\theta_2) \\ -V_2 10 \cos(\theta_2) & V_1 V_2 10 \sin(\theta_2) & 20V_2 - V_1 10 \cos(\theta_2) \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Example: Two Bus Case



- With a flat start guess we get

$$H(\mathbf{x}^0) = \begin{bmatrix} 0 & -10 & 0 \\ 10 & 0 & -10 \\ 0 & 10 & 0 \\ -10 & 0 & 10 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \mathbf{z} - \mathbf{f}(\mathbf{x}^0) = \begin{bmatrix} 2.02 \\ 1.5 \\ -1.98 \\ -1 \\ 0.01 \\ -0.13 \end{bmatrix}$$

$$\mathbf{R} = \begin{bmatrix} 0.0001 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.0001 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.0001 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.0001 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.0001 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.0001 \end{bmatrix}$$

Example: Two Bus Case



$$\mathbf{H}^T \mathbf{R}^{-1} \mathbf{H} = 1e^6 \times \begin{bmatrix} 2.01 & 0 & -2 \\ 0 & 2 & 0 \\ -2 & 0 & 2.01 \end{bmatrix}$$

$$\mathbf{x}^1 = \mathbf{x}^0 + \left[\mathbf{H}^T \mathbf{R}^{-1} \mathbf{H} \right]^{-1} \mathbf{H}^T \mathbf{R}^{-1} \begin{bmatrix} 2.02 \\ 1.5 \\ -1.98 \\ -1 \\ 0.01 \\ -0.13 \end{bmatrix} = \begin{bmatrix} 1.003 \\ -0.2 \\ 0.8775 \end{bmatrix}$$

Assumed SE Measurement Accuracy

- The assumed measurement standard deviations can have a significant impact on the resultant solution, or even whether the SE converges
- The assumption is a Gaussian (normal) distribution of the error with no bias

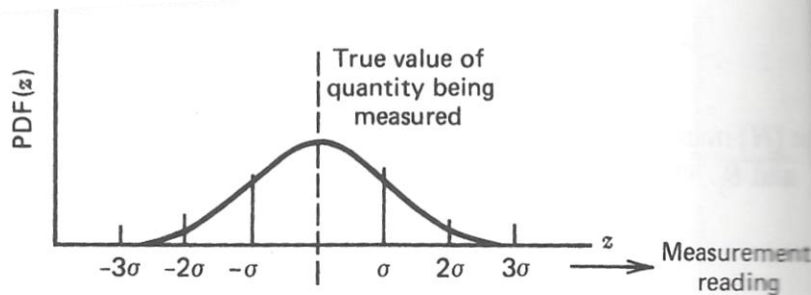


FIGURE 9.8 Normal distribution of meter errors.

SE Observability



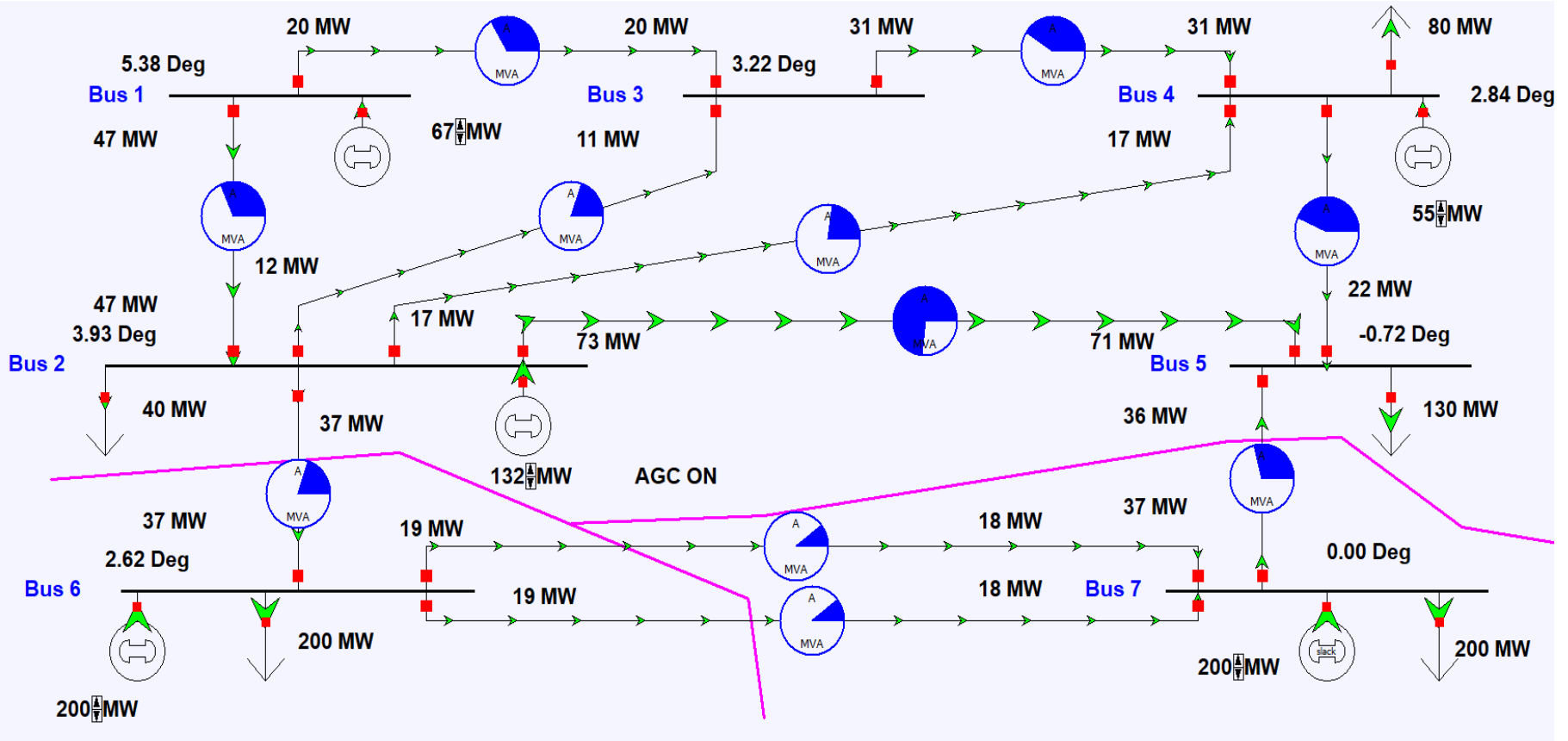
- In order to estimate all n states we need at least n measurements. However, where the measurements are located is also important, a topic known as observability
 - In order for a power system to be fully observable usually we need to have a measurement available no more than one bus away
 - At buses we need to have at least measurements on all the injections into the bus except one (including loads and gens)
 - Loads are usually flows on feeders, or the flow into a transmission to distribution transformer
 - Generators are usually just injections from the GSU

Pseudo Measurements



- Pseudo measurements are used at buses in which there is no load or generation; that is, the net injection into the bus is known with high accuracy to be zero
 - In order to enforce the net power balance at a bus we need to include an explicit net injection measurement
- To increase observability sometimes estimated values are used for loads, shunts and generator outputs
 - These “measurements” are represented as having a higher standard deviation

SE Observability Example



SE Bad Data Detection



- The quality of the measurements available to an SE can vary widely, and sometimes the SE model itself is wrong. Causes include
 - Modeling Errors: perhaps the assumed system topology is incorrect, or the assumed parameters for a transmission line or transformer could be wrong
 - Data Errors: measurements may be incorrect because of incorrect data specifications, like the CT ratios or even flipped positive and negative directions
 - Transducer Errors: the transducers may be failing or may have bias errors
 - Sampling Errors: SCADA does not read all values simultaneously and power systems are dynamic

SE Bad Data Detection

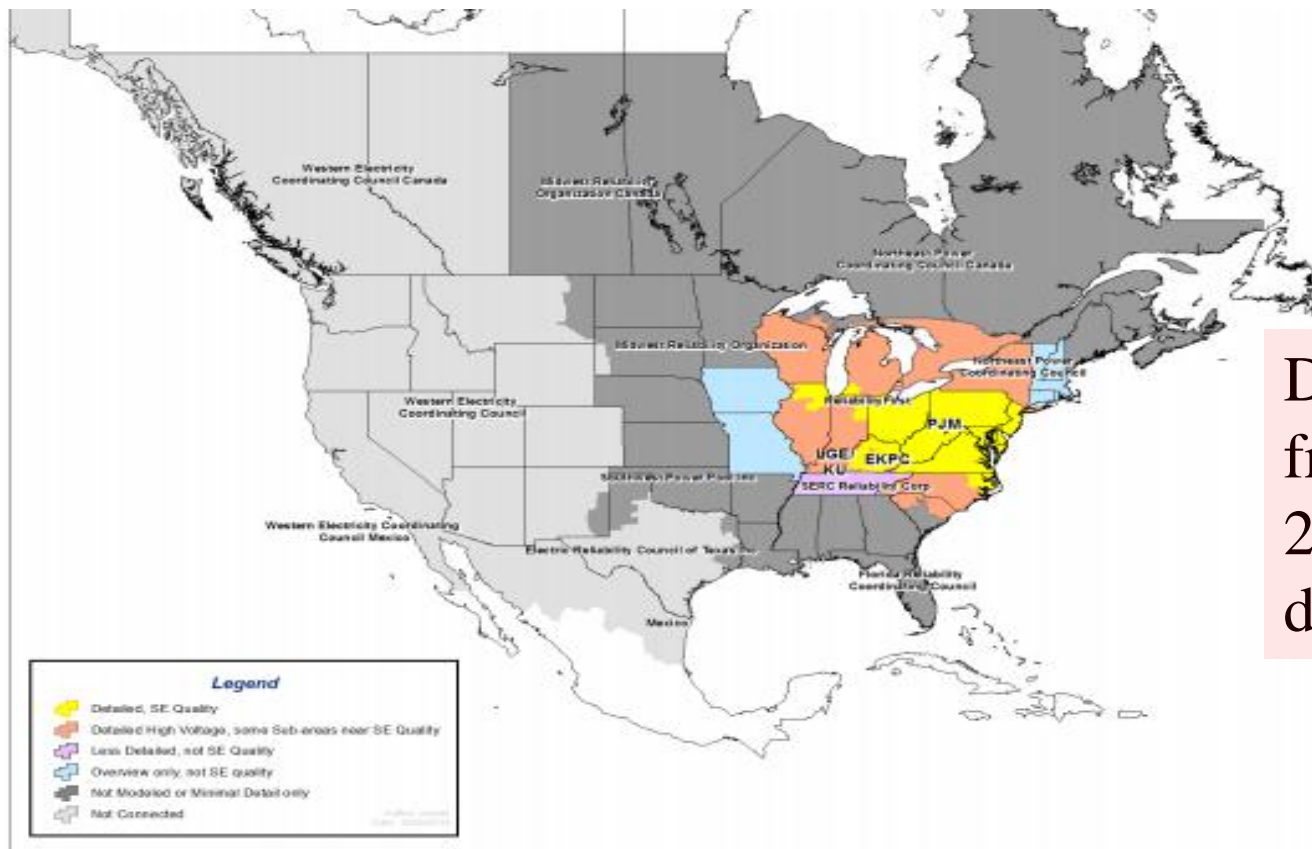


- The challenge for SE is to determine when there is likely a bad measurement (or multiple ones), and then to determine the particular bad measurements
- $J(\mathbf{x})$ is random number, with a probability density function (PDF) known as a chi-squared distribution, $\chi^2(K)$, where K is the degrees of freedom, $K=m-n$
- It can be shown the expected mean for $J(\mathbf{x})$ is K , with a standard deviation of $\sqrt{2K}$
 - Values of $J(\mathbf{x})$ outside of several standard deviations indicate possible bad measurements, with the measurement residuals used to track down the likely bad measurements
- SE can be re-run without the bad measurements

Example SE Application: PJM and MISO



- PJM provides information about their EMS model in
 - www.pjm.com/-/media/documents/manuals/m03a.ashx



Data here is from the Sept 2018 (Rev 16) document

Exhibit 4: PJM EMS Model Details

Example SE Application: PJM and MISO



- PJM measurements are required for 69 kV and up
- PJM SE is triggered to execute every minute
- PJM SE solves well over 98% of the time
- Below reference provides info on MISO SE from March 2015
 - 54,433 buses
 - 54,415 network branches
 - 6332 generating units
 - 228,673 circuit breakers
 - 289,491 mapped points

<https://www.naspi.org/sites/default/files/2017-05/3a%20MISO-NASPIWokshop-Synchrophasor%20Data%20and%20State%20Estimation.pdf>