



Fabiano Schmidt

Estimação de Estado Regularizada para Sistemas de Energia Elétrica

Campinas
2013



UNIVERSIDADE ESTADUAL DE CAMPINAS
FACULDADE DE ENGENHARIA ELÉTRICA E DE COMPUTAÇÃO

Fabiano Schmidt

Estimação de Estado Regularizada para Sistemas de Energia Elétrica

Orientador: Prof. Dr. Madson Cortes de Almeida

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Engenharia Elétrica da Faculdade de Engenharia Elétrica e de Computação da Universidade Estadual de Campinas para a obtenção do título de Mestre em Engenharia Elétrica. Área de concentração: Energia Elétrica.

ESTE EXEMPLAR CORRESPONDE À VERSÃO FINAL DA DISSERTAÇÃO DEFENDIDA PELO ALUNO FABIANO SCHMIDT E ORIENTADA PELO PROF. DR. MADSON CORTES DE ALMEIDA.

Campinas
2013

Ficha catalográfica
Universidade Estadual de Campinas
Biblioteca da Área de Engenharia e Arquitetura
Rose Meire da Silva - CRB 8/5974

Schmidt, Fabiano, 1989-
Ach52e Estimação de estado regularizada para sistemas de energia elétrica / Fabiano Schmidt. – Campinas, SP : [s.n.], 2013.

Orientador: Madson Cortes de Almeida.
Dissertação (mestrado) – Universidade Estadual de Campinas, Faculdade de Engenharia Elétrica e de Computação.

1. Sistemas de energia elétrica. 2. Mínimos quadrados. 3. Teoria da estimativa. 4. Estimativas a-priori. I. Almeida, Madson Cortes da, 1973-. II. Universidade Estadual de Campinas. Faculdade de Engenharia Elétrica e de Computação. III. Título.

Informações para Biblioteca Digital

Título em outro idioma: Regularized state estimation for electric power systems

Palavras-chave em inglês:

Power systems

Least squares

Estimation theory

A priori estimates

Área de concentração: Energia Elétrica

Titulação: Mestre em Engenharia Elétrica

Banca examinadora:

Madson Cortes de Almeida [Orientador]

José Roberto Sanches Mantovani

Carlos Alberto de Castro Júnior

Data de defesa: 12-07-2013

Programa de Pós-Graduação: Engenharia Elétrica

COMISSÃO JULGADORA - TESE DE MESTRADO

Candidato: Fabiano Schmidt

Data da Defesa: 12 de julho de 2013

Título da Tese: "Estimação de Estado Regularizada para Sistemas de Energia Elétrica"

Prof. Dr. Madson Cortes de Almeida (Presidente):

Prof. Dr. José Roberto Sanches Mantovani:

Prof. Dr. Carlos Alberto de Castro Júnior:

Dedico este trabalho à minha mãe, Anadir Rondora Barbosa.

AGRADECIMENTOS

Machado de Assis certa vez publicou em um jornal do Rio de Janeiro a seguinte crítica: “Não é desprezo pelo que é nosso, não é desdém pelo meu país. O país real, esse é bom, revela os melhores instintos; mas o país oficial, esse é caricato e burlesco”. Talvez eu até concorde com Ariano Suassuna quando ele diz que Machado estava movido por uma justa indignação quando classificou o nosso país como burlesco, também concordo com Ariano ao dizer que se nós não tomarmos cuidado corremos sério risco de nos tornarmos burlescos de fato. Eu, como um integrante do país oficial, gostaria de desculpar-me com o país real, tendo em vista que este trabalho não atinge as reais necessidades desse povo, mas ao mesmo tempo, gostaria de agradecê-los, já que esse povo honesto e trabalhador deu-me o privilégio de estudar numa instituição pública de qualidade.

Agradeço aos meus pais, Anadir Rondora Barbosa e Flávio Schmidt, ao meu irmão, Flávio Schmidt Júnior e ao amigo e padrasto Antônio Alexandre Tacca pela confiança, pelo incondicional apoio durante o desenvolvimento deste trabalho e por sempre apoiarem as minhas decisões.

Agradeço ao professor Madson Cortes de Almeida pela oportunidade, por confiar no meu trabalho, por sempre me atender de maneira gentil e pela orientação necessária para a conclusão deste trabalho. Também gostaria de agradecer aos professores do DSEE por sempre estarem prontamente disponíveis em momentos de dúvidas.

Agradeço aos amigos Jakson e Rubia pela acolhida, pelas conversas e por compartilharem dos bons e maus momentos que vivenciei nesse período.

Agradeço a todos os amigos do DSEE, pelo apoio, companhia, discussões, ajudas e pelas conversas descontraídas durante os intervalos.

Agradeço aos amigos do curso de engenharia elétrica da UFMT, por apoiarem a minha decisão de cursar o mestrado.

Agradeço ao Fundo de Apoio ao Ensino, à Pesquisa e à Extensão (FAEPEX) pelo apoio financeiro.

Agradeço à UNICAMP pelo título e por todos os benefícios que essa instituição proporciona aos seus estudantes.

“If you torture the data long enough, it will confess.”

Ronald Coase

RESUMO

A estimação de estado em sistemas de energia elétrica pode ser descrita como um conjunto de funções que visam fornecer o modelo em tempo real do sistema elétrico. Fazem parte do modelo, a topologia da rede, seus parâmetros elétricos e as variáveis de estado, frequentemente definidas como os fasores de tensões nas barras do sistema. Na obtenção das variáveis de estado, embora os sistemas de medição possam ser projetados para que a rede seja sempre observável, eventualmente, falhas de comunicação, mudanças topológicas ou falhas em medidores podem tornar a rede temporariamente não observável. Nessas situações, as equações normais do modelo de estimação por mínimos quadrados tornam-se mal postas inviabilizando a estimação do estado. Convencionalmente, esse problema é contornado através da restauração da observabilidade, onde pseudomedidas são adicionadas à rede tornando-a novamente observável. Esta dissertação trata da estimação de estado regularizada, que é um método alternativo para lidar com os problemas associados a não observabilidade temporária da rede. Com o estimador regularizado é possível evitar a etapa de restauração da observabilidade, simplificando o processo de estimação de estado. O esquema de regularização adotado é baseado na regularização de Tikhonov, que consiste na introdução de um conjunto de informações *a priori* que tornam o problema sempre factível. Esse conjunto de informações *a priori* é obtido da melhor estimativa conhecida para o estado da rede ou a partir de previsões baseadas no conhecimento do comportamento histórico da rede. Ao regularizar o problema, é desejável que o conjunto de informações introduzidas não deteriore a qualidade do estado estimado. Para tanto, deve haver um compromisso entre a qualidade das informações *a priori* e a suas respectivas ponderações. O uso de ponderações inadequadas pode deteriorar o estado estimado ou piorar o condicionamento numérico das matrizes envolvidas, dificultando a convergência do estimador regularizado. Tais problemas são estudados nesta dissertação. O método de regularização foi desenvolvido nas versões completa e desacoplada que foram combinadas em dois algoritmos visando melhorar as características de convergência do estimador regularizado. As versões desenvolvidas foram testadas com as redes de 14 e 118 barras do *Institute of Electrical and Electronics Engineers* (IEEE).

Palavras-chave: Sistemas de energia elétrica, Mínimos quadrados, Teoria da estimativa, Estimativas a-priori.

ABSTRACT

Power system state estimation can be described as a set of functions aimed to provide a real time power system electrical model. This model is composed of the network topology, its electrical parameters and a set of state variables, often defined as the bus voltage phasors. In obtaining of the state variables, although the measurement systems may be designed to ensure that the grid is always observable, eventually, communication failures, topological changes or meter failures can make the grid temporarily unobservable. In these situations the normal equations of the least squares estimation model becomes ill posed, invalidating the state estimation. Conventionally this problem is solved by the observability restoration, where pseudomeasurements are added to the network making it observable again. This thesis deals with the regularized state estimation, that is, an alternative method to deal with the problems associated to temporary network unobservability. With the regularized estimator is possible to avoid the observability restoration stage, simplifying the state estimation procedure. The regularization scheme used is based on the Tikhonov regularization, that consists in the introduction of an *a priori* information set that results in a problem always feasible. This set of *a priori* information is obtained from the best estimative for the state variables or by a forecast based on historical patterns of the grid. During regularization of the problem it is desirable that the information set does not deteriorate the quality of the state estimated. Therefore, must be a compromise between the quality of the *a priori* information and its weighing. The use of inadequate weighting can deteriorate the state estimated or worsen the numerical condition of the involved matrices, disturbing or even preventing the regularized estimator convergence. These problems are studied in this thesis. The regularization method was developed for both the complete and the decoupled variants. These variants are combined in two algorithms that aim to improve the convergence characteristics of the regularized state estimator. The developed variants are tested with the IEEEs 14 and 118 networks.

Keywords: Power systems, Least squares, Estimation theory, A priori estimates.

Lista de Figuras

1.1	Funções de um centro de controle.	3
2.1	Interpretação geométrica da solução de mínimos quadrados.	8
2.2	Função densidade de probabilidade χ^2 para diferentes valores de ν	26
3.1	Curva L genérica.	38
4.1	Rede de 14 barras do IEEE observável.	51
4.2	Magnitude das tensões - 14 barras IEEE observável	53
4.3	Ângulos das tensões - 14 barras IEEE observável	53
5.1	Rede de 14 barras do IEEE não observável 1.	61
5.2	Gráfico de Picard para a rede de 14 barras não observável 1.	63
5.3	Curva L para a rede de 14 barras não observável 1.	66
5.4	Variação do índice $F(x_\lambda)$ para a rede de 14 barras não observável 1.	66
5.5	Variação do índice $J(x_\lambda)$ para a rede de 14 barras não observável 1 usando medidas perfeitas.	68
6.1	Rede de 14 barras do IEEE não observável 2.	74
6.2	Comportamento de um caso divergente.	75
6.3	Variação do índice $J(x)$ para ambos os casos.	76
6.4	Número de iterações para ambos os casos.	76
6.5	Comportamento da função objetivo no estimador desacoplado.	78
6.6	Comportamento das correções nos fluxos de potência medidos.	80
6.7	Múltiplos mínimos e máximos na busca unidimensional.	82
6.8	Comportamento de um caso convergente.	85
6.9	Pivô nulo na matriz ganho fatorada.	86
6.10	Resíduos de estimação normalizados no teste 1 - Versão 1.	89
6.11	Magnitudes das tensões estimadas - Teste 1.	90
6.12	Aberturas angulares estimadas - Teste 1.	91

6.13	Magnitudes das tensões estimadas - Teste 2.	92
6.14	Aberturas angulares estimadas - Teste 2.	93
6.15	Rede de 118 barras do IEEE não observável.	95
6.16	Erro das aberturas angulares estimadas para a rede de 118 barras.	96

Lista de Tabelas

4.1	Valores reais e medidos para a rede de 14 barras observável.	52
4.2	Variâncias adotadas para as medidas.	52
4.3	Índices do estimador convencional durante o processo iterativo.	52
5.1	Estados da rede de 14 barras não observável - Ilhas independentes.	62
5.2	Resíduos estimados - Ilhas independentes.	62
5.3	Resultados do processo iterativo de estimação utilizando a SVD.	64
5.4	Resíduos estimados - SVD.	64
5.5	Estados da rede de 14 barras não observável 1 - SVD.	65
5.6	Estados da rede de 14 barras não observável - Regularização de Tikhonov.	67
5.7	Comparação entre as três soluções de Tikhonov.	68
5.8	Estados da rede de 14 barras não observável 1 - Método desacoplado.	71
5.9	Resultados do processo iterativo de estimação utilizando o método desacoplado.	72
6.1	Vetores de correção - Caso divergente.	77
6.2	Estados da rede de 14 barras não observável 2 - Caso divergente.	77
6.3	Estado estimado para um ponto de inicialização alternativo.	83
6.4	Valores reais e medidos para a rede de 14 barras não observável 2.	88
6.5	Comparação entre os resultados do teste 1.	91
6.6	Comparação entre os resultados do teste 2.	93
6.7	Comparação entre os resultados do teste com a rede de 118 barras.	97

Sumário

1	Introdução	1
2	Revisão Sobre o Método dos Mínimos Quadrados	5
2.1	Introdução	5
2.2	Mínimos Quadrados Lineares	6
2.3	Mínimos Quadrados Não Lineares	8
2.3.1	Critérios de parada e convergência	11
2.3.2	Busca unidimensional como estratégia de globalização	12
2.4	Métodos Diretos para Obtenção da Solução de Mínimos Quadrados	14
2.4.1	A decomposição em valores singulares (SVD)	15
2.4.2	O método das equações normais e a fatoração de Cholesky	18
2.4.3	O método de Peters e Wilkinson	18
2.4.4	A fatoração QR	19
2.5	Aspectos Estatísticos do Estimador de Mínimos Quadrados	20
2.5.1	Os vetores estimados	20
2.5.2	Matrizes de covariância dos vetores estimados	21
2.5.3	Erro médio quadrático	22
2.5.4	A distribuição de probabilidade dos resíduos estimados \hat{r}	23
2.5.5	A distribuição de probabilidade do índice $J(\hat{x})$	24
2.6	Conclusões	26
3	Problemas Inversos e Regularização	29
3.1	Introdução	29
3.2	A SVD Truncada	30
3.3	A Regularização de Tikhonov	32
3.3.1	Regularização de Tikhonov aplicada ao problema não linear	34
3.3.2	Métodos para escolha do parâmetro de regularização	35
3.4	Conclusões	40

4	Revisão Sobre a Estimação de Estado Convencional	41
4.1	Introdução	41
4.2	O Estimador de Estado Convencional	42
4.3	Observabilidade de Sistemas de Energia Elétrica	44
4.3.1	Observabilidade algébrica e observabilidade numérica	45
4.3.2	Observabilidade topológica	46
4.4	Processamento de Erros Grosseiros	46
4.4.1	Classificação dos erros	47
4.4.2	Classificação das medidas	47
4.4.3	Detecção e identificação de erros grosseiros	47
4.5	Estudo de Caso	51
4.6	Conclusões	54
5	A Estimação de Estado Regularizada	55
5.1	Introdução	55
5.2	O Estimador de Estado Regularizado	56
5.2.1	Alguns aspectos relevantes	58
5.2.2	Aspectos computacionais e de implementação	59
5.2.3	Processamento de erros grosseiros na estimação de estado regularizada	60
5.2.4	Estudo de caso	60
5.3	O Estimador Desacoplado Rápido Regularizado	69
5.3.1	Exemplo	71
5.4	Conclusões	71
6	Análise da Robustez dos Estimadores Regularizados	73
6.1	Introdução	73
6.2	Análise de Casos Críticos	74
6.2.1	O efeito da ponderação das pseudomedidas	75
6.2.2	Aplicação do estimador desacoplado regularizado	78
6.2.3	Alternativas preliminares	78
6.3	Os Métodos Propostos	83
6.3.1	Versão 1 - Inicialização com o estimador desacoplado	84
6.3.2	Versão 2 - Redução no conjunto de pseudomedidas	85
6.4	Testes Finais	88
6.4.1	Sistema teste IEEE 14 barras	88
6.4.2	Sistema teste IEEE 118 barras	93
6.5	Conclusões	96

7 Conclusões Gerais	99
Referências Bibliográficas	100
A Trabalho Publicado	105

Capítulo 1

Introdução

Em um cenário adequado, um sistema de energia elétrica deve operar sempre no estado dito normal seguro (A. Abur, 2004). Nesse estado, todas as cargas são atendidas e não há violação das restrições operativas da rede, mesmo após a ocorrência de contingências pré-selecionadas. Na prática, manter a rede sempre operando no estado normal seguro é economicamente inviável, assim, para classificar o estado operativo de um sistema foram estabelecidos os chamados níveis de segurança da rede (Dy Liacco, 1974). Quanto maior o nível de segurança de um sistema, menor a possibilidade de haver interrupções no suprimento de energia (Monticelli, 1999). Nesse contexto, para manter o sistema operando de forma mais segura é primordial monitorar continuamente as suas condições operativas.

O monitoramento das condições de operação de um sistema de energia requer a existência de informações estáticas e dinâmicas suficientes. Essas informações devem ser processadas e o modelo da rede, confiável e em tempo real, deve ser obtido para que sejam executadas as funções de análise de rede (Monticelli, 1999). Segundo Monticelli (1999) o modelo em tempo real de um sistema de energia elétrica é uma representação matemática, quase estacionária, das condições correntes de operação da rede. Esse modelo é construído periodicamente a partir de informações dinâmicas coletadas em tempo real, como as medidas e o estado de chaves e disjuntores, e de informações estáticas obtidas a partir de bancos de dados, como as configurações topológicas básicas da rede e seus parâmetros. A ferramenta que realiza esse processamento e, portanto, fornece o modelo em tempo real da rede é o estimador de estado, o qual é, portanto, uma ferramenta básica e essencial para a operação de um sistema elétrico.

O conceito de estimação de estado em sistemas de energia elétrica foi introduzido por F. D. Schweppe, J. Wildes e D. Rom (Schweppe and Wildes, 1970; Schweppe and Rom, 1970; Schweppe, 1970). Em (Larson et al., 1970b,a) podem ser encontrados os primeiros estudos de viabilidade da aplicação da estimação estática de estado em sistemas de energia reais. O objetivo do processo de estimação de estado é fornecer o estado corrente do sistema de

energia elétrica com a maior precisão possível. Portanto, essa ferramenta fornece informações necessárias para a execução, de forma precisa e eficiente, das funções de análise de rede nos centros de controle.

Sabe-se que a estimação de estado é baseada em medidas disponíveis em tempo real e medidas obtidas do comportamento histórico da rede, tradicionalmente chamadas de pseudo-medidas (Schweppe and Wildes, 1970). Normalmente as medidas que chegam a um centro de controle incluem fluxos de potência e correntes nas linhas de transmissão e transformadores, tensões em barramentos, despacho de geradores, informação sobre o *status* de disjuntores e dispositivos seccionadores, posição de *tap* de transformadores e banco de capacitores chaveados. Além disso, podem existir medidas virtuais tais como injeções nulas em barras de passagem. Em geral, os estimadores de estado são compostos pelas seguintes funções:

- Configurador: Agrupa as informações do *status* de disjuntores e seccionadores de modo a formar um diagrama corrente das conexões elétricas dos componentes da rede. Nesse diagrama são alocados os eventuais medidores disponíveis. Na abordagem tradicional, cria-se um modelo barra/ramo da rede;
- Analisador de observabilidade: Determina se é possível estimar o estado para toda a rede a partir do conjunto de medidas disponíveis. Do contrário, identifica os ramos não observáveis e as barras que formam ilhas observáveis. Entenda-se por barra ou ramo observável aqueles onde é possível estimar as grandezas associadas ao seu ponto de operação;
- Estimador de estado: Baseado nas medidas e informações estáticas disponíveis, estima, segundo algum critério estatístico, o estado mais provável de operação do sistema. Estimado o estado, é possível calcular todas as demais grandezas associadas ao ponto de operação da rede;
- Processador de erros grosseiros: Caso exista redundância suficiente nas medidas, detecta a presença de erros grosseiros nas mesmas, identifica e as elimina do conjunto de medidas, melhorando a confiabilidade do estado estimado.

A princípio pode-se pensar que as funções básicas de um estimador de estado são independentes. Entretanto, como ilustra a figura 1.1, após o processamento das informações dinâmicas pelo sistema *Supervisory Control and Data Acquisition* (SCADA), o resultado final do estimador de estado é obtido através da execução sequencial das funções de estimação de estado. Portanto, o sucesso da estimação de estado depende de quão bem foram executadas as suas funções.

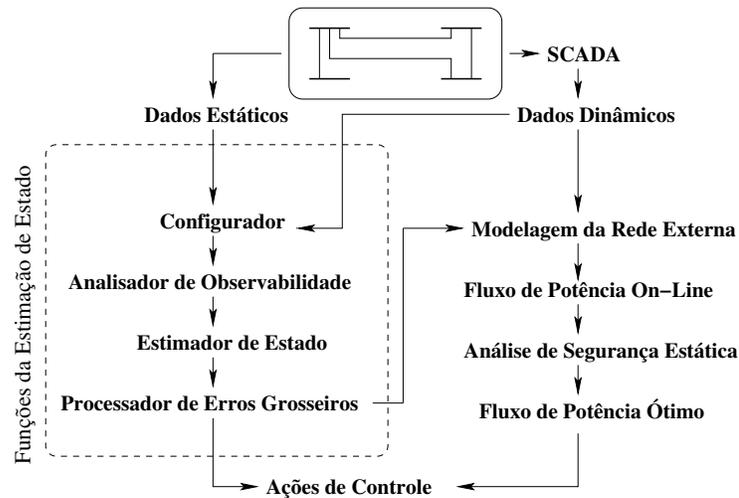


Figura 1.1: Funções de um centro de controle.

Fonte: (de Almeida, 2007)

Na abordagem convencional, ilustrada na figura 1.1, a função de análise de observabilidade tem o papel fundamental de verificar a factibilidade matemática de execução do procedimento de estimação, daí a necessidade de ela ser executada antes do procedimento de estimação do estado. Durante a operação de um sistema elétrico, após a perda temporária de medições, o sistema pode se tornar temporariamente não observável e isso acaba impedindo a obtenção da solução matemática do estimador. Tradicionalmente, esse problema é contornado executando a subfunção de restauração da observabilidade, cujo objetivo é alocar pseudomedidas críticas tornando a rede novamente observável.

A estimação de estado regularizada para sistemas de energia elétrica, proposta em (de Almeida et al., 2012), é uma metodologia baseada na regularização de Tikhonov, cuja principal característica é garantir a factibilidade matemática do estimador. Assim, mesmo quando a observabilidade é perdida, é possível estimar o estado da rede. Nesse método de estimação, a restauração da observabilidade não precisa ser realizada, pois em sua formulação matemática o estimador de estado regularizado utiliza informações *a priori* sobre o estado da rede que garantem a observabilidade.

Durante a aplicação da regularização de Tikhonov é necessário definir um parâmetro de regularização. Esse parâmetro reflete a ponderação que será dada para as informações *a priori* incluídas no problema. O sucesso da estimação de estado regularizada depende da escolha correta desse parâmetro e da qualidade das informações *a priori*.

Nesse contexto, esta dissertação analisa o comportamento do estimador de estado regularizado em função da qualidade e das ponderações das informações *a priori* utilizadas na regularização. Será mostrado que a escolha inadequada do parâmetro de regularização

pode criar dificuldades numéricas nos procedimentos matemáticos de estimação. Assim, são discutidas as origens dos problemas numéricos do estimador regularizado, são analisadas soluções preliminares e, por fim, são propostos dois algoritmos baseados nas versões completa e desacoplada do estimador regularizado. Nos algoritmos propostos, as vantagens da regularização de Tikhonov são agregadas ao processo de estimação de estado em sistemas de energia elétrica, porém, sem trazer complicações matemáticas e de implementação significativas.

Entre as contribuições deste trabalho é possível destacar: (i) a detecção e o estudo detalhado de casos onde o estimador regularizado pode apresentar problemas de convergência, (ii) o desenvolvimento da versão desacoplada rápida do estimador regularizado, e (iii) a proposição de mecanismos que permitem tornar o estimador regularizado mais robusto.

Este trabalho está organizado da seguinte maneira: no capítulo 2 é feita uma revisão dos aspectos algébricos, numéricos e estatísticos do método dos mínimos quadrados. Esse método é largamente utilizado na estimação de estado em sistemas de energia elétrica. Além disso, a regularização de Tikhonov se baseia no seu desenvolvimento. No capítulo 3 os conceitos de problemas inversos e a regularização de Tikhonov são apresentados. Nesse mesmo capítulo são apresentadas as principais metodologias para escolha do parâmetro de regularização. No capítulo 4 é feita uma breve revisão da estimação de estado convencional. O objetivo dessa revisão é apresentar os principais conceitos da estimação de estado, tendo em vista que eles serão utilizados nos desenvolvimentos e testes do modelo regularizado. No capítulo 5 são desenvolvidos o modelo de estimação de estado regularizado completo e o modelo desacoplado. Adicionalmente, são apresentados testes que ilustram as principais características desses estimadores. No capítulo 6 são mostrados casos onde podem ocorrer problemas numéricos com o modelo de estimação de estado regularizado. Para lidar com esses problemas, nesse mesmo capítulo, são propostos dois algoritmos em dois estágios que utilizam as versões completa e desacoplada do estimador regularizado. Ainda neste capítulo são apresentados testes com as redes de 14 e 118 barras do IEEE. Por fim, no capítulo 7 apresentam-se as conclusões deste trabalho.

Capítulo 2

Revisão Sobre o Método dos Mínimos Quadrados

2.1 Introdução

Um problema básico em ciências e engenharia é o de ajustar os parâmetros de um modelo matemático a um conjunto de observações sujeitas a erros. Intuitivamente, espera-se que quanto mais observações estiverem disponíveis melhor será a qualidade do modelo estimado. Entre os métodos de estimação de parâmetros um dos mais utilizados é o método dos mínimos quadrados. Historicamente há divergências quanto ao desenvolvimento original do método, pois enquanto a maioria dos historiadores atribuem o feito ao matemático Carl Friedrich Gauss alguns afirmam que a autoria é do também matemático Adrien Marie Legendre (Bjorck, 1996).

No desenvolvimento do método dos mínimos quadrados é essencial que algumas características estatísticas dos dados observados, isto é, esperança matemática, covariância e distribuição de probabilidade possuam comportamentos bastante específicos. Considere que a partir de um experimento obteve-se o seguinte modelo de medição:

$$z = z^{real} + e \tag{2.1}$$

onde e é o ruído de medição e z^{real} é o valor verdadeiro da variável observada.

No método dos mínimos quadrados admite-se que os ruídos de medição possuem distribuição de probabilidade normal com média nula, ou seja, $E[z] = \mu = z^{real}$. Além disso, considera-se que os dados são independentes, isto é, a matriz de covariância das medidas, $Cov[z] = R_z$, é diagonal. A variância da i -ésima medida é denotada por σ_i^2 , como sugere a equação 2.2.

$$R_z = \text{diag} \left[\sigma_1^2 \quad \sigma_2^2 \quad \cdots \quad \sigma_m^2 \right] \quad (2.2)$$

Os estimadores de estado para sistemas de energia elétrica, em sua maioria, utilizam o método dos mínimos quadrados. Assim, este capítulo apresenta uma revisão matemática do método dos mínimos quadrados, destacando os seus aspectos algébricos, numéricos e estatísticos. Os conceitos apresentados serão úteis para o desenvolvimento do estimador regularizado apresentado nesta dissertação.

2.2 Mínimos Quadrados Lineares

Nesta seção são tratados os problemas onde o modelo de regressão é linear nos parâmetros a serem estimados. Dado um vetor de m observações, z , e o vetor de n parâmetros a serem estimados, x , onde $m \geq n$, pode-se escrever um sistema linear de equações relacionando essas variáveis:

$$Hx = z \quad (2.3)$$

A princípio admite-se que a matriz $H \in \mathbb{R}^{m \times n}$ possui posto completo, isto é, $\text{posto}(H) = n$. Para que esse problema possua solução é necessário que o vetor z possa ser escrito como uma combinação linear das colunas da matriz H . Entretanto, como o conjunto de observações z normalmente contém erros, o sistema linear sobredeterminado 2.3 não admite solução. Apesar de não existir um conjunto de parâmetros x que satisfaça a equação 2.3, ainda há interesse em obter uma solução aproximada para o problema. Assim, considere o vetor de resíduos 2.4.

$$r = z - Hx \quad (2.4)$$

Para a obtenção de uma solução aproximada normalmente é utilizado um critério associado ao vetor de resíduos, r . Atribui-se o nome de solução de mínimos quadrados àquela que minimiza o quadrado da norma euclidiana do vetor de resíduos 2.4 (Aster et al., 2005).

Apesar da definição determinística dada acima, o método provém do critério de estimação de máxima verossimilhança. Na referência (Aster et al., 2005; Bard, 1974) é apresentado um amplo desenvolvimento das características estatísticas do método dos mínimos quadrados. Em geral, durante o processo de obtenção da solução os resíduos são ponderados pelo inverso dos desvios padrão das respectivas medidas:

$$r_w = R_z^{-\frac{1}{2}}[z - Hx] \quad (2.5)$$

Logo, a solução de mínimos quadrados ponderados é obtida minimizando-se o quadrado da norma euclidiana do vetor de resíduos ponderados 2.5 ou, de forma equivalente, a seguinte função objetivo:

$$\min_{x \in \mathbb{R}^n} J(x) = r_w^T r_w = [z - Hx]^T R_z^{-1} [z - Hx] \quad (2.6)$$

A função objetivo $J(x)$ frequentemente é chamada de índice de performance ou simplesmente índice $J(x)$. Aplicando as condições de otimalidade de primeira ordem em 2.6, tem-se:

$$\nabla J(x) = 2\nabla r_w^T r_w = -H^T R_z^{-1} [z - Hx] = 0 \quad (2.7)$$

Rearranjando

$$H^T R_z^{-1} Hx = H^T R_z^{-1} z \quad (2.8)$$

A solução da equação acima somente será única se a matriz $H_w = R_z^{-\frac{1}{2}} H$ possuir posto completo de colunas (Bjorck, 1996). Além disso, para que essa solução, \hat{x} , ocorra em um mínimo local da função objetivo 2.6, a condição de otimalidade de segunda ordem deve ser atendida, isto é, a matriz hessiana da função objetivo, $\nabla^2 J = H^T R_z^{-1} H$, deve ser positiva definida (Bazaraa et al., 1993), ou seja:

$$a^T H^T R_z^{-1} H a > 0, \forall a \in \mathbb{R}^n, a \neq 0 \quad (2.9)$$

As condições acima são asseguradas pois, desde que a matriz H_w tem posto completo, então para qualquer $a \neq 0 \Rightarrow H_w a \neq 0$, logo $a^T H^T R_z^{-1} H a = \|H_w a\|_2^2 > 0$. Pode-se concluir ainda que a solução \hat{x} corresponde a um mínimo global, pois a função objetivo 2.6 é estritamente convexa, ou seja, a sua matriz hessiana é positiva definida e constante em todo o \mathbb{R}^n .

O sistema de equações 2.8 recebe o nome de equações normais, pois o vetor de resíduos, r_w , é ortogonal (ou normal) ao subespaço coluna $\mathcal{R}(H_w)$. Esse fato pode ser interpretado geometricamente através da figura 2.1.

O problema ponderado equivalente à equação 2.3, $H_w x = z_w$, onde $z_w = R_z^{-\frac{1}{2}} z$, não admite solução pois z_w também não pode ser escrito como combinação linear das colunas de H_w , ou de forma equivalente, z_w não está no subespaço coluna $\mathcal{R}(H_w)$, como ilustra a figura 2.1. Entretanto, ainda é possível obter uma solução aproximada \hat{x} com $\hat{z}_w = H_w \hat{x}$, note que \hat{z}_w está no subespaço $\mathcal{R}(H_w)$. Uma boa escolha para a solução é aquela onde o vetor de resíduos $r_w = z_w - \hat{z}_w$ tenha o menor comprimento possível. Geometricamente isso ocorre quando o vetor de resíduos r_w é ortogonal a $\mathcal{R}(H_w)$. Matematicamente isso significa que os

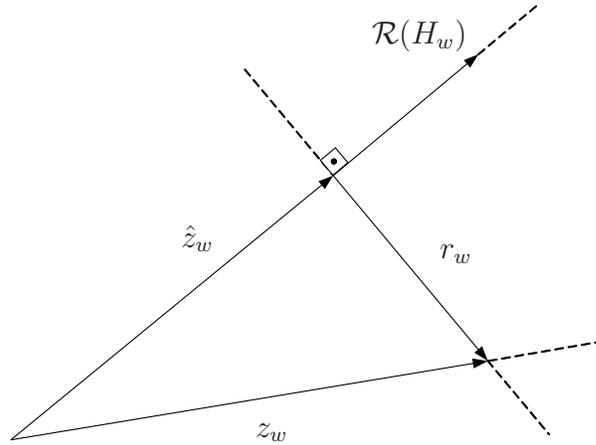


Figura 2.1: Interpretação geométrica da solução de mínimos quadrados.

produtos escalares entre as colunas de H_w e r_w são nulos, isto é:

$$H_w^T r_w = H^T R_z^{-1} H \hat{x} - H^T R_z^{-1} z = 0 \quad (2.10)$$

note que \hat{x} é solução das equações normais 2.8.

2.3 Mínimos Quadrados Não Lineares

Nesta seção são tratados os problemas onde os parâmetros a serem estimados se relacionam de maneira não linear com os dados observados. Os métodos de resolução de tais problemas são iterativos e a cada passo um problema linear relacionado ao problema original é resolvido. Considere o sistema de equações não lineares abaixo:

$$h(x) = z \quad (2.11)$$

onde h é um operador não linear que relaciona os n parâmetros a serem estimados, x , aos m dados observados, z , com $m \geq n$, os quais contêm ruídos gaussianos conforme o modelo de medição definido na equação 2.1.

De modo similar à seção anterior, pode-se escrever um novo sistema de equações levando em conta os desvios padrão dos ruídos de medição, isto é, $h_w(x) = z_w$, onde $h_w(x) = R_z^{-\frac{1}{2}} h(x)$. Devido à presença de ruídos nos dados observados, esse sistema de equações não admite solução. Para obter uma solução de mínimos quadrados será definida uma função objetivo similar à função 2.6, com a diferença que nesta nova função os resíduos ponderados, $r_w = z_w - h_w(x)$, são funções não lineares dos parâmetros. Portanto, a solução é obtida minimizando-se a seguinte função:

$$J(x) = r_w^T r_w = [z - h(x)]^T R_z^{-1} [z - h(x)] = \sum_{i=1}^m \left(\frac{z_i - h_i(x)}{\sigma_i} \right)^2 \quad (2.12)$$

Para solucionar esse problema de otimização irrestrita pode-se utilizar o método de Newton. Um problema de otimização irrestrita consiste em encontrar o mínimo, ou o máximo, de uma função escalar de várias variáveis em todo o espaço onde são definidas essas variáveis sem a presença de qualquer restrição, ou seja,

$$\min_{x \in \mathbb{R}^n} J(x) : \mathbb{R}^n \rightarrow \mathbb{R} \quad (2.13)$$

Quando não existe uma solução analítica para o problema acima utiliza-se algum algoritmo de otimização. Esses algoritmos são inicializados com uma aproximação da solução, x^ν , então uma nova solução é calculada através da seguinte regra: $x^{\nu+1} = x^\nu + \Delta x^\nu$. Em geral esses algoritmos calculam os vetores de correção, Δx^ν , de modo que eles estejam numa direção de descida, s^ν . Uma direção de descida atende a seguinte restrição: $\nabla J(x^\nu)^T s^\nu < 0$. Isso garante que existe um δ , na direção s^ν , suficientemente pequeno tal que $J(x^{\nu+1} + \delta) < J(x^\nu)$. Essa é uma característica básica de grande parte dos métodos de otimização (J. E. Dennis and Schnabel, 1996). Em particular, dada uma solução inicial, o método de Newton cria o seguinte modelo quadrático a partir da expansão em série de Taylor da função objetivo:

$$m_\nu(x) = J(x^\nu) + \nabla J(x^\nu)^T \Delta x^\nu + \frac{1}{2} \Delta x^{\nu T} \nabla^2 J(x^\nu) \Delta x^\nu \quad (2.14)$$

onde $\Delta x^\nu = x - x^\nu$. Aplicando a condição de otimalidade de primeira ordem para obter o minimizador do modelo quadrático $m_\nu(x)$, tem-se a seguinte aproximação para a solução do problema original:

$$\nabla m_\nu(x) = \nabla J(x^\nu) + \nabla^2 J(x^\nu) \Delta x^\nu = 0 \quad (2.15)$$

$$\nabla^2 J(x^\nu) \Delta x^\nu = -\nabla J(x^\nu) \quad (2.16)$$

Portanto, $\Delta x^\nu = -[\nabla^2 J(x^\nu)]^{-1} \nabla J(x^\nu)$. Note que o método depende da existência da matriz hessiana inversa. Após a solução da equação 2.16 o vetor de parâmetros é atualizado e, então, cria-se um novo modelo quadrático 2.14. O processo é repetido até que algum critério de parada seja atendido. O sistema de equações recursivo abaixo resume o método de Newton.

$$x^{\nu+1} = x^\nu - [\nabla^2 J(x^\nu)]^{-1} \nabla J(x^\nu) \quad (2.17)$$

No caso do problema de mínimos quadrados não lineares, o gradiente $\nabla J(x^\nu)$ e a matriz hessiana $\nabla^2 J(x^\nu)$ são definidos como se segue:

$$\nabla J(x^\nu) = -2H(x^\nu)^T R_z^{-1}[z - h(x^\nu)] \quad (2.18)$$

$$\nabla^2 J(x^\nu) = 2H(x^\nu)^T R_z^{-1} H(x^\nu) - 2 \sum_{i=1}^m \frac{1}{\sigma_i^2} [z_i - h_i(x^\nu)] \nabla^2 h_i(x^\nu)$$

onde $H(x^\nu)$ é a matriz Jacobiana do vetor $h(x^\nu)$ calculada no ponto x^ν . Apesar do desenvolvimento feito acima, a maioria dos problemas de mínimos quadrados não lineares são resolvidos pelo método de Gauss-Newton, que é uma variante do método desenvolvido acima. A seguir é apresentado o seu desenvolvimento.

Primeiramente, dada uma solução inicial, x^ν , expande-se $h(x)$ em série de Taylor retendo-se somente os termos de primeira ordem. Assim,

$$h(x) \cong h(x^\nu) + H(x^\nu)[x - x^\nu] \quad (2.19)$$

Então, substitui-se $h(x)$ na equação 2.12 pelo lado direito da equação 2.19. Substituindo e rearranjando, tem-se:

$$\min_{x \in \mathbb{R}^n} J_\nu(x) = [d^\nu - H(x^\nu)x]^T R_z^{-1} [d^\nu - H(x^\nu)x] \quad (2.20)$$

onde $d^\nu = z - h(x^\nu) + H(x^\nu)x^\nu$. Note que a nova função objetivo 2.20 é similar ao problema de mínimos quadrados lineares e sua solução pode ser obtida por inspeção das equações normais 2.8. Nesse sentido, o algoritmo de Gauss-Newton consiste em, a cada iteração ν , resolver um problema de mínimos quadrados linear, obtendo-se uma nova aproximação para o vetor de parâmetros, $x^{\nu+1}$. Esse processo é repetido até que algum critério de parada seja satisfeito. Tradicionalmente o vetor $x^{\nu+1}$ não é obtido diretamente, ao invés disso, resolve-se primeiro para Δx^ν e posteriormente $x^{\nu+1}$ é computado. As equações recursivas 2.21 resumem o método iterativo de Gauss-Newton. Estas equações também são denominadas equações normais. Deve ficar claro que o subproblema linear associado o problema não linear original é função de x e não de Δx .

$$H(x^\nu)^T R_z^{-1} H(x^\nu) \Delta x^\nu = H(x^\nu)^T R_z^{-1} [z - h(x^\nu)] \quad (2.21)$$

$$x^{\nu+1} = x^\nu + \Delta x^\nu$$

Note que ao desprezar o termo de segunda ordem na expressão da matriz hessiana 2.18 o método de Newton torna-se idêntico ao método de Gauss-Newton. Portanto, pode-se considerar o método de Gauss-Newton como uma simplificação no método de Newton.

Em geral, o algoritmo de Gauss-Newton funciona bem quando os problemas são pouco não lineares e os resíduos de estimação não são muito grandes, e como nesses casos o termo de segunda ordem da matriz hessiana é pequeno, geralmente pode ser desprezado. Todavia, nos casos onde isso não se aplica, o algoritmo apresenta convergência lenta e pode até falhar. Caso as expressões analíticas para as derivadas de segunda ordem estejam disponíveis, o método de Newton pode se tornar uma versão interessante, tendo em vista o baixo custo associado ao cômputo e armazenamento desses termos adicionais. Em (Bjorck, 1996; J. E. Dennis and Schnabel, 1996) podem ser encontradas discussões mais amplas sobre as características de convergência de ambos os métodos.

Para que a direção de correção dada pelo método de Newton, obtida pela equação 2.17, seja uma direção de descida, é preciso que:

$$\nabla J(x^\nu)^T \Delta x^\nu = -\nabla J(x^\nu)^T [\nabla^2 J(x^\nu)]^{-1} \nabla J(x^\nu) < 0$$

isto é, a matriz $[\nabla^2 J(x^\nu)]^{-1}$, ou de forma equivalente, a própria matriz hessiana $\nabla^2 J(x^\nu)$ seja positiva definida.

Um resultado interessante do método de Gauss-Newton é que, desde que a matriz Jacobiana $H(x^\nu)$ tenha posto completo, a aproximação para a matriz hessiana será positiva definida e, portanto, a direção da correção será de descida. Não obstante, no método de Newton, mesmo a matriz Jacobiana possuindo posto completo no ponto de cômputo, x^ν , a matriz hessiana pode não ser positiva definida, o que pode causar a falha ou divergência do algoritmo, pois sob essas circunstâncias, uma direção de crescimento da função objetivo poderá ser gerada.

2.3.1 Critérios de parada e convergência

Como os métodos descritos na seção anterior são iterativos, eles devem possuir um ou vários critérios de parada associados. Esses critérios devem indicar quando o processo de cálculo deve ser encerrado. Todavia, isso não necessariamente significa que o objetivo final foi alcançado, de fato, convergência e critério de parada tem conceitos distintos (Bard, 1974). Sob essa ótica, os critérios de parada funcionam também como mecanismos de segurança dos algoritmos computacionais, evitando cálculos excessivos ou desnecessários. Além disso, um bom algoritmo deve ser capaz de checar se de fato o objetivo final foi alcançado. Em geral, esses mecanismos também são associados aos critérios de parada.

Matematicamente, obter uma solução de mínimos quadrados não lineares consiste em encontrar o minimizador global da função objetivo 2.12, isto é, um ponto $\hat{x} \in \mathbb{R}^n$ onde $\nabla J(\hat{x}) = 0$ tal que a matriz hessiana $\nabla^2 J(\hat{x})$ seja positiva definida e a função atinja seu valor

mínimo.

Ambos os métodos apresentados foram desenvolvidos a partir da ideia de encontrar um ponto crítico da função objetivo ($\nabla J(\hat{x}) = 0$). Logo, o algoritmo busca um ponto onde $\nabla J(\hat{x}) \cong 0$, entretanto, devido a questões de escalonamento e dimensões dos problemas, atribuir um valor genérico para $\|\nabla J(\hat{x})\|_2$ pode ser impraticável (Aster et al., 2005; Bard, 1974). Felizmente, existem outras regras que funcionam bem na prática. Essas regras podem ser baseadas no número de iterações, no vetor de correções Δx^ν e no valor da função objetivo $J(x^\nu)$ a cada iteração. Pode-se utilizar, por exemplo, as seguintes condições como critérios de parada:

$$\nu \geq \text{limite de iterações}$$

$$\max |\Delta x^\nu| \leq tol_1 \quad (2.22)$$

$$\frac{J(x^{\nu+1}) - J(x^\nu)}{J(x^\nu)} \leq tol_2$$

A convergência dos métodos de Newton e Gauss-Newton depende das características da função objetivo e do ponto de partida do algoritmo, x^0 . Caso a estimativa inicial seja boa, admite-se que o ponto crítico encontrado ($\nabla J(\hat{x}) = 0$) é um minimizador local da função objetivo (Bard, 1974).

2.3.2 Busca unidimensional como estratégia de globalização

A inicialização do método de Newton pode determinar se o algoritmo vai conseguir encontrar um mínimo local ou falhar, podendo também determinar se a convergência vai ser rápida ou lenta. Existem diferentes técnicas para estimar bons pontos de inicialização, seja através de conhecimento *a priori*, algoritmos específicos ou através de “força bruta” (Bard, 1974). Uma boa inicialização possui um papel importante, entretanto, é preciso fazer um balanço entre o custo para sua obtenção e a taxa de convergência do algoritmo. Nesse contexto, existem as estratégias de globalização que visam obter convergência rápida, mesmo que local, a partir de inicializações não tão boas (J. E. Dennis and Schnabel, 1996). Intuitivamente, espera-se que o algoritmo de Newton produza $J(x^{\nu+1}) < J(x^\nu)$, todavia, o passo completo de Newton pode não satisfazer essa condição. Assim, uma simples modificação na regra de atualização do vetor de parâmetros, mostrada na equação 2.23, pode melhorar as suas características de convergência.

$$x^{\nu+1} = x^\nu + \alpha^\nu \Delta x^\nu \quad (2.23)$$

Na equação acima α^ν é um escalar que tem como objetivo controlar o tamanho do passo Δx^ν . Esse escalar é obtido pelos algoritmos de busca unidimensional, que para uma direção de descida, Δx^ν , buscam ajustar o tamanho do passo de modo a garantir que a função objetivo decresça. O termo unidimensional se deve ao fato de que essa busca é feita em apenas uma direção do \mathbb{R}^n . A busca unidimensional consiste em obter o minimizador da função abaixo.

$$\min_{\alpha > 0} \phi(\alpha) = J(x^\nu + \alpha \Delta x^\nu) \quad (2.24)$$

Existem duas abordagens para o tratamento do problema acima, a busca exata e a inexata. A diferença reside no fato de encontrar o minimizador da função $\phi(\alpha)$ ou, então, algum ponto que não necessariamente é o minimizador mas que produz um decréscimo suficiente em $J(x)$. Em (Sun and Yuan, 2006) tem-se uma revisão ampla sobre as diferentes abordagens. Um critério de aceitação para o tamanho do passo, largamente utilizado na prática, baseia-se na regra de Armijo, isto é, o tamanho do passo é aceito desde que a seguinte desigualdade seja satisfeita,

$$J(x^\nu + \alpha \Delta x^\nu) \leq J(x^\nu) + \lambda \alpha \nabla J(x^\nu)^T \Delta x^\nu \quad (2.25)$$

onde $\lambda \in (0,1)$. Tipicamente λ assume um valor pequeno, por exemplo, 10^{-4} . No contexto da busca inexata pode-se utilizar um mecanismo de interpolação para realizar a busca. Esse mecanismo consiste em criar um polinômio $p(\alpha)$ que interpola a função $\phi(\alpha)$. No caso quadrático temos:

$$\min_{\alpha > 0} p(\alpha) = a\alpha^2 + b\alpha + c \quad (2.26)$$

Para obter os coeficientes de $p(\alpha)$ precisa-se de três equações, por exemplo:

$$p(0) = \phi(0) = J(x^\nu)$$

$$p(1) = \phi(1) = J(x^\nu + \Delta x^\nu) \quad (2.27)$$

$$p'(0) = \phi'(0) = \nabla J(x^\nu)^T \Delta x^\nu$$

Note que esses valores são facilmente obtidos dos algoritmos de Newton ou Gauss-Newton. Logo, o polinômio $p(\alpha)$ fica:

$$p(\alpha) = [J(x^\nu + \Delta x^\nu) - J(x^\nu) - \nabla J(x^\nu)^T \Delta x^\nu] \alpha^2 + \nabla J(x^\nu)^T \Delta x^\nu \alpha + J(x^\nu) \quad (2.28)$$

cujos pontos críticos $\hat{\alpha}$ é dado por:

$$\hat{\alpha} = \frac{-\nabla J(x^\nu)^T \Delta x^\nu}{2[J(x^\nu + \Delta x^\nu) - J(x^\nu) - \nabla J(x^\nu)^T \Delta x^\nu]} \quad (2.29)$$

e a derivada de segunda ordem de $p(\alpha)$ é dada por:

$$p''(\alpha) = 2[J(x^\nu + \Delta x^\nu) - J(x^\nu) - \nabla J(x^\nu)^T \Delta x^\nu] \quad (2.30)$$

Considere uma situação onde a condição 2.25 não é atendida. Como $0 < \lambda < 1$, $p''(\alpha) > 0$, logo, $\hat{\alpha}$ é o minimizador de $p(\alpha)$. Note que $\hat{\alpha} > 0$, pois de acordo com os métodos de descida, $p'(0) = \nabla J(x^\nu)^T \Delta x^\nu < 0$. Dado o tamanho do passo, $\hat{\alpha}$, a função objetivo é calculada no novo ponto, x^ν , e a condição de Armijo 2.25 é testada novamente. Caso ela não seja atendida, cria-se um novo $p(\alpha)$ para o novo ponto e o processo é repetido até que a condição de Armijo seja satisfeita. Normalmente o multiplicador α é fixado dentro de um intervalo fechado $[l, u]$, ou seja, caso o $\hat{\alpha}$ calculado esteja fora dos limites, ele é forçado para dentro do intervalo. Um intervalo típico é $[\frac{1}{10}, \frac{1}{2}]$. O algoritmo 1 resume o processo iterativo da busca unidimensional inexata.

Algoritmo 1: Busca unidimensional inexata com regra de Armijo.

Dados: λ, l, u

$\alpha^\nu = 1;$

enquanto $J(x^\nu + \alpha^\nu \Delta x^\nu) > J(x^\nu) + \lambda \alpha^\nu \nabla J(x^\nu)^T \Delta x^\nu$ **faça**

 Calcule $\hat{\alpha}$ a partir da equação 2.29;

se $\hat{\alpha} < l \alpha^\nu$ **então**

 | $\alpha^\nu \leftarrow l \alpha^\nu;$

senão se $\hat{\alpha} > u \alpha^\nu$ **então**

 | $\alpha^\nu \leftarrow u \alpha^\nu;$

senão

 | $\alpha^\nu \leftarrow \hat{\alpha};$

fim se

fim enqto

$x^{\nu+1} = x^\nu + \alpha^\nu \Delta x^\nu$

2.4 Métodos Diretos para Obtenção da Solução de Mínimos Quadrados

Nas seções 2.2 e 2.3 o método dos mínimos quadrados foi aplicado aos problemas linear e não linear. Todavia, ainda é preciso desenvolver os métodos numéricos capazes de obter a solução dos problemas de mínimos quadrados formulados. Nesta seção são desenvolvidos alguns métodos para obtenção da solução de problemas de mínimos quadrados lineares.

Como mencionado na seção 2.3, o algoritmo de Gauss-Newton consiste na resolução de uma sequência de problemas lineares, portanto, os métodos que são desenvolvidos aqui também podem ser utilizados nos problemas não lineares.

A solução de um sistema de equações lineares genérico do tipo $Ax = b$ pode ser obtida de forma direta ou iterativa. Entende-se por método iterativo aquele que inicia com uma aproximação x^0 e, então, uma sequência de pontos $\{x^n | n \in \mathbb{N}\}$ é gerada objetivando refinar a solução até que a convergência seja alcançada. Alguns exemplos dessa classe de métodos são os métodos de Jacobi, Gauss-Seidel e Gradiente Conjugado. Os métodos diretos, que são tratados nesta seção, são baseados na fatoração da matriz A . Nos métodos diretos, caso os cálculos fossem realizados com precisão numérica infinita, a solução exata do problema seria obtida. Em particular, uma comparação entre métodos diretos e iterativos aplicados a problemas de sistemas de energia elétrica pode ser encontrada em (Alves et al., 1999). Uma ampla revisão dos algoritmos de fatoração que são tratados nesta seção pode ser encontrada em (Bjorck, 1996; Hansen, 1998).

2.4.1 A decomposição em valores singulares (SVD)

Dada uma matriz $A \in \mathbb{R}^{m \times n}$ com $m \geq n$ e $\text{posto}(A) = r \leq n$, a sua decomposição em valores singulares ou *Singular Value Decomposition* (SVD) é dada por:

$$A = U\Sigma V^T \quad (2.31)$$

onde $U = (u_1 \cdots u_m) \in \mathbb{R}^{m \times m}$ e $V = (v_1 \cdots v_n) \in \mathbb{R}^{n \times n}$ são matrizes ortonormais, isto é, $U^T U = I_m$ e $V^T V = I_n$. O fator $\Sigma = \text{diag}(\sigma_1 \cdots \sigma_n) \in \mathbb{R}^{m \times n}$ é uma matriz retangular diagonal, cujos elementos nas posições diagonais são números não negativos arranjados de maneira decrescente como indicado na equação 2.32. Os elementos remanescentes, $i \neq j$, são todos nulos.

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \geq 0 \quad (2.32)$$

Os elementos σ_i são chamados de valores singulares da matriz A , enquanto que os vetores u_i e v_i são os vetores singulares à esquerda e à direita, respectivamente. Note que para $r < n$ os $n - r$ valores singulares restantes são nulos. Caso apenas os primeiros r valores singulares sejam não nulos, então, Σ pode ser escrita da seguinte forma:

$$\Sigma = \begin{bmatrix} \Sigma_r & 0_{r \times n-r} \\ 0_{m-r \times r} & 0_{m-r \times n-r} \end{bmatrix} \quad (2.33)$$

onde Σ_r é uma matriz diagonal de dimensão $r \times r$. Dessa maneira o problema linear sobrede-

terminado, $H_w x = z_w$, apresentado na seção 2.2, pode ser resolvido utilizando a SVD como segue:

$$\begin{aligned}
 H_w x &= z_w \\
 U \Sigma V^T x &= z_w \\
 \Sigma V^T x &= U^T z_w \\
 \Sigma^T \Sigma V^T x &= \Sigma^T U^T z_w \\
 V^T x &= (\Sigma^T \Sigma)^{-1} \Sigma^T U^T z_w \\
 x_{\dagger} &= V (\Sigma^T \Sigma)^{-1} \Sigma^T U^T z_w
 \end{aligned}$$

Perceba que a quinta passagem acima só é possível se a matriz Σ tem posto completo de colunas, isto é, se todos os valores singulares de H_w forem não nulos. Definindo $\Sigma^T = [\Sigma_n^T \ 0^T]$, então, pode-se mostrar que $(\Sigma^T \Sigma)^{-1} \Sigma^T = [\Sigma_n^{-1} \ 0^T]$.

A seguir mostra-se que a solução x_{\dagger} é uma solução de mínimos quadrados.

$$\begin{aligned}
 x_{\dagger} &= V [\Sigma_n^{-1} \ 0^T] U^T z_w \\
 &= V \Sigma_n^{-2} V^T V \Sigma^T U^T z_w \\
 &= (V \Sigma_n^2 V^T)^{-1} V \Sigma^T U^T z_w \\
 &= (V \Sigma^T U^T U \Sigma V^T)^{-1} V \Sigma^T U^T z_w \\
 &= (H_w^T H_w)^{-1} H_w^T z_w \\
 &= (H^T R_z^{-1} H)^{-1} H^T R_z^{-1} z
 \end{aligned}$$

A decomposição SVD da matriz A pode ser reescrita da seguinte maneira:

$$A = [U_r \ U_0] \begin{bmatrix} \Sigma_r & 0 \\ 0 & 0 \end{bmatrix} [V_r \ V_0]^T \quad (2.34)$$

onde U_r e V_r são submatrizes contendo as r primeiras colunas de U e V respectivamente, enquanto que U_0 e V_0 contêm as colunas restantes das respectivas matrizes originais. De fato, essas submatrizes estão fortemente relacionadas com os subespaços fundamentais da matriz A . Enquanto os vetores coluna de U_r e V_r geram bases ortonormais para $\mathcal{R}(A)$ e $\mathcal{R}(A^T)$, respectivamente, os vetores coluna de U_0 e V_0 geram bases ortonormais para $\mathcal{N}(A^T)$ e $\mathcal{N}(A)$, respectivamente. A equação 2.34 pode ser escrita em sua forma compacta:

$$A = U_r \Sigma_r V_r^T \quad (2.35)$$

Um importante resultado dessa decomposição é a matriz inversa generalizada de Moore-Penrose $A^\dagger = V_r \Sigma_r^{-1} U_r^T$, também chamada de pseudoinversa (Aster et al., 2005; Bjorck, 1996; Hansen, 1998). Note que essa generalização permite a obtenção da inversa de qualquer matriz, seja ela retangular ou mesmo com deficiência de posto. Isso permite a obtenção da solução direta de qualquer sistema linear seja ele determinado, subdeterminado ou sobredeterminado. Uma caracterização das soluções obtidas através da SVD em diferentes classes de sistemas lineares pode ser obtida em (Aster et al., 2005).

Outro resultado importante da decomposição em valores singulares é a análise de sensibilidade da solução de um problema linear genérico $x_\dagger = A^\dagger b$. Esta análise indica quanto a solução x_\dagger é alterada devido a pequenas perturbações no vetor independente, $b + \delta b$. A medida dessa sensibilidade é dada pelo número de condição da matriz de coeficientes, matematicamente definido como:

$$\text{cond}(A) = \|A\|_2 \|A^\dagger\|_2 = \frac{\sigma_1}{\sigma_n} \quad (2.36)$$

O número de condição fornece uma ideia de quão imprecisa será a solução do sistema de equações devido a perturbações no vetor independente. Essa imprecisão também está relacionada com a quantidade de casas decimais que são representadas nas operações matemáticas, isto é, pelo epsilon da máquina ¹. Por exemplo, se o número de condição possuir a mesma ordem de grandeza do epsilon da máquina, então, a solução do sistema linear tem pouca ou nenhuma precisão. Quando um sistema linear tem um número de condição muito elevado diz-se que ele é mal condicionado. Note que, quando a matriz A tem posto incompleto seu condicionamento numérico é infinito.

A SVD possui propriedades muito importantes para análise de sistemas lineares, principalmente nos casos com deficiência de posto. A falta de algoritmos numericamente estáveis para o cômputo da decomposição tem limitado o seu uso na prática. Além disso, os algoritmos disponíveis são muito caros do ponto de vista computacional e questões como esparsidade tornam esse método pouco competitivo na aplicação em problemas de grande porte (Bjorck, 1996; Hansen, 1998). Nesse contexto, a seguir são desenvolvidos os algoritmos mais populares para a obtenção da solução de mínimos quadrados.

¹Representa um limite superior para os erros introduzidos durante cálculos em ponto flutuante. Em computadores de 64 bits com dupla precisão esse valor é aproximadamente igual a 10^{-16} .

2.4.2 O método das equações normais e a fatoração de Cholesky

Na seção 2.2 foi mostrado que a solução de mínimos quadrados é obtida através da solução das equações normais,

$$H^T R_z^{-1} H x = H^T R_z^{-1} z \quad (2.37)$$

Foi mostrado também que a matriz de coeficientes $G = H^T R_z^{-1} H \in \mathbb{R}^{n \times n}$ é simétrica e caso o posto(H) = n ela é positiva definida. Sob essas condições a matriz de coeficientes pode ser escrita unicamente como $G = LL^T$ onde $L \in \mathbb{R}^{n \times n}$ é uma matriz diagonal inferior com os elementos da diagonal positivos. Essa decomposição é conhecida como fatoração de Cholesky (Bjorck, 1996). Portanto, o sistema de equações 2.37 pode ser reescrito na seguinte forma:

$$LL^T x = H^T R_z^{-1} z \quad (2.38)$$

A solução do sistema acima pode ser obtida pela solução de dois subsistemas. Primeiramente define-se $Ly = H^T R_z^{-1} z$ onde y é obtido por substituição direta, em seguida x é encontrado resolvendo $L^T x = y$ por retrossubstituição.

2.4.3 O método de Peters e Wilkinson

Uma matriz genérica $A \in \mathbb{R}^{m \times n}$ com $m \geq n$ pode ser colocada numa forma triangular superior através da eliminação Gaussiana com pivoteamento parcial (Bjorck, 1996). Logo, a matriz H_w definida na seção 2.2 pode ser escrita na seguinte forma $H_w = LU$ onde $L \in \mathbb{R}^{m \times n}$ (diferente do fator de Cholesky) é uma matriz diagonal inferior e $U \in \mathbb{R}^{n \times n}$ é diagonal superior. Quando posto(H_w) = n , a matriz U é não singular. Portanto, pode-se escrever as equações normais 2.37 da seguinte maneira:

$$U^T L^T L U x = U^T L^T z_w \quad (2.39)$$

Como U é não singular, então,

$$L^T L U x = L^T z_w \quad (2.40)$$

Para a solução da equação acima primeiro admite-se que $Ux = a$, aplica-se a fatoração de Cholesky ao termo $L^T L$ e, então, encontra-se a através de substituição direta e retrossubstituição, conforme mostrado na seção anterior. Em seguida, obtém-se x por retrossubstituição no subproblema $Ux = a$. A vantagem desse método é que o termo $L^T L$ tem um

condicionamento numérico melhor que o termo $H_w^T H_w$ (Bjorck, 1996).

2.4.4 A fatoração QR

A matriz $H_w \in \mathbb{R}^{m \times n}$ pode ser escrita como $H_w = QR$, onde $Q \in \mathbb{R}^{m \times m}$ é uma matriz ortonormal e $R \in \mathbb{R}^{m \times n}$ é uma matriz diagonal superior. Sabe-se que uma matriz ortonormal satisfaz $QQ^T = Q^T Q = I$, isto é, $Q^{-1} = Q^T$. Além disso, para qualquer vetor $a \in \mathbb{R}^m$, $\|Qa\|_2 = \|a\|_2$ e para quaisquer $a, b \in \mathbb{R}^m$, $a^T b = (Qa)^T (Qb)$.

Os fatores Q e R podem ser particionados da seguinte maneira $Q^T = [Q_1^T \ Q_2^T]$ e $R^T = [R_1^T \ 0^T]$. Onde $Q_1 \in \mathbb{R}^{m \times n}$, $Q_2 \in \mathbb{R}^{m \times m-n}$ e $R_1 \in \mathbb{R}^{n \times n}$. Quando $\text{posto}(H) = n$, R_1 é não singular.

O problema 2.6 pode ser reescrito segundo a equação 2.41.

$$\min \|H_w x - z_w\|_2^2 \quad (2.41)$$

Tendo em vista que a multiplicação de uma matriz ortonormal por um vetor não altera sua norma, o problema abaixo também é equivalente ao problema original 2.6

$$\min \|Q^T (H_w x - z_w)\|_2^2 \quad (2.42)$$

Sabe-se que $Q^T H_w = R$, então, temos:

$$\min \|R x - Q^T z_w\|_2^2 \quad (2.43)$$

ou de forma equivalente,

$$\min \|R_1 x - Q_1^T z_w\|_2^2 + \|0x - Q_2^T z_w\|_2^2 \quad (2.44)$$

Note que o segundo termo da equação acima é constante, isto é, para qualquer x seu valor permanece inalterado. Portanto, a solução é obtida minimizando-se o primeiro termo. Tendo em vista que R_1 é não singular, a solução é obtida por retrosubstituição no sistema $R_1 x = Q_1^T z_w$. Os algoritmos tradicionais para obtenção da matriz Q são as transformações de Householder e as rotações de Givens (Bjorck, 1996).

Na seção 2.4.1 mostrou-se que H_w pode ser escrita como $H_w = U \Sigma V^T$. Aplicando-se a mesma decomposição às equações normais temos que $G = H^T R_z^{-1} H = V \Sigma_n^2 V^T$. Portanto, podemos concluir que $\text{cond}(G) = \left(\frac{\sigma_1}{\sigma_n}\right)^2 = \text{cond}(H_w)^2$. Isso evidencia que, encontrar a solução de mínimos quadrados através de mecanismos que fatoram a matriz H_w , como o método da fatoração QR, ao invés da matriz de coeficientes $G = H^T R_z^{-1} H$, pode ser muito mais

estável numericamente, pois ao se utilizar as equações normais o condicionamento numérico das equações é elevado ao quadrado.

Para contornar o problema do elevado custo computacional associado à fatoração QR é possível adotar uma abordagem híbrida. Considere as equações normais 2.8, aplicando-se a fatoração $H_w = QR$ ao seu lado esquerdo elas podem ser reescritas como $R^T R x = H^T R_z^{-1} z$ ou na forma compacta $R_1^T R_1 x = H^T R_z^{-1} z$. Nessa forma elas são denominadas equações semi normais. A solução dessas equações é obtida facilmente por substituição direta e retrosubstituição. A abordagem híbrida é idêntica ao método da fatoração de Cholesky, a diferença é que a obtenção dos fatores triangulares através da fatoração ortogonal é muito mais estável do ponto de vista numérico.

2.5 Aspectos Estatísticos do Estimador de Mínimos Quadrados

Nesta seção são apresentados alguns aspectos estatísticos básicos do estimador de mínimos quadrados linear. É dado apenas o enfoque necessário para o desenvolvimento do trabalho. Por analogia, ao longo do trabalho esses aspectos serão estendidos para o caso não linear. Um desenvolvimento matemático mais elegante, ou mais adequado, das características estatísticas do estimador de mínimos quadrados não linear, pode ser encontrado em (Gallant, 1987).

2.5.1 Os vetores estimados

Na seção 2.2 foi discutido como estimar uma solução para o seguinte problema:

$$z = Hx$$

onde $z = z^{real} + e$, sendo e um vetor de variáveis aleatórias com $E[e] = 0$ e $Cov[z] = E[ee^T] = R_z$. Naquela seção foi utilizado o método dos mínimos quadrados, cuja estimativa para a solução do problema acima é dada por:

$$\begin{aligned} \hat{x} &= (H^T R_z^{-1} H)^{-1} H^T R_z^{-1} z \\ &= (H^T R_z^{-1} H)^{-1} H^T R_z^{-1} (Hx^{real} + e) \\ &= x^{real} + (H^T R_z^{-1} H)^{-1} H^T R_z^{-1} e \end{aligned} \tag{2.45}$$

O vetor de medidas estimadas pode ser escrito como:

$$\begin{aligned}
 \hat{z} &= H\hat{x} \\
 &= H(H^T R_z^{-1} H)^{-1} H^T R_z^{-1} z \\
 &= H(H^T R_z^{-1} H)^{-1} H^T R_z^{-1} (Hx^{real} + e) \\
 &= Hx^{real} + H(H^T R_z^{-1} H)^{-1} H^T R_z^{-1} e
 \end{aligned} \tag{2.46}$$

De forma equivalente, o vetor de resíduos estimados é dado por:

$$\begin{aligned}
 \hat{r} &= z - H\hat{x} \\
 &= z - H(H^T R_z^{-1} H)^{-1} H^T R_z^{-1} z \\
 &= [I - H(H^T R_z^{-1} H)^{-1} H^T R_z^{-1}] z \\
 &= [I - H(H^T R_z^{-1} H)^{-1} H^T R_z^{-1}] (Hx^{real} + e) \\
 &= [I - H(H^T R_z^{-1} H)^{-1} H^T R_z^{-1}] e
 \end{aligned} \tag{2.47}$$

Frequentemente defini-se a matriz de sensibilidade $S = I - H(H^T R_z^{-1} H)^{-1} H^T R_z^{-1}$.

2.5.2 Matrizes de covariância dos vetores estimados

Considere que dois vetores de variáveis aleatórias, y e v , se relacionam da seguinte maneira:

$$y = Av \tag{2.48}$$

onde A é uma matriz de coeficientes constantes representando uma transformação linear. Nesse caso aplicam-se as seguintes propriedades:

$$E[y] = AE[v] \tag{2.49}$$

$$Cov[y] = R_y = AR_v A^T \tag{2.50}$$

onde R_v é a matriz de covariância de v . Utilizando a equação 2.45 e considerando as equações 2.48 e 2.50 a matriz de covariância do vetor \hat{x} é dada por:

$$Cov[\hat{x}] = R_{\hat{x}} = (H^T R_z^{-1} H)^{-1} H^T R_z^{-1} Cov[z] R_z^{-T} H (H^T R_z^{-1} H)^{-T} = (H^T R_z^{-1} H)^{-1} \quad (2.51)$$

A matriz de covariância do vetor de medidas estimadas pode ser escrita como:

$$Cov[\hat{z}] = R_{\hat{z}} = H Cov[\hat{x}] H^T = H R_{\hat{x}} H^T \quad (2.52)$$

Equivalentemente, a matriz de covariância do vetor de resíduos estimados é dada por:

$$Cov[\hat{r}] = R_{\hat{r}} = S Cov[z] S^T = R_z - H R_{\hat{x}} H^T = R_z - R_{\hat{z}} \quad (2.53)$$

2.5.3 Erro médio quadrático

O erro médio quadrático de uma estimativa \hat{x} , com relação ao conjunto de parâmetros reais, x^{real} , é definido como:

$$EMQ(\hat{x}) = E(\hat{x} - x^{real})^2 \quad (2.54)$$

A equação acima pode ser rearranjada da seguinte maneira:

$$\begin{aligned} EMQ(\hat{x}) &= E[(\hat{x} - E(\hat{x})) + (E(\hat{x}) - x^{real})]^2 \\ &= E[\hat{x} - E(\hat{x})]^2 + [E(\hat{x}) - x^{real}]^2 \\ &= Cov(\hat{x}) + [E(\hat{x}) - x^{real}]^2 \end{aligned}$$

onde o termo $E(\hat{x}) - x^{real}$ é chamado de vício ou viés do estimador. O viés de um estimador dá uma ideia da sua precisão (Casella and Berger, 2002). Quando o viés de um estimador é nulo, diz-se que ele é não viesado. Considerando a equação 2.49, a esperança da solução de mínimos quadrados, \hat{x} , é dada por:

$$E[\hat{x}] = (H^T R_z^{-1} H)^{-1} H^T R_z^{-1} E[z] \quad (2.55)$$

tendo em vista que $E[z] = z^{real}$ e $H x^{real} = z^{real}$ pode-se concluir $E[\hat{x}] = x^{real}$, portanto, o estimador de mínimos quadrados linear é não viesado.

2.5.4 A distribuição de probabilidade dos resíduos estimados \hat{r}

Talvez a principal consideração feita neste capítulo foi a de que o vetor de ruídos, e , é composto por variáveis aleatórias independentes, com distribuição normal de probabilidade. Dentre outros aspectos, isso significa que o i -ésimo elemento do vetor e , tem uma distribuição de probabilidade dada pela seguinte função:

$$f(e_i) = \frac{1}{\sigma_i \sqrt{2\pi}} \exp\left(-\frac{1}{2\sigma_i^2}(e_i - \mu_i)^2\right) \quad (2.56)$$

onde $E[e_i] = \mu_i = 0$ e $Var[e_i] = \sigma_i^2 = R_{z_i,i}$. Utilizando a matriz de sensibilidade S e considerando a equação 2.47, temos que $\hat{r} = Se$, isto é, o vetor de resíduos estimados pode ser escrito como uma combinação linear dos elementos do vetor de ruídos. Assim, o vetor \hat{r} também possui distribuição normal. Entretanto, como S não é necessariamente diagonal, os elementos i e j de \hat{r} podem ser correlacionados, portanto, o vetor de resíduos estimados possui distribuição normal multivariada. Logo, a função densidade de probabilidade conjunta de \hat{r} é da pela seguinte função:

$$f(\hat{r}) = \frac{1}{(2\pi)^{m/2}} \frac{1}{\sqrt{\det(Cov[\hat{r}])}} \exp\left(-\frac{(\hat{r} - \mu)^T (Cov[\hat{r}])^{-1} (\hat{r} - \mu)}{2}\right) \quad (2.57)$$

onde $E[\hat{r}] = SE[e] = \mu = 0$ e $Cov[\hat{r}] = R_{\hat{r}} = R_z - R_{\hat{z}}$. Note que a equação 2.57 só possui sentido quando $\det(Cov[\hat{r}]) > 0$, isto é, quando a matriz $R_{\hat{r}}$ for positiva definida. Todavia, isso nunca ocorre, pois, por construção a matriz $R_{\hat{r}}$ é sempre positiva semidefinida. De fato, $posto(R_{\hat{r}}) = m - n$, onde m e n correspondem às dimensões de \hat{r} e \hat{x} , respectivamente.

Suponha agora um subconjunto \hat{r}_k de \hat{r} com k elementos, tal que $k \leq m - n$. Caso o posto da matriz de covariância de \hat{r}_k seja completo, isto é, $posto(R_{\hat{r}_k}) = k$, então, podemos definir uma função densidade de probabilidade conjunta associada a esses k elementos, por:

$$f(\hat{r}_k) = \frac{1}{(2\pi)^{k/2}} \frac{1}{\sqrt{\det(R_{\hat{r}_k})}} \exp\left(-\frac{(\hat{r}_k - \mu_k)^T R_{\hat{r}_k}^{-1} (\hat{r}_k - \mu_k)}{2}\right) \quad (2.58)$$

Quando os elementos de \hat{r} são fracamente correlacionados, isto é, quando os elementos fora da diagonal de $R_{\hat{r}}$ são pequenos se comparados com os respectivos elementos da diagonal, é possível fazer uma simplificação no problema. Essa simplificação consiste em ignorar a correlação entre as variáveis e , portanto, considerar que cada elemento de \hat{r} é uma variável aleatória independente com distribuição normal. Alternativamente, pode-se obter as funções de densidade de probabilidade marginais para cada variável aleatória. Considerando a primeira situação, podemos definir a seguinte função densidade de probabilidade para o i -ésimo elemento de \hat{r} :

$$f(\hat{r}_i) = \frac{1}{\sqrt{2\pi R_{\hat{r}_i,i}}} \exp\left(-\frac{1}{2R_{\hat{r}_i,i}}(\hat{r}_i - \mu_i)^2\right) \quad (2.59)$$

Além disso, como $\hat{r}_i \sim N(\mu_i, R_{\hat{r}_i,i})$, a relação $\hat{r}_i^N = \frac{\hat{r}_i - \mu_i}{\sqrt{R_{\hat{r}_i,i}}}$ transforma a variável aleatória \hat{r}_i numa variável aleatória $\hat{r}_i^N \sim N(0,1)$. Para ilustrar isso, considere que estejamos interessados em calcular a probabilidade de \hat{r}_i ser menor ou igual a um determinado valor K , ou seja:

$$P(\hat{r}_i \leq K) = \int_{-\infty}^K \frac{1}{\sqrt{2\pi R_{\hat{r}_i,i}}} \exp\left(-\frac{1}{2R_{\hat{r}_i,i}}(\hat{r}_i - \mu_i)^2\right) d\hat{r}_i$$

fazendo uma mudança de variáveis,

$$P(\hat{r}_i \leq K) = \int_{-\infty}^{\frac{K-\mu_i}{\sqrt{R_{\hat{r}_i,i}}}} \frac{\sqrt{R_{\hat{r}_i,i}}}{\sqrt{2\pi R_{\hat{r}_i,i}}} \exp\left(-\frac{(\hat{r}_i^N)^2}{2}\right) d\hat{r}_i^N = \int_{-\infty}^C \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(\hat{r}_i^N)^2}{2}\right) d\hat{r}_i^N$$

portanto,

$$P(\hat{r}_i \leq K) = P(\hat{r}_i^N \leq C)$$

onde $C = \frac{K-\mu_i}{\sqrt{R_{\hat{r}_i,i}}}$ e $\mu_i = 0$. Note que o integrando, função de \hat{r}_i^N , possui distribuição normal com média nula e variância unitária, isto é, distribuição normal padronizada. Essa propriedade é bastante útil, pois fazendo uma transformação adequada em uma variável aleatória qualquer com distribuição normal, sempre é possível utilizar a distribuição normal padronizada.

2.5.5 A distribuição de probabilidade do índice $J(\hat{x})$

Considere um vetor de variáveis aleatórias $y \sim N(\mu, V)$ e uma matriz simétrica e definida positiva Q . Assim, uma forma quadrática pode ser definida como:

$$q(y) = y^T Q y \quad (2.60)$$

Considerando que q é função de variáveis aleatórias, intuitivamente pode-se dizer que $q(y)$ também é uma variável aleatória. De fato, $q(y)$ possui distribuição de probabilidade χ^2 , leia-se qui quadrado, se e somente se QV é indepoente, isto é, $QVQV = QV$. A função densidade de probabilidade é dada por:

$$f_{\chi^2}(q) = \frac{1}{2^{\nu/2}\Gamma(\nu/2)} q^{\frac{\nu}{2}-1} \exp\left(-\frac{q}{2}\right) \quad (2.61)$$

onde ν é o número de graus de liberdade da distribuição e a função gamma é:

$$\Gamma(x) = \int_0^{\infty} t^{x-1} \exp(-t) dt$$

Como consequência, temos que:

$$E[q(y)] = \text{tr}(QV) + \mu^T Q\mu \quad (2.62)$$

$$\text{Var}[q(y)] = 2\text{tr}(QVQV) + 4\mu^T QVQ\mu \quad (2.63)$$

Dada a solução de mínimos quadrados, \hat{x} , o índice de performance $J(\hat{x})$ pode ser escrito como uma forma quadrática da seguinte maneira:

$$J(\hat{x}) = \hat{r}^T R_z^{-1} \hat{r} = [z - H\hat{x}]^T R_z^{-1} [z - H\hat{x}] \quad (2.64)$$

Como discutido na seção anterior, $\hat{r} \sim N(0, R_{\hat{r}})$, logo, aplicando as equações 2.62 e 2.63 ao índice $J(\hat{x})$ e substituindo os termos:

$$E[J(\hat{x})] = \text{tr}(R_z^{-1} R_{\hat{r}}) \quad (2.65)$$

$$\text{Var}[J(\hat{x})] = 2\text{tr}(R_z^{-1} R_{\hat{r}} R_z^{-1} R_{\hat{r}}) \quad (2.66)$$

desenvolvendo a equação 2.65, temos:

$$\begin{aligned} \text{tr}(R_z^{-1} R_{\hat{r}}) &= \text{tr}(R_z^{-1} [R_z - R_{\hat{z}}]) \\ &= \text{tr}(I_m) - \text{tr}(R_z^{-1} R_{\hat{z}}) \\ &= \text{tr}(I_m) - \text{tr}(R_{\hat{z}} R_z^{-1}) \\ &= \text{tr}(I_m) - \text{tr}(H R_{\hat{x}} H^T R_z^{-1}) \\ &= \text{tr}(I_m) - \text{tr}(H^T R_z^{-1} H R_{\hat{x}}) \\ &= \text{tr}(I_m) - \text{tr}(I_n) \\ &= m - n \end{aligned}$$

e desenvolvendo a equação 2.66, temos:

$$\begin{aligned}
2\text{tr}(R_z^{-1}R_{\hat{r}}R_z^{-1}R_{\hat{r}}) &= 2\text{tr}(R_z^{-1}(R_z - R_{\hat{z}})R_z^{-1}(R_z - R_{\hat{z}})) \\
&= 2\text{tr}((I_m - R_z^{-1}R_{\hat{z}})(I_m - R_z^{-1}R_{\hat{z}})) \\
&= 2\text{tr}(I_m - R_z^{-1}R_{\hat{z}}) \\
&= 2(m - n)
\end{aligned}$$

Portanto, $E[J(\hat{x})] = m - n$ e $\text{Var}[J(\hat{x})] = 2(m - n)$. Onde $\nu = m - n$ é o número de graus de liberdade da distribuição χ^2 . Como ilustra a figura 2.2, quando o número de graus de liberdade cresce a distribuição χ^2 tende a uma distribuição normal com $\mu = \nu$ e $\sigma = \sqrt{2\nu}$. Em geral, quando $\nu > 30$ assume-se que $J(\hat{x}) \sim N(\nu, \sqrt{2\nu})$. Similarmente ao que foi feito na seção anterior, é possível definir uma variável $\xi \sim N(0,1)$ onde

$$\xi = \frac{J(\hat{x}) - \nu}{\sqrt{2\nu}} \quad (2.67)$$

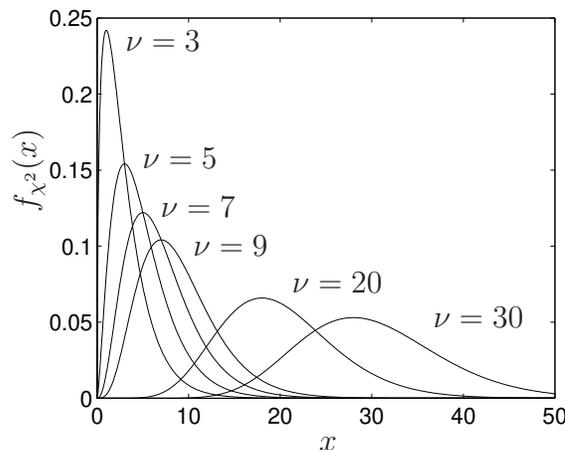


Figura 2.2: Função densidade de probabilidade χ^2 para diferentes valores de ν .

2.6 Conclusões

As principais características matemáticas do método dos mínimos quadrados foram apresentados neste capítulo. O modelo de estimação de parâmetros foi desenvolvido para as versões linear e não linear. Além disso, foram desenvolvidos os principais métodos para obtenção da solução desses problemas. Como será apresentado no capítulo 4, esse método de estimação é o mais utilizado no contexto da estimação de estado em sistemas de energia elé-

trica. Os conceitos matemáticos apresentados neste capítulo irão fomentar o desenvolvimento do método de regularização de Tikhonov, que será apresentado no capítulo 3.

Capítulo 3

Problemas Inversos e Regularização

3.1 Introdução

No campo das engenharias, os sistemas ou processos físicos são comumente descritos por modelos matemáticos, os quais são definidos pelos parâmetros do modelo ou entrada x , pelas observações ou saída z , e por um operador (linear ou não linear) $H : X \rightarrow Z$, onde, X e Z são espaços vetoriais com produto interno tal que $x \in X$ e $z \in Z$. Essas relações podem ser descritas conforme a equação 3.1.

$$H(x) = z \tag{3.1}$$

Dada a entrada x , um problema direto consiste em obter ou prever a saída z . Por essa razão H é normalmente chamado de operador direto. Note que o problema direto possui apenas uma solução. Um problema inverso, por sua vez, consiste em obter os parâmetros x dado um conjunto de observações z (Aster et al., 2005; Tarantola, 2005). Durante a solução de problemas inversos algumas questões essenciais devem ser levadas em conta:

1. Existência: Pode não haver um conjunto de parâmetros x que satisfaça o modelo 3.1. A inexistência de solução pode estar associada a erros na modelagem matemática, H , ou nas observações z .
2. Unicidade: Caso exista uma solução exata para o problema ela pode não ser única. De fato, podem existir muitas ou infinitas soluções. O problema da unicidade está relacionado com a existência do operador inverso, ou seja, H^{-1} .
3. Estabilidade. O processo de cômputo da solução de um problema inverso pode ser instável, no sentido de que pequenas alterações nas observações z podem produzir grandes alterações nos parâmetros estimados x .

Esses três aspectos são fundamentais para a compreensão das dificuldades associadas à resolução de problemas inversos. Atribui-se ao matemático Jacques Salomon Hadamard o conceito de problema bem posto. Diz-se que um problema é bem posto quando ele atende a todas as condições acima, do contrário o problema é classificado como mal posto. Um problema inverso não linear naturalmente pode apresentar múltiplas soluções, portanto, diz-se que um problema inverso não linear é mal posto, principalmente, quando ele viola a primeira e a terceira condição (Ramlau, 2004).

Conforme foi discutido no capítulo 2, normalmente, no problema de estimação 2.3 não há um conjunto de parâmetros x que satisfaça as observações z . Para contornar esse problema foi proposta a aplicação do método dos mínimos quadrados, que leva a uma solução única somente quando o operador H possui posto completo. Portanto, ao problema mal posto, isto é, sem solução, foi aplicado o método dos mínimos quadrados, que sob as condições de posto completo de H e número de condicionamento adequado, tornou-se um problema bem posto.

Entretanto, nem sempre é possível satisfazer as condições estabelecidas quanto ao condicionamento e ao posto do operador H . De fato, existem situações onde o espectro dos valores singulares de H é muito extenso, assim, o número de condição é elevado e ainda podem ocorrer grandes saltos entre valores singulares adjacentes. Situações mais complexas envolvem valores singulares nulos, evidenciando que as matrizes possuem posto incompleto. Essas deficiências resultam em instabilidade numérica e não unicidade da solução do problema de mínimos quadrados 2.6. Nesse contexto, diante de problemas numericamente mal condicionados e com posto deficiente, os métodos de regularização podem ser aplicados para se obter soluções aceitáveis e/ou estabilizar o processo de resolução do problema 3.1. A seguir são apresentados dois métodos clássicos de regularização encontrados na literatura. A SVD truncada será discutida para que se possa expor aspectos críticos associados a essas metodologias e o método de regularização de Tikhonov será apresentado pois será utilizado na proposta principal deste trabalho.

3.2 A SVD Truncada

Considere um problema inverso linear $H_w x = z_w$ onde $H_w \in \mathbb{R}^{m \times n}$ e $\text{posto}(H) = r \leq \min(m, n)$. Na seção 2.4.1 mostrou-se que uma solução x_{\dagger} pode ser obtida através do uso da inversa generalizada de Moore-Penrose. Note que a solução x_{\dagger} pode ser obtida mesmo quando $m = n$, $m > n$ ou $m < n$, como segue:

$$x_{\dagger} = H_w^{\dagger} z_w = V_r \Sigma_r^{-1} U_r^T z_w \quad (3.2)$$

Os elementos do vetor $U_r^T z_w$ correspondem aos produtos escalares entre as r primeiras colunas

de U e o vetor de observações z_w , isto é:

$$U_r^T z_w = \begin{bmatrix} u_1^T z_w \\ u_2^T z_w \\ \vdots \\ u_r^T z_w \end{bmatrix} \quad (3.3)$$

Multiplicando 3.3 à esquerda por Σ_r^{-1} obtemos

$$\Sigma_r^{-1} U_r^T z_w = \begin{bmatrix} \frac{u_1^T z_w}{\sigma_1} \\ \frac{u_2^T z_w}{\sigma_2} \\ \vdots \\ \frac{u_r^T z_w}{\sigma_r} \end{bmatrix} \quad (3.4)$$

Ao multiplicar 3.4 à esquerda por V_r nós obtemos uma combinação linear das colunas de V_r que, por sua vez, pode ser escrita como:

$$x_{\dagger} = V_r \Sigma_r^{-1} U_r^T z_w = \sum_{i=1}^r \frac{u_i^T z_w}{\sigma_i} v_i \quad (3.5)$$

A expressão da solução generalizada na forma 3.5 permite fazer uma análise da influência dos valores singulares, σ_i , na obtenção da solução. Note que, em se tratando de problemas mal postos, valores singulares muito pequenos no denominador da equação 3.5 podem produzir grandes coeficientes e, portanto, fazer com que a solução final tenha pouco ou nenhum sentido prático. Isso mostra que a inversa generalizada pode ser um amplificador dos eventuais ruídos contidos nas observações z_w . Portanto, o número de condição definido na equação 2.36 é um bom indicador de quão instável será o processo de obtenção da solução x_{\dagger} . Uma alternativa para a solução desse problema é a SVD truncada (Hansen, 1987). Esse método consiste em definir um $p < r$ e, então, a solução da equação 3.5 é obtida truncando-se os termos indesejáveis, isto é:

$$x_{\dagger} = \sum_{i=1}^p \frac{u_i^T z_w}{\sigma_i} v_i \quad (3.6)$$

Pode-se utilizar a condição discreta de Picard como uma metodologia para definir quando truncar o somatório 3.6 (Aster et al., 2005; Hansen, 1990). Essa condição garante que a solução da equação 3.5 será estável se os termos $u_i^T z_w$ decaírem à zero mais rapidamente que os valores singulares σ_i . Portanto, o ponto onde a taxa de decrescimento dos valores singulares for maior que a taxa de decrescimento dos termos $u_i^T z_w$ pode ser usado como ponto de truncamento.

Se a SVD está disponível, a solução truncada pode ser obtida rapidamente através da equação 3.6. Entretanto, como fora comentado na seção 2.4.1, na prática pode ser muito caro ou impraticável obter essa decomposição. A SVD truncada é um exemplo dos métodos de regularização. Intuitivamente, nota-se que a solução truncada, por descartar termos indesejáveis, não deve se ajustar tão bem aos dados observados em comparação com a solução obtida a partir da equação 3.5, caso fosse possível obtê-la de maneira numericamente estável. Além disso, o estimador dado pela SVD truncada é um estimador enviesado (vide seção 2.5). Portanto, durante a definição de um mecanismo de regularização é preciso fazer um balanço entre o quão estável será o processo de solução e quanto esse mecanismo irá influenciar na solução real do problema.

3.3 A Regularização de Tikhonov

No capítulo 2 utilizou-se o método dos mínimos quadrados para obter uma solução para problema inverso 3.1. Mostrou-se também que no caso linear a solução de mínimos quadrados pode ser obtida minimizando-se o seguinte índice:

$$\min_{x \in \mathbb{R}^n} J(x) = \|R_z^{-\frac{1}{2}}(z - Hx)\|_2^2 \quad (3.7)$$

Sabe-se, entretanto, que quando o operador H possui deficiência de posto, existem infinitas soluções para o problema 3.7. Para superar essa dificuldade pode-se utilizar a SVD através da inversa generalizada de Moore-Penrose. Nela uma solução particular para 3.7 pode ser obtida, muito embora essa solução possa ter pouco sentido prático. Em casos ainda mais graves pode-se utilizar a SVD truncada, descrita na seção anterior, como método de regularização.

Nos casos em que o problema 3.7 possui infinitas soluções, é possível incluir no problema original um conjunto de características desejadas, viabilizando a obtenção de uma solução conveniente. Isso pode ser feito adicionando-se um novo funcional $\Omega(x)$ ao problema original. Por exemplo, caso estejamos procurando uma solução que possua uma norma pequena, o funcional $\Omega(x)$ pode ser escrito como:

$$\Omega(x) = \|x\|_2^2 \quad (3.8)$$

Caso tenhamos alguma informação *a priori* da solução da equação 3.7, é possível definir $\Omega(x)$ tal que a solução final não se distancie muito da informação *a priori* disponível no vetor u , isto é:

$$\Omega(x) = \|u - x\|_2^2 \quad (3.9)$$

De uma forma mais geral, existe uma definição para $\Omega(x)$ onde se utiliza um operador linear sobre a diferença $u - x$,

$$\Omega(x) = \|L(u - x)\|_2^2 \quad (3.10)$$

Tipicamente L é a matriz identidade, entretanto, pode-se utilizar esse operador como derivada discreta sobre o vetor $u - x$. Por exemplo, no caso de primeira ordem L é escrito como:

$$L = \begin{bmatrix} -1 & 1 & & & \\ & -1 & 1 & & \\ & & \dots & & \\ & & & -1 & 1 \\ & & & & -1 & 1 \end{bmatrix} \quad (3.11)$$

O método de regularização de Tikhonov é um dos mais utilizados na regularização de problemas mal postos (Aster et al., 2005). Caso esteja disponível um conjunto de informações *a priori* que descreva a solução desejada, é possível introduzi-las no problema original regularizando-o. Matematicamente isso é feito reescrevendo o problema 3.7 como:

$$\min_{x \in \mathbb{R}^n} F(x) = J(x) + \lambda^2 \Omega(x) = \|R_z^{-\frac{1}{2}}(z - Hx)\|_2^2 + \lambda^2 \|L(u - x)\|_2^2 \quad (3.12)$$

Note que $\Omega(x)$ tem como função estabilizar a solução de 3.7. A variável λ é conhecida como parâmetro de regularização. Caso o parâmetro de regularização seja muito grande, o termo referente às observações, $J(x)$, terá pouca influência na solução obtida \hat{x} , no limite, $\lim_{\lambda \rightarrow \infty} \hat{x} = u$. Por outro lado, se λ for pequeno o termo $J(x)$ será dominante. Quando λ tende a zero voltamos ao problema de mínimos quadrados original. Aplicando as técnicas apresentadas no capítulo 2, a solução da equação 3.12 é obtida através do seguinte sistema de equações:

$$[H^T R_z^{-1} H + \lambda^2 L^T L]x = H^T R_z^{-1} z + \lambda^2 L^T L u \quad (3.13)$$

Note que, caso $\mathcal{N}(H_w) \cap \mathcal{N}(L) = \{0\}$, isto é, o espaço nulo da matriz de coeficientes é trivial, a solução das equações acima é única.

O problema de otimização 3.12 pode ser reescrito da seguinte maneira

$$\min_{x \in \mathbb{R}^n} \left\| \begin{pmatrix} z_w \\ u \end{pmatrix} - \begin{pmatrix} H_w \\ \lambda L \end{pmatrix} x \right\|_2^2 \quad (3.14)$$

Fazendo $A^T = [H_w^T \ \lambda L^T]$, $b^T = [z_w^T \ u^T]$ e $Ax = b$ é fácil verificar que os métodos desenvolvidos na seção 2.4 podem ser aplicados diretamente ao problema regularizado 3.12.

Considere o caso onde $L = I$ e $u = 0$. Reescrevendo as equações 3.13 temos

$$[H^T R_z^{-1} H + \lambda^2 I]x = H^T R_z^{-1} z \quad (3.15)$$

Aplicando a decomposição SVD na matriz H_w as equações 3.15 podem ser reescritas como:

$$[V \Sigma_n^2 V^T + \lambda^2 I]x = V \Sigma^T U^T z_w \quad (3.16)$$

valendo-se do fato de que $VV^T = I$, é possível mostrar que a solução das equações 3.16 pode ser obtida por:

$$x_\lambda = \sum_{i=1}^n \frac{\sigma_i^2}{\sigma_i^2 + \lambda^2} \frac{u_i^T z_w}{\sigma_i} v_i \quad (3.17)$$

onde o termo

$$\frac{\sigma_i^2}{\sigma_i^2 + \lambda^2} = \hat{f}_i \quad (3.18)$$

é chamado de fator de filtro. Note que quando $\sigma_i \gg \lambda$, $\hat{f}_i \approx 1$ e $\sigma_i \ll \lambda$, $\hat{f}_i \approx 0$. Portanto, o parâmetro de regularização deve ser ajustado de modo que a solução seja fortemente dominada pelos dados observados e os efeitos negativos provocados por valores singulares muito pequenos sejam contornados (Aster et al., 2005).

3.3.1 Regularização de Tikhonov aplicada ao problema não linear

Nas seções anteriores foi considerado apenas o problema inverso linear, consideraremos agora o caso onde o problema inverso tem a forma $h(x) = z$, onde $h(x) : \mathbb{R}^n \rightarrow \mathbb{R}^m$ é um vetor de funções não lineares. Para aplicar o método de regularização de Tikhonov o problema de otimização 3.12 é reescrito da seguinte forma:

$$\min_{x \in \mathbb{R}^n} F(x) = J(x) + \lambda^2 \Omega(x) = [z - h(x)]^T R_z^{-1} [z - h(x)] + \lambda^2 [u - x]^T L^T L [u - x] \quad (3.19)$$

A resolução deste problema pode ser feita utilizando técnicas de otimização não linear.

Aplicando, por exemplo, o método de Gauss-Newton, desenvolvido na seção 2.3, linearizamos $h(x)$ e x em torno de x^ν , então, o sistema 3.19 é substituído por:

$$\min_{\Delta x^\nu \in \mathbb{R}^n} F(\Delta x^\nu) = [\Delta z^\nu - H(x^\nu)\Delta x^\nu]^T R_z^{-1} [\Delta z^\nu - H(x^\nu)\Delta x^\nu] + \lambda^2 [\Delta u^\nu - \Delta x^\nu]^T L^T L [\Delta u^\nu - \Delta x^\nu] \quad (3.20)$$

onde $\Delta u^\nu = u - x^\nu$ e $\Delta z^\nu = z - h(x^\nu)$. Note que o problema de otimização 3.20 é um problema de mínimos quadrados lineares onde Δx^ν é o vetor de parâmetros a serem estimados. A solução do problema acima pode ser obtida por inspeção a partir da equação 3.13. Logo, a solução regularizada do problema inverso não linear pode ser obtida pelo seguinte processo iterativo:

$$[H(x^\nu)^T R_z^{-1} H(x^\nu) + \lambda^2 L^T L] \Delta x^\nu = H(x^\nu)^T R_z^{-1} [z - h(x^\nu)] + \lambda^2 L^T L [u - x^\nu] \quad (3.21)$$

$$x^{\nu+1} = x^\nu + \Delta x^\nu$$

É possível adotar os mesmos critérios de parada apresentados na seção 2.3.1.

O ajuste adequado do parâmetro de regularização e a disponibilidade de informações *a priori* de boa qualidade, geralmente, são suficientes para estabilizar o processo de resolução do problema 3.19. Entretanto, em casos mais críticos é possível adotar, por exemplo, a busca unidimensional desenvolvida na seção 2.3.2 juntamente com o método de Gauss-Newton. A seguir são apresentados alguns métodos para escolha do parâmetro de regularização.

3.3.2 Métodos para escolha do parâmetro de regularização

Todo método de regularização deve possuir um esquema de escolha do seu parâmetro de regularização. Na SVD truncada foi mostrado que o parâmetro de regularização p pode ser escolhido a partir da condição discreta de Picard. No esquema de regularização de Tikhonov o parâmetro de regularização é a variável contínua λ . Nessa seção serão apresentados alguns dos esquemas mais conhecidos para a escolha do parâmetro de regularização de Tikhonov aplicados a problemas lineares.

Na seção 2.1 mostrou-se que o lado direito da equação 3.1 pode ser escrito como $z = z^{real} + e$. Onde e representa o ruído associado ao valor verdadeiro da variável observada, z^{real} . Nesse sentido, os esquemas de escolha do parâmetro de regularização podem ser divididos em duas categorias dependendo do tratamento dado ao vetor de ruídos e (Hansen, 1998).

1. Esquemas baseados no conhecimento, ou boa estimativa da norma dos ruídos, isto é,

$\|e\|_2$;

2. Esquemas que não requerem $\|e\|_2$ diretamente, mas tentam buscar essa informação a partir do vetor de observações z .

O princípio da discrepância

O princípio da discrepância, normalmente atribuído a Morozov (1984), pode ser considerado como o método que se baseia na norma do erro, $\|e\|_2$, mais utilizado. Caso o problema 3.1 seja consistente no sentido de que ele pode ser escrito como $Hx^{real} = z^{real}$, então, o método deve escolher o parâmetro λ tal que a norma do vetor de resíduos seja igual a um limite superior δ_e para $\|e\|_2$ (Hansen, 1998). Assim,

$$\|Hx_\lambda - z\|_2 = \delta_e, \text{ onde } \|e\|_2 \leq \delta_e \quad (3.22)$$

Um algoritmo básico para aplicar esse princípio consiste em, dada uma estimativa inicial para o parâmetro de regularização, λ^0 , gerar uma sequência de soluções x_λ , através do método de Tikhonov, até que a condição 3.22 seja atendida. Note que a decomposição SVD truncada também poderia ser utilizada se considerarmos $\|Hx_p - z\|_2 \leq \delta_e$. Na referência (Morozov, 1984) podem ser encontrados outros métodos para solucionar 3.22.

A aplicação do princípio da discrepância de Morozov só é possível caso exista uma boa estimativa para $\|e\|_2$. Uma das principais considerações feitas no capítulo 2 é que os ruídos e , são independentes e possuem distribuição normal com média nula, isto é, $e \sim N(0, R_z)$. A partir dos desenvolvimentos feitos na seção 2.5 pode-se mostrar que $E[\|e\|_2^2] = \text{tr}(R_z)$, assim, uma estimativa bastante razoável para $\|e\|_2$ pode ser obtida por $E[\|e\|_2] = [\text{tr}(R_z)]^{\frac{1}{2}}$.

Validação cruzada generalizada

A validação cruzada generalizada é um método bastante popular para a escolha do parâmetro de regularização λ que dispensa o conhecimento da norma do ruído $\|e\|_2$. Esse método é baseado na consideração estatística de que uma solução x_λ com um bom parâmetro de regularização deve predizer alguma eventual observação perdida z_i (Hansen, 1998). Assim, calcula-se a solução de Tikhonov x_λ^i quando a medida z_i é retirada do conjunto z , isto é:

$$x_\lambda^i = [(H^i)^T R_z^{-1} (H^i) + \lambda^2 I]^{-1} (H^i)^T R_z^{-1} z^i \quad (3.23)$$

Onde H^i representa a matriz H sem a i -ésima linha e z^i o vetor z sem o i -ésimo elemento. Logo, o parâmetro de regularização pode ser escolhido de modo que o erro de predição seja o menor possível, isto é:

$$Hx_\lambda^i = \hat{z} \approx z \quad (3.24)$$

Uma alternativa é minimizar o erro médio quadrático de $\hat{z} - z$:

$$\min_{\lambda \in \mathbb{R}} \frac{1}{m} \sum_{i=1}^m (H(i, :)x_\lambda^i - z_i)^2 \quad (3.25)$$

O problema apresentado acima é conhecido como validação cruzada deixa um fora. Infelizmente o problema de otimização 3.25 requer a solução de m sistemas de equações de Tikhonov 3.23. A validação cruzada generalizada é uma alternativa para esse problema, em última análise ela consiste em encontrar a solução para o seguinte problema de otimização:

$$\min_{\lambda \in \mathbb{R}} \frac{m \|H_w x_\lambda - z_w\|_2^2}{\text{tr}(I - H_w H_w^\#)^2} \quad (3.26)$$

Onde $H_w^\#$ representa a inversa regularizada de H_w , isto é, $H_w^\# = [H^T R_z^{-1} H + \lambda^2 I]^{-1} H_w^T$. Mais detalhes desse método aplicado a problemas inversos lineares podem ser encontrados em (Aster et al., 2005; Hansen, 1998) enquanto que um desenvolvimento para problemas inversos não lineares pode ser encontrado em (Santos, 1995).

A curva L

Ao definirmos a regularização de Tikhonov foi mencionado que estávamos procurando um solução particular, dentre um conjunto de soluções factíveis, para o problema de otimização 3.7. Segundo a formulação 3.13 pode-se dizer que a solução x_λ deve ser tal que $\Omega(x_\lambda) \leq \epsilon$, onde ϵ é um limite superior para a qualidade da solução procurada. Aplicando o princípio da discrepância foi possível definir um limite superior para a norma dos resíduos, ou seja, $J(x_\lambda) \leq \delta$. Essas restrições embutidas na equação 3.13 implicam que o parâmetro λ deve satisfazer, simultaneamente, as condições $J(x_\lambda) \leq \delta$ e $\Omega(x_\lambda) \leq \epsilon$. Portanto, qualquer x_λ que atenda estas restrições é uma solução aceitável para o problema inverso.

Um resultado interessante é que ao plotar uma curva em escala logarítmica dos valores $\Omega(x)$ versus $J(x)$ parametrizados por λ o resultado é uma curva em forma de um L. Isso acontece, pois ao aumentar λ o funcional $\Omega(x)$ decresce e $J(x)$ cresce (Aster et al., 2005). A figura 3.1 representa uma curva L genérica com os limites desejáveis para as soluções.

Baseados no formato e nas características da curva L existe outro método bastante popular para escolha do parâmetro de regularização. Esse método foi proposto por Hansen (1992), consistindo em encontrar o parâmetro de regularização λ referente ao ponto da curva com a maior curvatura. Normalmente esse ponto se encontra no canto inferior da curva.

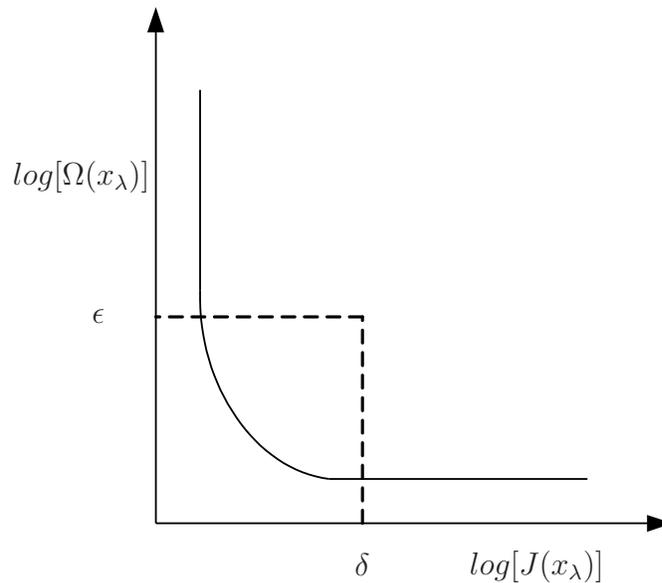


Figura 3.1: Curva L genérica.

Intuitivamente, esse canto sugere um balanço entre os termos da equação 3.13.

O uso da distribuição χ^2

Na seção 2.5.5 mostrou-se que $J(\hat{x})$ possui distribuição de probabilidade χ^2 com $m - n$ graus de liberdade, $E[J(\hat{x})] = m - n$ e $Var[J(\hat{x})] = 2(m - n)$. Essas propriedades estatísticas de $J(\hat{x})$ podem ser usadas no desenvolvimento de uma metodologia para o ajuste de λ , que é similar ao princípio da discrepância. Essa metodologia consiste em obter uma solução x_λ tal que $J(x_\lambda)$ seja igual a um limite superior δ_{m-n} para $m - n$, isto é:

$$J(x_\lambda) = \|R_z^{-\frac{1}{2}}(z - Hx_\lambda)\|_2^2 = \delta_{m-n}, \text{ onde } m - n \leq \delta_{m-n} \quad (3.27)$$

No entanto, a distribuição χ^2 só é definida para $m - n \geq 1$, logo, nem sempre é possível utilizar essa metodologia.

Considere que as informações *a priori*, u , introduzidas pela regularização de Tikhonov, também sejam variáveis aleatórias normalmente distribuídas, independentes entre si e com relação às medidas. Sob essas circunstâncias, analogamente ao índice $J(\hat{x})$, a função objetivo $F(\hat{x})$ do problema regularizado 3.12 também possui distribuição de probabilidade χ^2 . Nesse caso, o número de graus de liberdade da distribuição será $m - n + l$, onde l é o número de linhas não nulas de L . Note que quando $L = I$ o número de graus de liberdade é m . Note ainda que na equação 3.19 o termo $\lambda^{-2}I$ representa a matriz de covariância do vetor u . Portanto, pode-se reescrever a equação 3.27 para $F(x_\lambda)$ da seguinte maneira,

$$F(x_\lambda) = \|R_z^{-\frac{1}{2}}(z - Hx_\lambda)\|_2^2 + \lambda^2 \|L(u - x_\lambda)\|_2^2 = \delta_{m-n+l}, \text{ onde } m - n + l \leq \delta_{m-n+l} \quad (3.28)$$

Uma metodologia alternativa que também utiliza a distribuição χ^2 foi proposta por Mead (2008). Quando o número de graus de liberdade, $\nu = m - n + l$, é grande, a distribuição χ^2 tende a uma distribuição normal com média ν e variância 2ν . Portanto, a transformação $\xi = \frac{F(x_\lambda) - \nu}{\sqrt{2\nu}}$ possui distribuição normal padronizada. Considere a seguinte probabilidade:

$$P(-\gamma_{\frac{\alpha}{2}} \leq \xi \leq \gamma_{\frac{\alpha}{2}}) = 1 - \alpha$$

substituindo e rearranjando os termos, temos:

$$P(\nu - \sqrt{2\nu}\gamma_{\frac{\alpha}{2}} \leq F(x_\lambda) \leq \nu + \sqrt{2\nu}\gamma_{\frac{\alpha}{2}}) = 1 - \alpha$$

A constante $\gamma_{\frac{\alpha}{2}}$ pode ser obtida partir de:

$$\int_{-\infty}^{\gamma_{\frac{\alpha}{2}}} f_N(\xi) d\xi = 1 - \frac{\alpha}{2}$$

onde f_N representa a função densidade de probabilidade normal padronizada e α é o nível de significância. Dessa forma, com a confiança de $1 - \alpha$, pode-se criar o seguinte intervalo:

$$\nu - \sqrt{2\nu}\gamma_{\frac{\alpha}{2}} \leq F(x_\lambda) \leq \nu + \sqrt{2\nu}\gamma_{\frac{\alpha}{2}} \quad (3.29)$$

Por exemplo, escolhendo $\alpha = 0,05$ temos um intervalo com 95% de confiança dado por:

$$\nu - 2,7719\sqrt{\nu} \leq F(x_\lambda) \leq \nu + 2,7719\sqrt{\nu}$$

Nesse sentido, a metodologia proposta em (Mead, 2008) consiste em, dado um α , encontrar um λ tal que a desigualdade 3.29 seja satisfeita. O funcional $F(x_\lambda)$ pode ser escrito, com x_λ implícito, da seguinte maneira:

$$F(\lambda) = [z - Hu]^T [\lambda^{-2}H(L^T L)^{-1}H^T + R_z]^{-1} [z - Hu] \quad (3.30)$$

Portanto, o parâmetro de regularização que satisfaz a desigualdade 3.29 é computado a partir da equação 3.30. Repare que essa expressão só é válida quando L possui posto completo de colunas.

3.4 Conclusões

Neste capítulo foram apresentados métodos de regularização para a solução de problemas numericamente mal condicionados e com posto deficiente. Dentre esses métodos, a regularização de Tikhonov foi adotada na concepção do estimador regularizado desenvolvido nesta dissertação. Para garantir o bom funcionamento dos métodos de regularização ficou claro que deve haver um bom mecanismo de escolha do parâmetro de regularização. Alguns mecanismos de escolha desse parâmetro foram apresentados. O estudo desses mecanismos mostrou que eles apresentam nível de complexidade e carga computacional inadequados para o problema de estimação de estado. Nesse contexto, foi adotado um esquema simplificado de seleção do parâmetro de regularização visando a obtenção de soluções dominadas pelos dados observados e não pelas informações utilizadas para regularizar o problema. O parâmetro de regularização deve permitir ainda que o estimador apresente boas características de convergência. A seleção do parâmetro de regularização é retomada no Capítulo 5, onde apresenta-se o estimador regularizado desenvolvido.

Capítulo 4

Revisão Sobre a Estimação de Estado Convencional

4.1 Introdução

O estado de um sistema pode ser definido como o conjunto mínimo de variáveis necessárias para caracterizar a situação física do sistema em um determinado instante de tempo. Em sistemas de energia elétrica o estado é tipicamente formado pelos módulos e ângulos dos fasores de tensão em todas as barras do sistema. Entretanto, outros conjuntos de variáveis, tais como as partes reais e imaginárias dos fasores de tensão ou as correntes complexas nos ramos da rede também podem ser adotadas como variáveis de estado.

Para determinar os valores das variáveis de estado em um determinado instante, os estimadores utilizam um conjunto redundante de medidas. Assim, os eventuais ruídos presentes nas medidas podem ser filtrados e o estado mais provável do sistema pode ser encontrado segundo algum critério estatístico estabelecido. Nos sistemas modernos podem existir, além das medições convencionais de potência e magnitude de tensão, medições fasoriais sincronizadas de corrente elétrica e tensão. Nesse contexto, como os sistemas de energia elétrica, geralmente, possuem uma dinâmica lenta, admite-se que para uma janela de tempo suficientemente pequena o sistema opera em regime estacionário e, portanto, medidas coletadas dentro dessa janela podem ser usadas para estimar o estado da rede. Esse procedimento é chamado de estimação estática de estado (Freitas, 1995).

Como já citado, normalmente o estado de um sistema de energia elétrica é definido pelos fasores de tensão nas barras. Logo, a topologia e os parâmetros dos elementos da rede devem ser conhecidos e considerados corretos. Entretanto, sabe-se, por exemplo, que a topologia da rede pode variar durante a operação. Nesse contexto, o estimador de estado generalizado proposto por A. Monticelli, apresentado detalhadamente em (A. Abur, 2004;

Monticelli, 1999), permite incluir o estado de chaves/disjuntores no conjunto de variáveis de estado e, portanto, estimar a topologia da rede ao invés de simplesmente considerá-la correta.

Neste capítulo é feita uma breve revisão dos aspectos principais da estimação de estado em sistemas de energia elétrica. Primeiramente o modelo de estimação de estado é apresentado. Em seguida, são discutidos os subproblemas de análise de observabilidade e o processamento de erros grosseiros. No final do capítulo é apresentado um estudo de caso.

4.2 O Estimador de Estado Convencional

A maioria dos estimadores de estado implementados nos centros de controle baseiam-se no método dos mínimos quadrados ponderados (Freitas, 1995). No capítulo 2 foi mostrado que na resolução de um problema de mínimos quadrados não linear deve-se encontrar um vetor de parâmetros x que minimiza a função de custo $J(x)$. No contexto da estimação de estado em sistemas de energia elétrica, o vetor de parâmetros é chamado de vetor de estado. Considerando um sistema com N barras o vetor de estado, x , é composto por todos os fasores de tensão das barras, que na forma polar é expresso por:

$$x = [\theta_2 \ \theta_3 \ \cdots \ \theta_N \ V_1 \ V_2 \ V_3 \ \cdots \ V_N]^T$$

onde θ_i e V_i representam o ângulo de fase e a magnitude da tensão da i -ésima barra, respectivamente. Devido às leis que regem o comportamento dos circuitos elétricos, é necessário especificar o valor de uma variável de ângulo para ser a referência angular, logo, o número de variáveis de estado será igual a $2N - 1$. Quando há medidas fasoriais sincronizadas de tensão não é necessário incluir a referência angular para tornar o processo de cálculo factível (Zhu and Abur, 2007).

Dada a janela de tempo em que será realizada a estimação de estado, os valores medidos dentro dessa janela são armazenados em um vetor z de dimensão m . Conforme estabelecido no capítulo 2, admite-se que essas medidas são variáveis aleatórias independentes com distribuição normal, com variâncias conhecidas e os ruídos de medição possuem média nula. Logo, o vetor de medidas pode ser escrito como:

$$z = z^{real} + e$$

e a matriz de covariância é dada por:

$$R_z = \text{diag} [\sigma_1^2 \ \sigma_2^2 \ \cdots \ \sigma_m^2]$$

onde e é o vetor de ruídos nas medições e z^{real} é o vetor dos valores verdadeiros da variáveis

observadas. A variância da i -ésima medida é denotada por σ_i^2 .

Para cada medida z_i é possível escrever uma função $h_i(x)$ que descreve a medida em função das variáveis de estado. Logo, é possível encontrar o vetor de estado resolvendo o seguinte sistema sobredeterminado de equações não lineares:

$$h(x) = z$$

Todavia, sabe-se que devido aos ruídos e esse sistema de equações não admite solução. Logo, aplicando-se o método dos mínimos quadrados ponderados, o vetor de estado, \hat{x} , pode ser estimado resolvendo-se o seguinte problema de otimização:

$$\min_{x \in \mathbb{R}^n} J(x) = [z - h(x)]^T R_z^{-1} [z - h(x)] \quad (4.1)$$

onde $n = 2N - 1$.

A solução do problema acima, \hat{x} , pode ser obtida iterativamente pelo método de Gauss-Newton, isto é:

$$H(x^\nu)^T R_z^{-1} H(x^\nu) \Delta x^\nu = H(x^\nu)^T R_z^{-1} [z - h(x^\nu)] \quad (4.2)$$

$$x^{\nu+1} = x^\nu + \Delta x^\nu$$

O termo $H(x)^T R_z^{-1} H(x)$ é conhecido como matriz ganho ou matriz de informação $G(x) \in \mathbb{R}^{n \times n}$. Do ponto de vista computacional, a matriz G é estrutural e numericamente simétrica além de ser altamente esparsa. Apesar de os termos de R_z^{-1} terem um caráter estatístico intrínseco, frequentemente eles são tratados apenas como fatores de ponderação. O termo $H(x)$ é a matriz Jacobiana do vetor $h(x)$, estruturalmente

$$H(x) = \begin{bmatrix} \frac{\partial h_1(x)}{\partial x_1} & \frac{\partial h_1(x)}{\partial x_2} & \cdots & \frac{\partial h_1(x)}{\partial x_n} \\ \frac{\partial h_2(x)}{\partial x_1} & \frac{\partial h_2(x)}{\partial x_2} & \cdots & \frac{\partial h_2(x)}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial h_m(x)}{\partial x_1} & \frac{\partial h_m(x)}{\partial x_2} & \cdots & \frac{\partial h_m(x)}{\partial x_n} \end{bmatrix} \quad (4.3)$$

As equações normais 4.2 podem ser solucionadas diretamente através da fatoração de Cholesky, desenvolvida na seção 2.4.2. Sabe-se que as equações normais, por construção,

possuem condicionamento numérico ruim, todavia, tendo em vista que os computadores atuais possuem boa representação numérica, em uma grande quantidade de casos a solução de mínimos quadrados pode ser obtida através das equações normais. Não obstante, existem fatores que deterioram fortemente o desempenho das equações normais, dentre esses podemos destacar:

- Uso de ponderações muito elevadas para medidas virtuais;
- A presença de linhas longas e curtas conectadas ao mesmo barramento;
- A disponibilidade de muitas medidas de injeção.

Além disso, destaca-se que a matriz ganho, presente nas equações normais, é significativamente menos esparsa que a matriz Jacobiana. Um estudo detalhado descrevendo os problemas associados às equações normais pode ser encontrado em (Monticelli et al., 1985).

Para contornar tais problemas numéricos é possível adotar os estimadores numericamente robustos. O método da fatoração ortogonal, apresentado na seção 2.4.4, normalmente é uma boa alternativa. Os primeiros estimadores ortogonais aplicados a sistemas de energia foram propostos por Simoes-Costa and Quintana (1981a,b). Em (Monticelli et al., 1985) foi proposta uma versão híbrida do estimador ortogonal. Além dos estimadores ortogonais existem outros métodos numericamente robustos. Dentre eles, talvez os mais populares sejam o método das equações normais com restrições de igualdade e o método da matriz aumentada ou *tableau* de Hachtel. Em (Holten et al., 1988) pode ser encontrado um estudo comparativo bastante amplo dos estimadores numericamente robustos.

4.3 Observabilidade de Sistemas de Energia Elétrica

A viabilidade de se realizar a estimação do estado de um sistema de energia elétrica depende da quantidade, do tipo e da localização das medidas disponíveis no sistema. Caso as medidas sejam suficientes para tornar a estimação do estado de todo o sistema possível, o sistema é dito observável (Monticelli and Wu, 1985a). Quando o modelo barra/ramo é adotado para a rede, as seguintes definições são aplicáveis (Monticelli, 1999):

- Definição 1: Uma ilha física é uma parte conexa da rede onde os ramos representam linhas de transmissão, transformadores e capacitores série, e as barras representam os nós;
- Definição 2: Uma ilha observável é uma ilha física onde os fluxos em todos os ramos podem ser estimados a partir das medidas disponíveis, independentemente do valor adotado para a referência angular.

A função de análise de observabilidade deve ter a capacidade de verificar se a rede é observável e, caso não seja, identificar as ilhas observáveis. Portanto, ela é uma função que deve ser executada em tempo real antes da função de estimação de estado. Além disso, a função de análise de observabilidade também é utilizada nas etapas de planejamento e projeto de sistemas de medição.

4.3.1 Observabilidade algébrica e observabilidade numérica

A solução de mínimos quadrados do estimador de estado é obtida iterativamente através das equações normais 4.2. Para que a solução das equações normais a cada iteração seja única é preciso que a matriz ganho G possua posto completo, isto é, $\text{posto}(G) = n$. Tendo em vista que R_z^{-1} é positiva definida, então, $\text{posto}(G) = \text{posto}(H) = n$. Como $H \in \mathbb{R}^{m \times n}$ e como o $\text{posto}(H) \leq \min(m, n)$, tem-se que $m \geq n$. Portanto, dado um conjunto de medição diz-se que um sistema de energia é algebricamente observável se a matriz Jacobiana H possui posto igual a n .

Como o procedimento de obtenção da solução de mínimos quadrados é iterativo, dependendo do estado no ponto de cômputo da matriz Jacobiana o seu posto pode ser alterado. Portanto, diz-se que um sistema de energia elétrica é numericamente observável se for possível obter a solução de mínimos quadrados, inicializando adequadamente o processo iterativo, por exemplo, com o perfil plano de tensões, isto é, 1,0 pu para magnitudes de tensão e 0 rad para os ângulos de fase das tensões.

Os algoritmos numéricos de análise de observabilidade não calculam diretamente o posto das matrizes. Em geral, eles utilizam técnicas de fatoração e cálculos em ponto flutuante para avaliar as matrizes ganho ou Jacobiana extraíndo informações sobre a observabilidade do sistema.

Como normalmente as medições das potências são realizadas ao pares, isto é, para cada medida ativa P há uma medida reativa Q , é possível desacoplar as variáveis de ângulo e magnitude de tensão linearizando a análise de observabilidade que pode, então, ser feita utilizando somente o subsistema linearizado $P-\theta$ (A. Abur, 2004; Monticelli and Wu, 1985b). Caso as medidas não sejam realizadas aos pares, é possível analisar a observabilidade a partir dos dois subsistemas linearizados $P-\theta$ e $Q-V$. Em sistemas de subtransmissão e distribuição existe uma presença bastante significativa de medidas de magnitude de corrente. Nesses casos não é possível realizar o desacoplamento entre as variáveis de ângulo e magnitude de tensão, portanto, faz-se necessária uma análise mais cuidadosa do problema (Abur and Gomez-Exposito, 1997).

4.3.2 Observabilidade topológica

A observabilidade topológica relaciona o conjunto de medições disponíveis com o grafo da rede elétrica, onde as arestas do grafo representam os ramos e os vértices representam as barras do sistema. A ideia geral dos algoritmos de análise de observabilidade topológica consiste em tentar criar uma árvore geradora no grafo da rede atribuindo as medidas de fluxo aos ramos onde elas são realizadas e as medidas de injeção aos ramos adjacentes às barras onde elas são alocadas. Caso uma árvore geradora seja criada diz-se que a rede é observável (Krumpholz et al., 1980). Naturalmente esses algoritmos acabam caindo na classe de problemas combinatoriais. Em geral esses algoritmos utilizam o modelo linearizado $P - \theta$ da rede elétrica.

4.4 Processamento de Erros Grosseiros

As medidas realizadas no sistema elétrico e disponibilizadas no centro de controle naturalmente contêm ruídos. Esses ruídos aparecem devido a precisão finita dos medidores e por interferências no canal de comunicação entre o centro de controle e o ponto de medição. Caso exista um nível de redundância adequado, o estimador de estado tem a capacidade de filtrar esses ruídos, e a eficiência dessa filtragem depende do critério estatístico utilizado no processo de estimação. Eventualmente, algumas medidas podem estar contaminadas por erros relativamente grandes. Quando isso ocorre, diz-se que essas medidas contêm erros grosseiros. Esse tipo de erro pode surgir, por exemplo, devido a falhas nos medidores, conexão errada nos instrumentos de medição ou mesmo durante a transmissão dos dados. Além dos erros grosseiros em medidas, podem surgir também erros topológicos. Esse tipo de erro ocorre com menor frequência, entretanto, pode causar grande impacto no processo de estimação, podendo levar até à não convergência do estimador. O uso da estimação de estado generalizada (Monticelli, 1999), que possui a capacidade de modelar elementos de impedância nula como chaves e disjuntores, é uma boa alternativa para lidar com erros topológicos.

Algumas medidas claramente contêm erros grosseiros que podem ser identificados na fase de pré-filtragem, antes da estimação propriamente dita. A pré-filtragem é realizada por regras simples, por exemplo, comparando-se os valores medidos com os valores nominais e esperados para as grandezas, ou então comparando os fluxos que chegam e deixam os nós (A. Abur, 2004). Todavia, esse procedimento de pré-filtragem não consegue detectar e identificar todos os erros grosseiros presentes no conjunto de medição. Por isso, a análise de erros grosseiros é requerida. No caso da estimação estática de estado essa análise é feita após o processo de estimação.

4.4.1 Classificação dos erros

Os erros grosseiros podem aparecer de diferentes maneiras, dependendo do tipo, da localização e do número de medidas contaminadas. Normalmente eles são classificados como se segue:

- Erros simples: Apenas uma medida possui erro grosseiro;
- Erros múltiplos: Mais de uma medida possui erro grosseiro;

 Não interativos: Erros em medidas com baixa correlação estatística, isto é, podem ser analisados separadamente;

 Interativos: Erros em medidas com elevada correlação estatística;

 Não Conformativos: Existe correlação entre as medidas, entretanto, os erros não concordam entre si;

 Conformativos: Existe correlação entre as medidas e os erros concordam entre si;

4.4.2 Classificação das medidas

Além das características estatísticas e dos valores das medidas disponíveis nos sistemas de energia elétrica, o grau de redundância do conjunto de medições influencia no processo de estimação de estado e no processamento de erros grosseiros. Normalmente as medidas classificam-se da seguinte maneira quanto ao grau de redundância:

- Medida crítica: Medida que ao ser removida torna o sistema não observável;
- Medida redundante: Medida que não é crítica;
- Par crítico: Formado por duas medidas redundantes. Se uma medida é eliminada a outra torna-se crítica;
- Conjunto crítico de ordem k : Formado por k medidas redundantes. Se uma medida é eliminada as outras tornam-se críticas.

4.4.3 Detecção e identificação de erros grosseiros

A detecção consiste em determinar se há ou não medidas contendo erros grosseiros e a identificação busca encontrar as medidas corrompidas. Os processos de detecção e identificação de erros grosseiros dependem da configuração do sistema de medição, dos valores medidos e de suas propriedades estatísticas. Um erro grosseiro não pode ser detectado se a

medida corrompida for crítica. Se a medida corrompida pertence a um conjunto crítico, o erro poderá ser detectado, mas a medida corrompida não poderá ser identificada (A. Abur, 2004). Portanto, o sucesso do processamento de erros grosseiros depende do grau de redundância do conjunto de medições.

Na estimação estática de estado, a detecção e a identificação de erros grosseiros são realizadas após o processo de estimação do estado. As metodologias mais tradicionais para o tratamento de erros são baseadas no índice $J(\hat{x})$ e nos resíduos de estimação normalizados.

Detecção utilizando o índice $J(\hat{x})$

Na seção 2.5.5 foi visto que, desde que os erros de medição tenham distribuição normal, o índice $J(\hat{x})$ é uma variável aleatória com distribuição de probabilidade χ^2 com $m - n$ graus de liberdade, portanto, $E[J(\hat{x})] = m - n$ e $Var[J(\hat{x})] = 2(m - n)$. Essa distribuição está definida para o intervalo $[0, +\infty)$, não obstante, a função $J(x)$, computada em \hat{x} , retorna apenas um valor dessa distribuição. Frequentemente estamos interessados em obter a probabilidade de $J(\hat{x})$ pertencer a um dado intervalo, isto é:

$$P(J(\hat{x}) \leq C) = \int_0^C f_{\chi^2}(J(\hat{x})) dJ(\hat{x})$$

Essa verificação permite determinar se de fato $J(\hat{x})$ pertence a uma distribuição χ^2 hipotética. Considere um teste de hipóteses onde H_0 é a hipótese nula e H_1 a hipótese alternativa:

$$H_0 : E[J(\hat{x})] = m - n;$$

$$H_1 : E[J(\hat{x})] \neq m - n.$$

Os resultados do teste podem ser definidos como:

- se $J(\hat{x}) > C$, então, rejeite a hipótese H_0 ;
- se $J(\hat{x}) \leq C$, então, aceite a hipótese H_0 .

onde C é uma constante a ser determinada. Imagine que $P(J(\hat{x}) \leq C) = 1 - \alpha$, onde α é o nível de significância ou probabilidade de falso alarme, logo C pode ser obtido por:

$$\int_0^C f_{\chi^2}(J(\hat{x})) dJ(\hat{x}) = 1 - \alpha$$

Na prática o valor escolhido para α situa-se entre 0,01 e 0,05, note que $P(J(\hat{x}) > C) = \alpha$. Considerando que H_0 representa a hipótese de não existência de erros grosseiros, pode-se

concluir que:

- se $J(\hat{x}) > C$, existem erros grosseiros;
- se $J(\hat{x}) \leq C$, não existem erros grosseiros.

Detecção e identificação utilizando o máximo resíduo normalizado

O método do índice $J(\hat{x})$ permite apenas a detecção de erros grosseiros. Alternativamente, um método que possui bastante sucesso para detecção e identificação, exceto para erros múltiplos conformativos, é o método do máximo resíduo normalizado. Os resíduos estimados são definidos como:

$$\hat{r} = z - h(\hat{x}) = z - \hat{z} \quad (4.4)$$

Por outro lado, o resíduo normalizado para a i -ésima medida é definido segundo a equação abaixo,

$$\hat{r}_i^N = \frac{\hat{r}_i}{\sqrt{R_{\hat{r}_{i,i}}}} \quad (4.5)$$

onde $R_{\hat{r}_{i,i}}$ é o i -ésimo elemento da diagonal da matriz de covariância dos resíduos estimados, isto é:

$$R_{\hat{r}} = R_z - H(\hat{x})[H(\hat{x})^T R_z^{-1} H(\hat{x})]^{-1} H(\hat{x})^T \quad (4.6)$$

note que o termo $[H(\hat{x})^T R_z^{-1} H(\hat{x})]^{-1}$, isto é, a inversa da matriz ganho G , é a matriz de covariância do estado estimado, \hat{x} .

Na seção 2.5.4 mostrou-se que, desde que os erros de medição tenham distribuição normal, $\hat{r}_i^N \sim N(0,1)$. Em (Monticelli and Garcia, 1983) mostra-se que, dado um conjunto de medidas onde somente uma delas possui ruído, então, o maior resíduo normalizado estará associado à medida corrompida. Logo, considere que \hat{r}_i^N corresponde ao máximo resíduo normalizado. Similarmente ao que fora feito na seção anterior, pode-se verificar se realmente o máximo resíduo normalizado possui distribuição normal padronizada. Para isso considera-se duas hipóteses:

$$H_0 : E[\hat{r}_i^N] = 0;$$

$$H_1 : E[\hat{r}_i^N] \neq 0.$$

Os resultados do teste podem ser definidos como:

- se $|\hat{r}_i^N| > C$, então, rejeite a hipótese H_0 ;
- se $|\hat{r}_i^N| \leq C$, então, aceite a hipótese H_0 .

a constante C , para um determinado nível de significância α , pode ser obtida a partir de:

$$\int_{-C}^C f_N(\hat{r}_i^N) d\hat{r}_i^N = 1 - \alpha$$

onde f_N representa a função densidade de probabilidade normal padronizada. Note que $P(-C > \hat{r}_i^N > C) = \alpha$. A equação acima pode ser reescrita da seguinte maneira:

$$\int_{-\infty}^C f_N(\hat{r}_i^N) d\hat{r}_i^N = 1 - \frac{\alpha}{2}$$

normalmente essa integral é definida em tabelas para diferentes valores de C , logo, não é preciso obter a expressão analítica para essa integral. Por exemplo, para $\alpha = 0,01$, $C = 2,57$ e para $\alpha = 0,0027$, $C = 3$. Considerando que H_0 representa a hipótese de a medida i conter erro grosseiro, pode-se concluir que:

- se $|\hat{r}_i^N| > C$, então, a medida i contém erro grosseiro;
- se $|\hat{r}_i^N| \leq C$, então, a medida i não contém erro grosseiro.

Do exposto acima, o método consiste em calcular os resíduos normalizados para todas as medidas e, então, encontrar a medida i que possui o maior resíduo normalizado. Caso $\hat{r}_i^N > C$, admite-se que a medida possui erro grosseiro. Como os resíduos normalizados \hat{r}_i^N possuem distribuição normal padrão, adotando por exemplo $C = 3$, significa que existe 99,7% de probabilidade de que $\hat{r}_i^N \leq 3$. Portanto, se $\hat{r}_i^N > 3$, a medida i é indicada como portadora de erro grosseiro.

Após o processo de detecção e identificação a medida i é descartada ou, se possível, corrigida (Monticelli and Garcia, 1983) e o estado é reestimado. Esse procedimento é repetido até que não seja detectado nenhum erro grosseiro no conjunto de medidas. A maior dificuldade associada ao método do máximo resíduo normalizado é a necessidade do cômputo, a cada ciclo de estimação, da matriz de covariância $R_{\hat{r}}$. Essa dificuldade está associada a obtenção da matriz ganho inversa G^{-1} . Todavia, há métodos de inversão esparsa, tais como o utilizado em (Broussolle, 1978), que permitem a obtenção dos elementos da diagonal da matriz $R_{\hat{r}}$ de maneira mais eficiente.

4.5 Estudo de Caso

Considere a rede de 14 barras representada na figura 4.1. Os parâmetros elétricos dessa rede podem ser encontrados em (PSTCA, 2013). O plano de medição é composto de medidores convencionais de fluxos de potências em ramos, injeções de potências em barras e magnitude de tensão em barras. As medições das potências são feitas aos pares. As medidas disponíveis tornam a rede observável. A tabela 4.1 apresenta os valores medidos e os valores reais das grandezas medidas. A tabela 4.2 apresenta as variâncias das medidas.

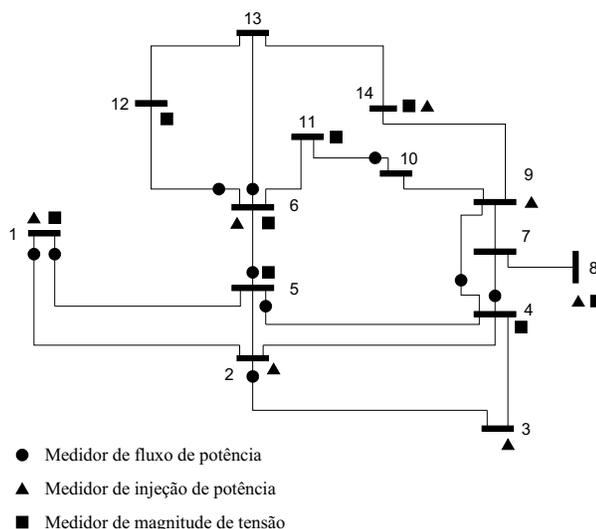


Figura 4.1: Rede de 14 barras do IEEE observável.

Para obter o estado da rede foi utilizado o método de Gauss-Newton. O ângulo da tensão da barra 1 foi adotado como referência angular, logo, o problema possui 27 variáveis de estado. O estimador foi inicializado com o perfil plano de tensões. Alguns índices que podem ser usados na definição do critério de convergência são mostrados na tabela 4.3 para as cinco primeiras iterações do processo iterativo. Com base nesses índices é possível afirmar que a partir da quarta iteração o processo iterativo pode ser considerado convergido. Note ainda que, por exemplo, a norma do gradiente da função objetivo, $\|\nabla J\|_2$, tende rapidamente a zero indicando boas características de convergência do estimador. De fato, para o ponto de inicialização, $cond(H_w) = 59,3852$ e $cond(G) = 3,5266 \times 10^3$, que para este tipo de problema indicam matrizes bem condicionadas. Por fim, estabelecendo uma tolerância de 10^{-4} para o máximo valor absoluto do vetor de correções, Δx , a convergência teria sido alcançada após a quarta iteração.

As figuras 4.2 e 4.3 ilustram uma comparação entre as magnitudes de tensão reais e estimadas, e entre ângulos de fase reais e estimados, respectivamente. O máximo resíduo

Tabela 4.1: Valores reais e medidos para a rede de 14 barras observável.

Tipo	Real	Medido	Tipo	Real	Medido	Tipo	Real	Medido
P_1	2,4692	2,4977	P_{1-2}	1,6711	1,6623	P_{6-13}	0,1869	0,1784
Q_1	-0,1862	-0,2468	Q_{1-2}	-0,2276	-0,2667	Q_{6-13}	0,0777	0,0088
P_2	0,1721	0,1632	P_{1-5}	0,7981	0,8342	P_{10-11}	-0,0404	-0,0269
Q_2	0,3517	0,3354	Q_{1-5}	0,0414	-0,0073	Q_{10-11}	-0,0201	-0,0586
P_3	-0,9891	-0,9638	P_{2-3}	0,7707	0,8034	P_{5-4}	0,6480	0,6597
Q_3	0,0885	0,1271	Q_{2-3}	0,0320	-0,0095	Q_{5-4}	-0,1488	-0,1524
P_6	-0,1176	-0,0805	P_{4-7}	0,2941	0,2904	V_1	1,0600	1,0173
Q_6	0,0733	0,0675	Q_{4-7}	-0,0959	-0,1011	V_4	1,0162	1,0277
P_8	0,0000	0,0063	P_{4-9}	0,1684	0,2116	V_5	1,0181	1,0038
Q_8	0,1854	0,1830	Q_{4-9}	-0,0021	-0,0559	V_6	1,0700	1,0948
P_9	-0,3098	-0,3044	P_{5-6}	0,4644	0,5238	V_8	1,0900	1,1291
Q_9	-0,1743	-0,2097	Q_{5-6}	0,1205	0,1954	V_{11}	1,0556	1,0897
P_{14}	-0,1565	-0,2008	P_{6-12}	0,0820	0,0798	V_{12}	1,0543	1,0532
Q_{14}	-0,0525	-0,0721	Q_{6-12}	0,0268	0,0514	V_{14}	1,0327	1,0168

Nota: Valores dados em pu

Tabela 4.2: Variâncias adotadas para as medidas.

Tipo de Medida	Variância
Magnitude de tensão	1×10^{-3}
Fluxo de potência (ativa e reativa)	$1/9 \times 10^{-2}$
Injeção de potência (ativa e reativa)	$1/8 \times 10^{-2}$

Tabela 4.3: Índices do estimador convencional durante o processo iterativo.

Iteração	$\ \nabla J\ _2$	$\max \Delta x $	$ \Delta J/J $
1	$2,8478 \times 10^4$	$3,4575 \times 10^{-1}$	$9,6764 \times 10^{-1}$
2	$9,2740 \times 10^2$	$2,9413 \times 10^{-2}$	$9,5423 \times 10^{-1}$
3	$9,8163 \times 10^{-1}$	$7,2575 \times 10^{-4}$	$1,7184 \times 10^{-2}$
4	$6,8362 \times 10^{-4}$	$3,7640 \times 10^{-6}$	$2,4022 \times 10^{-8}$
5	$8,9881 \times 10^{-6}$	$2,3663 \times 10^{-8}$	$4,1092 \times 10^{-13}$

normalizado corresponde a medida de fluxo de potência reativa Q_{5-6} , $\hat{r}^N = 2,8428$. Note que $\hat{r}^N < 3$, portanto, pode-se dizer que o estado estimado é confiável. A função objetivo $J(x)$ é uma variável aleatória com distribuição de probabilidade χ^2 com 15 graus de liberdade. Para o estado estimado, $J(\hat{x}) = 15,8001$. Esse valor também é estatisticamente aceitável,

pois $P(J(x) \leq J(\hat{x})) = 60,45 \%$, de fato, $E[J(x)] = 15$.

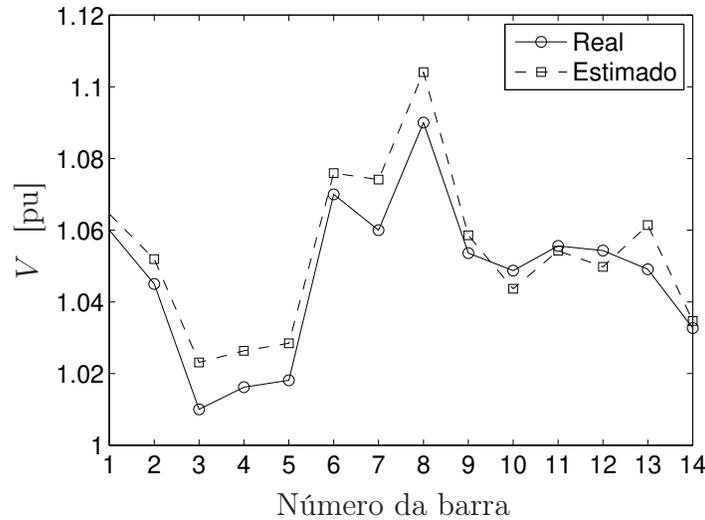


Figura 4.2: Magnitude das tensões - 14 barras IEEE observável

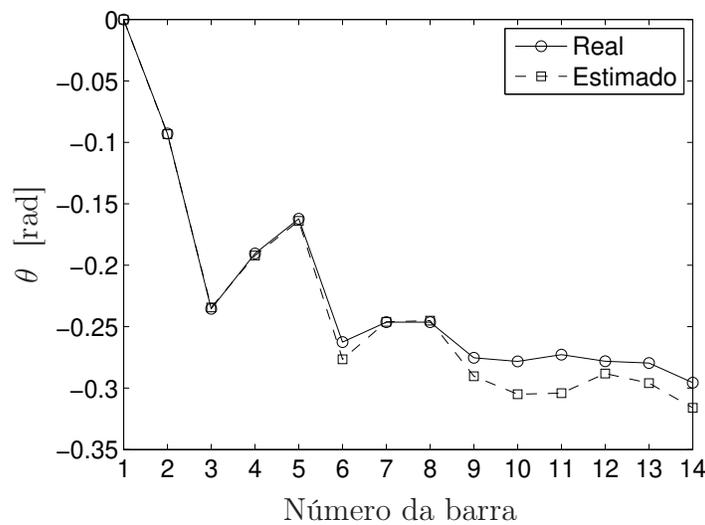


Figura 4.3: Ângulos das tensões - 14 barras IEEE observável

Para ilustrar a presença de erros grosseiros no resultado da estimação, considere que a medida de fluxo de potência reativa Q_{5-6} assume o valor 0,2205 pu (o valor correto de Q_{5-6} é 0,1205) e todas as demais medidas continuam com os valores apresentados na tabela 4.1. Inicializando o estimador de estado com o perfil plano de tensões e adotando uma tolerância de convergência igual a 10^{-4} para o maior valor absoluto do vetor de correções, Δx , o estado estimado é obtido após quatro iterações. Nessas condições o maior resíduo

normalizado obtido foi $r_{Q_{5-6}}^N = 3,2$. Assim, pode-se afirmar que a medida Q_{5-6} contém erro grosseiro e, portanto, o estado estado estimado não é plenamente confiável. Por outro lado, o valor da função objetivo, $J(\hat{x}) = 17,9521$, indica que o estado estimado é razoável pois, $P(J(x) \leq J(\hat{x})) = 73,48 \%$. Isso indica que o teste do índice $J(x)$ nem sempre é adequado para a detecção de erros grosseiros (Monticelli and Garcia, 1983). Após a retirada da medida Q_{5-6} o estado foi reestimado e o maior resíduo normalizado calculado vale 1,6031, além disso no ponto de convergência $J(\hat{x}) = 7,7426$, portanto, o processo de estimação é encerrado.

4.6 Conclusões

Neste capítulo foram revisadas as principais características e funções do processo de estimação de estado em sistemas de energia elétrica. O principal objetivo desta revisão é estabelecer os conceitos necessários para o desenvolvimento, no capítulo seguinte, do estimador regularizado. Como será visto, os resultados da função de análise de observabilidade tem um papel muito importante para justificar a aplicação do estimador regularizado, além disso, esses resultados também servem para analisar o comportamento desse estimador. Por fim, o estudo de caso apresentado, visa expor o comportamento numérico das variáveis envolvidas no processo de estimação convencional, de tal maneira que os mesmos possam servir de referência para os resultados do estimador regularizado.

Capítulo 5

A Estimação de Estado Regularizada

5.1 Introdução

Embora os sistemas de medição possam ser projetados para garantir que durante a operação as redes sejam observáveis, eventualmente, falhas de comunicação, mudanças topológicas imprevistas ou falhas em medidores podem tornar os sistemas de energia temporariamente não observáveis. Assim, na abordagem tradicional, antes de estimar o estado da rede é necessário verificar a observabilidade da rede. Caso a rede seja não observável, a função de análise de observabilidade deve identificar as porções observáveis da rede, as chamadas ilhas observáveis. Após a identificação é possível estimar o estado de cada ilha observável alocando referências angulares em cada uma delas. Nesse caso, não haverá qualquer informação sobre as porções não observáveis da rede. Alternativamente, pode-se executar a etapa de restauração da observabilidade, onde pseudomedidas críticas são alocadas na rede tornando-a completamente observável e garantindo que o estado das ilhas observáveis não será alterado pelas pseudomedidas (Monticelli and Wu, 1985b). As pseudomedidas podem, por exemplo, ser geradas a partir da análise do comportamento histórico da rede usando mecanismos de previsão de carga como é mostrado em (Brown Do Coutto Filho and Stacchini de Souza, 2009). Alguns relatos recentes sobre a utilização de pseudomedidas em sistemas reais podem ser vistos em (Atanackovic and Dwernychuk, 2011; Boddeti et al., 2011).

Do ponto de vista matemático, a solução das equações normais 4.2 é única somente quando a matriz Jacobiana possui posto completo, condição que não é atendida quando a rede não é observável. Assim, o problema de estimação de estado de redes não observáveis pode ser incluído na classe dos problemas mal postos. Nesse contexto, de Almeida et al. (2012) propuseram a aplicação do esquema de regularização de Tikhonov ao estimador de mínimos quadrados ponderados. Essa abordagem elimina a infactibilidade causada pela falta de medidas. Como consequência as etapas de análise e restauração da observabilidade

não precisam ser realizadas para que o estado da rede possa ser estimado. O estimador regularizado é mais recomendado para os casos onde a rede torna-se não observável devido a perda temporária de medições. A seguir o estimador regularizado é apresentado.

5.2 O Estimador de Estado Regularizado

Seja um sistema de energia elétrica contendo z medidas, esse sistema pode ou não ser observável. Aplicando a regularização de Tikhonov, desenvolvida na seção 3.3, ao problema de mínimos quadrados não linear 4.1, é obtido o seguinte problema regularizado:

$$\min_{x \in \mathbb{R}^n} F(x) = J(x) + \lambda^2 \Omega(x) = [z - h(x)]^T R_z^{-1} [z - h(x)] + \lambda^2 [u - x]^T L^T L [u - x] \quad (5.1)$$

onde λ é o parâmetro de regularização, u representa um conhecimento *a priori* sobre as variáveis de estado e L é uma matriz a ser definida.

Na seção 3.3 foi mencionado que L é tipicamente a matriz identidade. Idealmente a matriz L deve ser tal que o funcional de regularização, $\Omega(x)$, exerça mínima influência sobre a solução obtida utilizando somente o termo $J(x)$. Para simplificar a obtenção de L , dado um sistema de n variáveis de estado a matriz $L \in \mathbb{R}^{n \times n}$ é definida por:

$$L(i, j) = \begin{cases} 0 & i \neq j \\ \begin{cases} 0, & \text{se a variável } i \text{ é medida} \\ 1, & \text{se a variável } i \text{ não é medida} \end{cases} & i = j \end{cases}$$

Portanto, as informações *a priori* são consideradas apenas para as variáveis de estado que não são diretamente medidas e L tem a forma geral,

$$L = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix}$$

Por simplicidade, assumindo que $\lambda^2 L^T L = S$, o problema regularizado 5.1 pode ser reescrito da seguinte maneira:

$$\min_{x \in \mathbb{R}^n} F(x) = J(x) + \lambda^2 \Omega(x) = [z - h(x)]^T R_z^{-1} [z - h(x)] + [u - x]^T S [u - x] \quad (5.2)$$

note que S tem a mesma estrutura de L , entretanto, os elementos não nulos da diagonal são substituídos por λ^2 , isto é:

$$S = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 \\ 0 & \lambda^2 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \lambda^2 \end{bmatrix}$$

Rearranjando os termos do problema acima e particionando o conjunto de medidas z , tem-se:

$$z = \begin{bmatrix} \bar{z} \\ \bar{u} \end{bmatrix}$$

onde \bar{u} representa um subconjunto contendo somente medidas de variáveis de estado e \bar{z} representa o subconjunto com todas as medidas restantes. Substituindo os elementos de u , correspondentes às barras onde existem medidas de variáveis de estado, pelos respectivos elementos de \bar{u} e substituindo os elementos nulos da matriz S pelo inverso das variâncias das respectivas medidas contidas em \bar{u} , as matrizes u e S ficam:

$$u = \begin{bmatrix} \bar{u}_1 \\ u_2 \\ \bar{u}_3 \\ \vdots \\ u_n \end{bmatrix}, \quad S = \begin{bmatrix} \frac{1}{\sigma_1^2} & 0 & 0 & \dots & 0 \\ 0 & \lambda^2 & 0 & \dots & 0 \\ 0 & 0 & \frac{1}{\sigma_3^2} & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \lambda^2 \end{bmatrix}$$

Note que o novo vetor u inclui tanto medidas como informações *a priori* sobre o estado. A partir deste ponto as informações *a priori* também serão denominadas pseudomedidas de estado. A matriz S , por sua vez, passa a incluir tanto as ponderações das medidas como as ponderações das pseudomedidas. Portanto, reescrevendo o problema 5.2 temos:

$$\min_{x \in \mathbb{R}^n} F(x) = [\bar{z} - \bar{h}(x)]^T R_{\bar{z}}^{-1} [\bar{z} - \bar{h}(x)] + [u - x]^T S [u - x] \quad (5.3)$$

Note que os termos $J(x)$ e $\Omega(x)$ não aparecem explicitamente na equação acima. Aplicando o método de Gauss-Newton a solução desse problema de otimização pode ser obtida através do seguinte processo iterativo:

$$[H(x^\nu)^T R_{\bar{z}}^{-1} H(x^\nu) + S] \Delta x^\nu = H(x^\nu)^T R_{\bar{z}}^{-1} [\bar{z} - h(x^\nu)] + S[u - x^\nu] \quad (5.4)$$

$$x^{\nu+1} = x^\nu + \Delta x^\nu$$

O termo $H(x)^T R_{\bar{z}}^{-1} H(x) + S$ passa a ser chamado de matriz ganho regularizada $G_r(x)$. O espaço nulo de S é trivial, logo, a condição $\mathcal{N}(H^T R_{\bar{z}}^{-1} H) \cap \mathcal{N}(S) = \{0\}$ é atendida e, portanto, a cada iteração a solução do sistema de equações acima é única. Colocando o sistema de equações 5.4 em uma forma compacta, tem-se:

$$\Delta \tilde{z}(x^\nu) = \begin{bmatrix} \bar{z} - \bar{h}(x^\nu) \\ u - x^\nu \end{bmatrix}, \quad \tilde{W} = \begin{bmatrix} R_{\bar{z}}^{-1} & 0 \\ 0 & S \end{bmatrix}, \quad \tilde{H}(x^\nu) = \begin{bmatrix} H(x^\nu) \\ I \end{bmatrix}$$

e as equações normais 5.4 podem ser reescritas como:

$$\tilde{H}(x^\nu)^T \tilde{W} \tilde{H}(x^\nu) \Delta x^\nu = \tilde{H}(x^\nu)^T \tilde{W} \Delta \tilde{z} \quad (5.5)$$

$$x^{\nu+1} = x^\nu + \Delta x^\nu$$

A despeito das características puramente matemáticas associadas ao modelo regularizado, é possível afirmar que sempre é possível obter o estado do sistema de energia regularizado, tendo em vista que existem medidas ou pseudomedidas de ângulo de fase e magnitude tensão em todas as barras do sistema.

A incorporação de informações *a priori* nos estimadores de estado de mínimos quadrados para sistemas de energia elétrica, embora em contextos diferentes, já vem sendo desenvolvida em trabalhos anteriores. Em (Costa and Albuquerque, 2011) foi desenvolvido um estimador de dois estágios com fatoração ortogonal onde, dada uma rede com medição convencional e fasorial, no primeiro estágio o estado é estimado considerando somente as medidas convencionais e no segundo estágio consideram-se as medidas fasoriais e o estado estimado no primeiro estágio como informação *a priori*.

5.2.1 Alguns aspectos relevantes

No processo de restauração da observabilidade requerido pelas metodologias clássicas de estimação de estado, as pseudomedidas alocadas devem permitir a estimação das variáveis de estado das partes não observáveis da rede sem que as variáveis de estado das ilhas observáveis

sejam afetadas por essas pseudomedidas, portanto, as metodologias clássicas sugerem que as pseudomedidas alocadas devem ser críticas (Monticelli and Wu, 1985b).

No modelo proposto em (de Almeida et al., 2012), apresentado acima, como a matriz de ganho regularizada é sempre não singular, a restauração da observabilidade não é requerida. Entretanto, ainda é necessário garantir que o estado estimado para as ilhas observáveis não seja deteriorado pelas pseudomedidas incluídas no vetor u . Para tanto, as ponderações das pseudomedidas devem ser tanto menores quanto menor a confiança das pseudomedidas. Contudo, o uso de ponderações muito pequenas pode piorar o condicionamento numérico da matriz de ganho regularizada, fato que pode dificultar a convergência do estimador. Portanto, a escolha adequada dos valores das pseudomedidas contidas no vetor u e de suas variâncias, irá assegurar a qualidade do estado estimado juntamente com boas características de convergência do algoritmo. Note que a escolha das ponderações das pseudomedidas corresponde à escolha do parâmetro de regularização discutida no capítulo 3. Esse aspecto será tratado mais detalhadamente ainda neste capítulo.

5.2.2 Aspectos computacionais e de implementação

A estimação de estado é uma função executada em tempo real nos centros de controle. Isso sugere que o estimador de estado deve ser uma ferramenta confiável e que apresente seus resultados em um tempo hábil permitindo o monitoramento e definição de ações de controle sobre a rede. Com base nisso, os programas computacionais desenvolvidos buscam reduzir o tempo de processamento, valendo-se das características algébricas das matrizes e vetores que compõem o problema de estimação de estado.

No caso do estimador regularizado, a matriz ganho regularizada, $H(x)^T R_{\bar{z}}^{-1} H(x) + S$, apresenta as mesmas características da matriz ganho usada no estimador convencional, ou seja, ela é esparsa, simétrica e definida positiva. Equivalentemente, a matriz Jacobiana \tilde{H} também preserva as características algébricas da matriz Jacobiana convencional. Isso permite utilizar as técnicas de fatoração e armazenamento compacto comumente adotadas nos estimadores convencionais. Portanto, para a implementação do estimador regularizado a partir de um programa baseado na formulação clássica são necessárias apenas pequenas alterações. Além disso, desde que a ponderação para as pseudomedidas tenha sido especificada, o tempo computacional associado a uma iteração do estimador regularizado será o mesmo de uma iteração de um estimador convencional.

5.2.3 Processamento de erros grosseiros na estimação de estado regularizada

Ao introduzirmos o conceito de regularização de Tikhonov na seção 3.3, e posteriormente ao aplicarmos esse método na estimação de estado em sistemas de energia elétrica, foi dado um enfoque puramente determinístico, isto é, não foram feitas considerações estatísticas sobre as pseudomedidas introduzidas no problema. Quando analisamos o problema regularizado, na forma como está escrito no sistema de equações 5.3, fica claro que estamos lidando com um problema de mínimos quadrados ordinário. Isso fica mais evidente quando o escrevemos da seguinte maneira:

$$\min_{x \in \mathbb{R}^n} \left\| \begin{array}{c} R_{\bar{z}}^{-\frac{1}{2}}[\bar{z} - \bar{h}(x)] \\ S^{\frac{1}{2}}[u - x] \end{array} \right\|_2^2$$

Assim, a hipótese de que as medidas são variáveis aleatórias independentes com distribuição de probabilidade normal, média nula e variância conhecida, continua verdadeira. Essa mesma hipótese é aplicada às pseudomedidas sem perda de generalidade. De fato, a matriz S^{-1} , que representa a matriz de covariância das medidas e pseudomedidas de estado contidas em u , é diagonal. Além disso, admite-se que as variáveis contidas em \bar{z} e u também são independentes, isto é, $\text{Cov}[\bar{z}, u] = 0$. Portanto, a regra do máximo resíduo normalizado pode ser aplicada ao estimador de estado regularizado.

Assim, o vetor de resíduos estimados do estimador regularizado pode ser escrito como:

$$\tilde{r} = \begin{bmatrix} \bar{z} - \bar{h}(\hat{x}) \\ u - \hat{x} \end{bmatrix} \quad (5.6)$$

e a matriz de covariâncias dos resíduos estimados pode ser obtida por inspeção a partir da equação 4.6, resultando na equação 5.7. Finalmente, a partir das equações 5.6 e 5.7 os resíduos normalizados são calculados do mesmo modo que no estimador convencional.

$$R_{\tilde{r}} = \begin{bmatrix} R_{\bar{z}} & 0 \\ 0 & S^{-1} \end{bmatrix} - \begin{bmatrix} H(\hat{x})G_r^{-1}(\hat{x})H(\hat{x})^T & H(\hat{x})G_r^{-1}(\hat{x}) \\ G_r^{-1}(\hat{x})H(\hat{x})^T & G_r^{-1}(\hat{x}) \end{bmatrix} \quad (5.7)$$

5.2.4 Estudo de caso

Considere que a rede estudada na seção 4.5 tornou-se não observável devido a perda de algumas medidas. A figura 5.1 representa essa rede e as medidas remanescentes. Note que após a perda das medidas formaram-se duas ilhas observáveis e cinco barras ficaram isoladas.

Nos testes que seguirão, considera-se que os valores medidos são aqueles mostrados na tabela 4.1. As variâncias dessas medidas são mostradas na tabela 4.2. As ilhas observáveis 1 e 2 são compostas pelas barras $\{1,2,3,4,5,6,13\}$ e $\{10,11\}$, respectivamente.

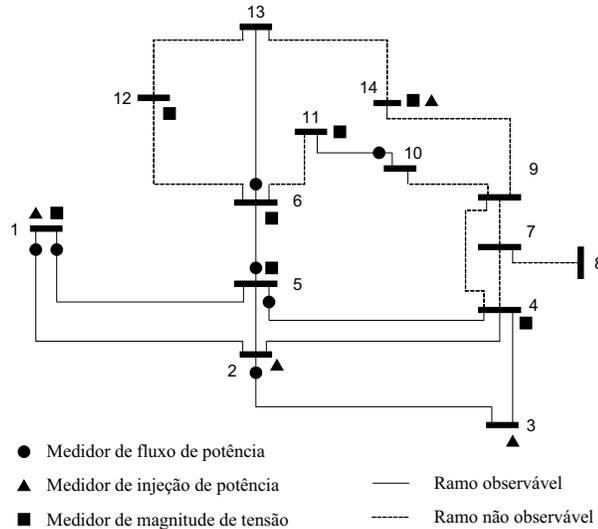


Figura 5.1: Rede de 14 barras do IEEE não observável 1.

Estimação do estado das ilhas observáveis

Em uma situação real, o estimador de estado poderia, dependendo da sua concepção, estimar o estado das ilhas observáveis independentemente ou incluir pseudomedidas no conjunto de medidas de modo a restaurar a observabilidade e, então, estimar o estado para toda a rede. Por ora, iremos estimar o estado das ilhas observáveis 1 e 2 de maneira independente. Para obter o estado dessas ilhas observáveis elas foram transformadas em ilhas físicas e adotou-se as barras 1 e 11 como barras de referência angular. A tabela 5.1 apresenta os resultados do processo de estimação por mínimos quadrados ponderados em comparação com o estado real da rede. A tabela 5.2 apresenta os resíduos estimados para as medidas de ambas as ilhas.

O estimador de estado foi inicializado com o perfil plano de tensões para ambos os casos, além disso, adotou-se a tolerância de convergência igual a 10^{-4} para o máximo valor no vetor de correções Δx . Após quatro iterações ambos os estimadores convergiram. O valor da função objetivo para a ilha 1, $J_1(\hat{x}) = 6,1259$, indica que o estado estimado é aceitável, pois $P(J_1(x) \leq J_1(\hat{x})) = 27,27\%$. Por outro lado, para a ilha 2 temos que $J_2(\hat{x}) = 0$. De fato, as medidas disponíveis na ilha 2 são críticas e, portanto, o estimador não pode filtrar os ruídos dessas medidas. Observe na tabela 5.2 que os resíduos estimados para as medidas da ilha 2 são todos nulos.

Tabela 5.1: Estados da rede de 14 barras não observável - Ilhas independentes.

Barra	Estado real		Estado estimado	
	V	θ	V	θ
1	1,0600	0	1,0543	0,0000
2	1,0450	-5,3206	1,0416	-5,4525
3	1,0100	-13,4885	1,0127	-13,7002
4	1,0162	-10,9142	1,0163	-11,2305
5	1,0181	-9,2928	1,0183	-9,5832
6	1,0700	-15,0394	1,0567	-16,1623
10	1,0487	-15,9453	1,0772	-0,0175
11	1,0556	-15,6294	1,0897	0,0000
13	1,0491	-16,0197	1,0447	-17,3383

Nota: Valores de tensão em pu e ângulos em graus

Tabela 5.2: Resíduos estimados - Ilhas independentes.

Tipo	$\hat{r} \times 10^2$	Tipo	$\hat{r} \times 10^2$	Tipo	$\hat{r} \times 10^2$	Tipo	$\hat{r} \times 10^2$
P_1	0,2234	Q_{1-2}	0,4870	P_{5-6}	-0,1117	V_6	3,8057
Q_1	1,1406	P_{1-5}	2,2982	Q_{5-6}	0,9689	P_{10-11}	0,0000
P_2	-1,4892	Q_{1-5}	-2,0664	P_{6-13}	0,0000	Q_{10-11}	0,0000
Q_2	1,0786	P_{2-3}	3,3178	Q_{6-13}	0,0000	V_{11}	0,0000
P_3	1,0925	Q_{2-3}	-1,0385	V_1	-3,6995	—	—
Q_3	-0,1824	P_{5-4}	0,0871	V_4	1,1373	—	—
P_{1-2}	-2,1948	Q_{5-4}	-0,1934	V_5	-1,4522	—	—

Nota: Valores dados em pu

Utilizando a SVD

Neste caso estimou-se o estado de toda a rede. Considere novamente a rede não observável da figura 5.1. Tendo em vista que a matriz Jacobiana do modelo de medição tem deficiência de posto, a cada iteração, as equações normais possuem infinitas soluções. Objetivando encontrar uma delas, iremos aplicar a SVD ao seguinte problema iterativo.

$$H_w(x^\nu)x^{\nu+1} = d_w^\nu \quad (5.8)$$

onde $d_w^\nu = R_z^{-\frac{1}{2}}[z - h(x^\nu) + H(x^\nu)x^\nu]$. A barra 1 foi definida como barra de referência angular e o problema foi inicializado com o perfil plano de tensões. O espectro dos valores singulares e os coeficientes $|u_i^T d_w|$ para a primeira iteração estão mostrados na figura 5.2.

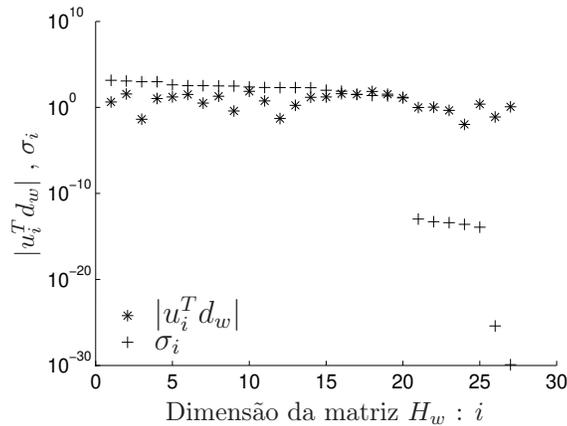


Figura 5.2: Gráfico de Picard para a rede de 14 barras não observável 1.

A figura 5.2 mostra que sete valores singulares são significativamente menores que os demais, portanto, $\text{posto}(H_w) = 20$. Nessa mesma figura também é possível verificar que a condição discreta de Picard não é satisfeita, isto é, na média, os coeficientes $|u_i^T d_w|$ decaem a zero mais lentamente que os valores singulares. Além disso, existem descontinuidades no espectro dos valores singulares. Esses fatos reforçam a ideia de que, qualquer tentativa convencional para obter o estado da rede diretamente a partir da equação 5.8, ou seja, a partir das equações normais, poderá falhar. Podemos, pois, utilizar a SVD compacta, desprezando os sete últimos valores singulares, que em última instância levará à inversa generalizada de Moore-Penrose. Alternativamente, pode-se utilizar a SVD truncada com um parâmetro de regularização $p < 20$.

A tabela 5.3 apresenta alguns resultados da estimação utilizando a SVD. Adotou-se a tolerância de convergência igual a 10^{-4} para o maior valor absoluto do vetor de correções Δx . Através dos valores dispostos na tabela 4.1 é possível calcular a norma do erro embutido nas medidas, isto é, $\|e\|_2 = 0,1932$, por outro lado, se essa informação não estivesse disponível poderíamos estimá-la da seguinte maneira, $E[\|e\|_2] = [\text{tr}(R_z)]^{\frac{1}{2}} = 0,1804$. A partir dos resultados expostos na tabela 5.3 é possível notar que a norma dos resíduos de estimação da SVD compacta é menor que a estimativa para a norma do erro das medidas, portanto, somente as soluções truncadas convergentes satisfizeram o princípio da discrepância (vide equação 3.22). Contudo, levando em conta a norma do gradiente, $\|\nabla J(\hat{x})\|_2$, e o valor da função objetivo $J(\hat{x})$ no ponto de convergência, a melhor solução foi obtida pela SVD compacta.

As tabelas 5.4 e 5.5 apresentam, respectivamente, os resíduos estimados e o estado estimado pela SVD compacta. Note que, para as ilhas observáveis o estado estimado é o mesmo encontrado pelo estimador convencional mostrado na tabela 5.1. Por fim, considerando que a rede em estudo não é observável, a matriz de covariância dos resíduos estimados, $R_{\hat{r}}$, não é

Tabela 5.3: Resultados do processo iterativo de estimação utilizando a SVD.

SVD	n° de iterações	$\ \nabla J(\hat{x})\ _2$	$J(\hat{x})$	$\ \hat{r}\ _2$
Compacta	10	$1,582 \times 10^{-5}$	6,1259	$8,0759 \times 10^{-2}$
Truncada $p = 19$	7	375,6961	177,3664	$4,2439 \times 10^{-1}$
Truncada $p = 18$	10	375,6666	177,3335	$4,2435 \times 10^{-1}$
Truncada $p < 18$	diverge!	—	—	—

definida, já que depende do inverso da matriz ganho. Portanto, não é possível fazer considerações estatísticas adicionais sobre o índice $J(\hat{x})$ e sobre os resíduos estimados. Para que tais análises fossem realizadas seria necessário identificar e descartar as barras não observáveis, o que pode ser feito, por exemplo, a partir da matriz V_0 , subproduto da SVD. Nesse caso, teríamos uma matriz $R_{\hat{r}}$ definida e as análises de erro poderiam ser realizadas normalmente.

Tabela 5.4: Resíduos estimados - SVD.

Tipo	$\hat{r} \times 10^2$	Tipo	$\hat{r} \times 10^2$	Tipo	$\hat{r} \times 10^2$	Tipo	$\hat{r} \times 10^2$
P_1	0,2234	Q_{1-2}	0,4870	P_{5-6}	-0,1117	V_6	3,8057
Q_1	1,1406	P_{1-5}	2,2982	Q_{5-6}	0,9689	P_{10-11}	0,0000
P_2	-1,4892	Q_{1-5}	-2,0664	P_{6-13}	0,0000	Q_{10-11}	0,0000
Q_2	1,0786	P_{2-3}	3,3178	Q_{6-13}	0,0000	V_{11}	0,0000
P_3	1,0925	Q_{2-3}	-1,0385	V_1	-3,6995	P_{14}	0,0000
Q_3	-0,1824	P_{5-4}	0,0871	V_4	1,1373	Q_{14}	0,0000
P_{1-2}	-2,1948	Q_{5-4}	-0,1934	V_5	-1,4522	V_{12}	0,0000
—	—	—	—	—	—	V_{14}	0,0000

Nota: Valores dados em pu

Cabe ressaltar que a SVD tem muitas propriedades interessantes para análise de problemas de mínimos quadrados. Todavia, o elevado custo computacional associado a sua obtenção limita bastante o seu uso para aplicações reais.

Aplicando a regularização de Tikhonov

Para aplicar o método de regularização de Tikhonov é necessário introduzir informações adicionais no problema de estimação original. O primeiro passo é definir quais informações *a priori* ou pseudomedidas serão utilizadas e, então, definir a ponderação adequada para elas, ou de forma equivalente, definir o parâmetro de regularização λ . Para definir a ponderação adequada para as pseudomedidas analisaremos o comportamento do estimador em função dessa variável.

Tabela 5.5: Estados da rede de 14 barras não observável 1 - SVD.

Barra	Estado real		Estado estimado	
	V	θ	V	θ
1	1,0600	0,0000	1,0543	0,0000
2	1,0450	-5,3206	1,0416	-5,4525
3	1,0100	-13,4885	1,0127	-13,7002
4	1,0162	-10,9142	1,0163	-11,2305
5	1,0181	-9,2928	1,0183	-9,5832
6	1,0700	-15,0394	1,0567	-16,1623
7	1,0600	-14,1153	0,0000	0,0000
8	1,0900	-14,1153	0,0000	0,0000
9	1,0536	-15,7754	1,0556	-1,5525
10	1,0487	-15,9453	1,0772	-0,0088
11	1,0556	-15,6294	1,0897	0,0088
12	1,0543	-15,9375	1,0532	0,0000
13	1,0491	-16,0197	1,0447	-17,3383
14	1,0327	-16,9365	1,0168	-9,7473

Nota: Valores de tensão em pu e ângulos em graus

O problema de estimação de estado regularizada consiste em obter a solução do seguinte problema de otimização:

$$\min_{x \in \mathbb{R}^n} F(x) = [\bar{z} - \bar{h}(x)]^T R_{\bar{z}}^{-1} [\bar{z} - \bar{h}(x)] + [u - x]^T S [u - x]$$

onde u é um vetor que contém medidas e pseudomedidas de variáveis de estado e \bar{z} contém as medidas. As matrizes S e $R_{\bar{z}}^{-1}$ são as respectivas matrizes de ponderação. Note que os elementos de S associados às pseudomedidas contidas em u são definidos por λ^2 . Neste estudo, será admitido que não existe nenhuma informação anterior sobre do estado da rede, portanto, utilizaremos o perfil plano de tensões como pseudomedidas. Os valores utilizados para as medidas são os mesmos dos estudos anteriores. A figura 5.3 ilustra a curva L para o problema em estudo. No eixo horizontal está representada a soma ponderada dos quadrados dos resíduos de todas as medidas e no eixo vertical está a soma dos quadrados dos resíduos das pseudomedidas. Note que nos eixos estão dispostos os logaritmos das variáveis. O ponto da curva L com a maior curvatura corresponde a um parâmetro de regularização $\lambda^2 = 0,5337$.

A figura 5.4 mostra o comportamento do índice $F(x)$ em função do parâmetro de regularização. Repare que existe uma ampla faixa de valores para λ em que a função $F(x_\lambda)$ praticamente não se altera. Assim, ao adotar qualquer valor de ponderação menor que 10^0

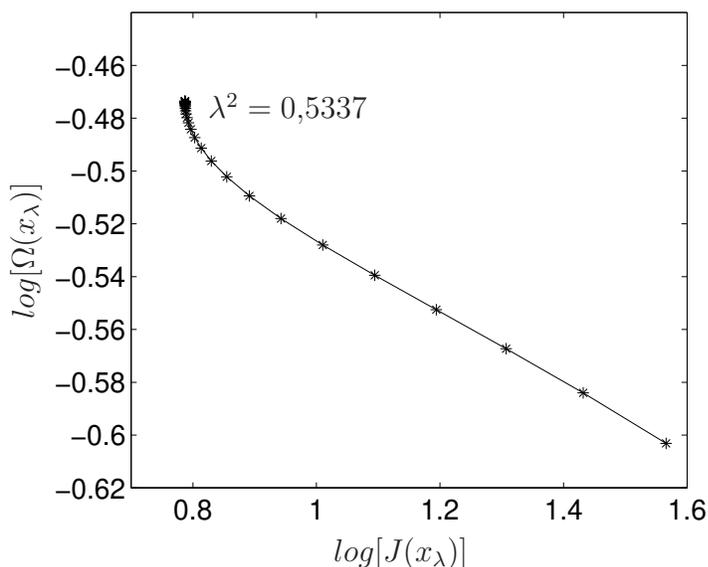


Figura 5.3: Curva L para a rede de 14 barras não observável 1.

tem-se praticamente o mesmo estado estimado.

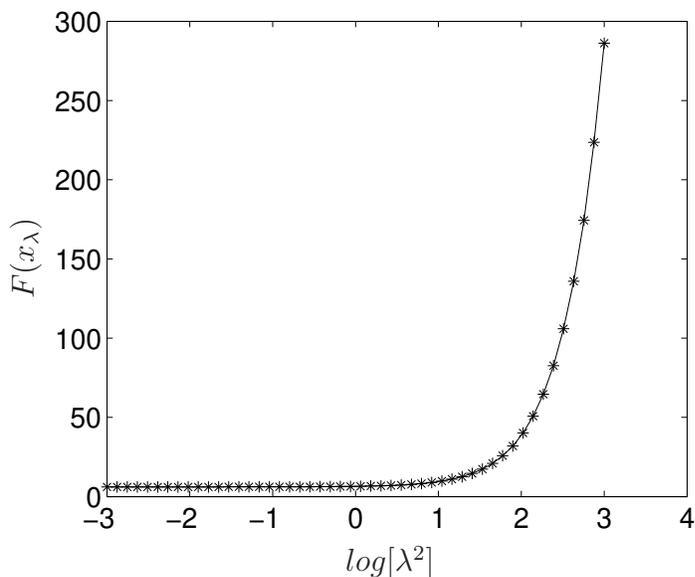


Figura 5.4: Variação do índice $F(x_\lambda)$ para a rede de 14 barras não observável 1.

O plano de medição mostrado na figura 5.1 é composto por 22 medidas de potência e 7 medidas de magnitude de tensão. Descontando a referência angular, o problema possui 27 variáveis de estado. Para esse caso, no estimador de estado regularizado foram incluídas 20 pseudomedidas de variáveis de estado. Portanto, pode-se concluir que a função objetivo $F(x)$ é uma variável aleatória que possui distribuição de probabilidade χ^2 com 22 graus de

liberdade. Considerando a metodologia apresentada na seção 3.3.2, pode-se criar um intervalo com 95 % de confiança e procurar por um λ que o satisfaça, isto é:

$$8,9986 \leq F(x_\lambda) \leq 35,0014$$

Para ilustrar os resultados que seriam obtidos por essa metodologia, por busca exaustiva foram encontradas as ponderações associadas aos limites desse intervalo de confiança. Para o limite inferior e superior foram obtidos $\lambda^2 = 8,5754$ e $\lambda^2 = 88,561$, respectivamente. A tabela 5.6 apresenta os estados estimados para esses dois parâmetros de regularização e para o parâmetro obtido através do método da curva L.

Tabela 5.6: Estados da rede de 14 barras não observável - Regularização de Tikhonov.

Barra	Estados estimados					
	$\lambda^2 = 0,5337$		$\lambda^2 = 8,5754$		$\lambda^2 = 88,561$	
	V	θ	V	θ	V	θ
1	1,0544	0,0000	1,0555	0,0000	1,0650	0,0000
2	1,0417	-5,4516	1,0429	-5,4388	1,0523	-5,3325
3	1,0128	-13,6980	1,0139	-13,6657	1,0235	-13,3975
4	1,0164	-11,2286	1,0175	-11,1998	1,0266	-10,9583
5	1,0184	-9,5815	1,0195	-9,5561	1,0286	-9,3430
6	1,0568	-16,1585	1,0579	-16,1031	1,0672	-15,6263
7	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000
8	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000
9	1,0674	3,4040	1,0651	3,4124	1,0481	3,4538
10	1,0772	-0,0088	1,0765	-0,0088	1,0705	-0,0088
11	1,0897	0,0088	1,0890	0,0088	1,0834	0,0088
12	1,0532	0,0000	1,0532	0,0000	1,0532	0,0000
13	1,0448	-17,3342	1,0459	-17,2729	1,0552	-16,7460
14	1,0167	-6,8717	1,0160	-6,8568	1,0113	-6,7213

Nota: Valores de tensão em pu e ângulos em graus

A tabela 5.7 apresenta alguns indicadores dos resultados de estimação para os três parâmetros de regularização. Apesar de haver uma diferença significativa entre os valores de $F(x_\lambda)$, note que o valor do índice $J(x_\lambda)$ é muito próximo para as três soluções. Perceba que para todas as soluções o princípio da discrepância não é atendido, isto é, as normas dos resíduos estimados são inferiores a norma dos ruídos das medidas, $\|e\|_2$. Uma maneira alternativa para avaliar a qualidade dos estados estimados é através da norma euclidiana do vetor de desvios entre o estado real da rede, x^{real} , e o estado estimado, x_λ (Costa and Albuquerque, 2011). Note que essa norma é muito parecida nos três casos, o que sugere que

os estados estão igualmente próximos ao estado real da rede.

Tabela 5.7: Comparação entre as três soluções de Tikhonov.

λ^2	$F(x_\lambda)$	$J(x_\lambda)$	$\ r_\lambda\ _2$	$\ x^{real} - x_\lambda\ _2$
0,5337	6,3052	6,1259	$8,0758 \times 10^{-2}$	0,7134
8,5754	8,9986	6,1353	$8,0802 \times 10^{-2}$	0,7133
88,561	35,0014	6,8976	$8,5461 \times 10^{-2}$	0,7136

Admitindo uma situação hipotética onde os ruídos das medidas disponíveis para o estimador regularizado são todos nulos, ou seja, as medidas são perfeitas, para assegurar que o termo de regularização não deteriore o estado das lhas observáveis, buscaremos um parâmetro de regularização que leve o índice $J(x_\lambda)$ tão próximo de zero quanto for possível. Note que essa restrição deve ser atendida independentemente da qualidade das pseudomedidas. A figura 5.5 ilustra o comportamento do índice $J(x_\lambda)$ em função de λ , dados os valores reais das medidas dispostos na tabela 4.1. A partir da figura 5.5 é possível notar que para valores de ponderação, $\lambda^2 \leq 10^{-2}$, assegura-se que as pseudomedidas estão interferindo o mínimo possível no estado das porções observáveis da rede. Não obstante, valores de ponderação muito pequenos podem gerar problemas numéricos.

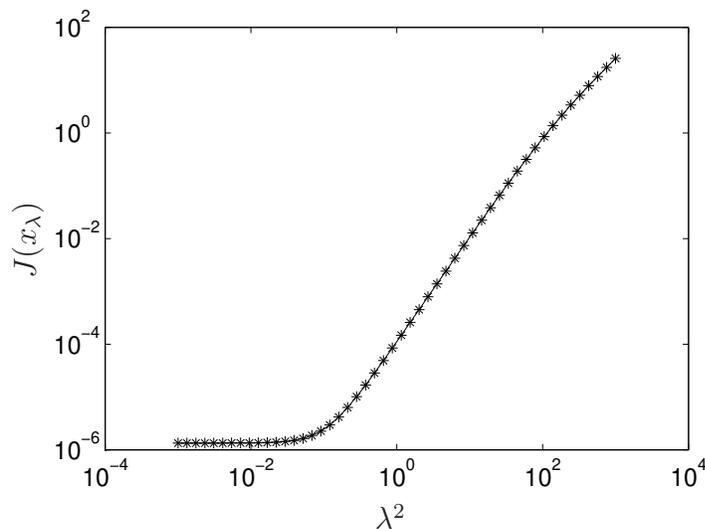


Figura 5.5: Variação do índice $J(x_\lambda)$ para a rede de 14 barras não observável 1 usando medidas perfeitas.

5.3 O Estimador Desacoplado Rápido Regularizado

O principal esforço computacional na resolução do problema da estimação de estado por mínimos quadrados ponderados está associado à construção e à fatoração, a cada iteração, da matriz ganho. Assim como no problema de fluxo de carga em sistemas de transmissão em alta tensão, para as medidas de potência, verifica-se um desacoplamento, com relação as variáveis de estado θ e V , entre as grandezas ativas e reativas. Isso permite a formulação de uma versão desacoplada para o problema de estimação de estado. Diferentemente das considerações normalmente feitas para justificar esses desacoplamentos, em (Monticelli and Garcia, 1990) uma abordagem matemática mais formal é dada para o tema. No estimador desacoplado as medidas são separadas em dois conjuntos, a saber:

- Medidas ativas: Fluxos e injeções de potência ativa e ângulo de fase em barras. Essas medidas serão designadas pelo índice A .
- Medidas reativas: Fluxos e injeções de potência reativa e magnitudes de tensão em barras. Essas medidas serão designadas pelo índice R .

As matrizes e vetores descritos no estimador regularizado podem ser rearranjados da seguinte maneira:

$$\bar{z} = \begin{bmatrix} \bar{z}_A \\ \bar{z}_R \end{bmatrix}, \quad u = \begin{bmatrix} u_A \\ u_R \end{bmatrix}$$

$$S = \begin{bmatrix} H_{AA} & H_{AR} \\ H_{AR} & H_{RR} \end{bmatrix}, \quad R_z = \begin{bmatrix} R_A & 0 \\ 0 & R_R \end{bmatrix}, \quad S = \begin{bmatrix} S_A & 0 \\ 0 & S_R \end{bmatrix}$$

Ignorando os blocos fora da diagonal da matriz Jacobiana, isto é, H_{AR} e H_{RA} , a matriz ganho regularizada desacoplada pode ser escrita como:

$$G_r = \begin{bmatrix} H_{AA}^T R_A^{-1} H_{AA} + S_A & 0 \\ 0 & H_{RR}^T R_R^{-1} H_{RR} + S_R \end{bmatrix} \quad (5.9)$$

Nos métodos de estimação de estado desacoplados no algoritmo, a matriz ganho desacoplada é calculada para o perfil plano de tensões e mantida constante durante todo o processo iterativo (Garcia et al., 1979). Essa alteração no método de Gauss-Newton não produz nenhuma mudança na solução final do problema. De fato, desde que o lado direito da equação 5.4 não é alterado, a única condição imposta para que direções de descida, Δx ,

sejam obtidas, é que a matriz ganho seja positiva definida. Como se sabe, quando a rede é observável H_{AA} e H_{RR} possuem posto completo, logo a matriz ganho desacoplada é positiva definida por construção. Por outro lado, quando a rede é não observável a presença das matrizes S_A e S_R asseguram o posto completo.

Os estimadores desacoplados no modelo, ou estimadores desacoplados rápidos, simplificam ainda mais o problema de estimação. Essas simplificações consistem em utilizar, em ambos os lados da equação 5.4, uma matriz H , calculada para o perfil plano de tensões, constante e desacoplada. Após a adoção de todas essas simplificações não é possível garantir que o estado estimado será o mesmo obtido por um estimador convencional (Garcia et al., 1979), entretanto, tem-se uma boa aproximação.

Cabe ressaltar que sob essas simplificações, a matriz H passa a ser função apenas dos parâmetros elétricos da rede. Nesse contexto, existem diferentes abordagens para lidar com esses parâmetros durante a montagem da matriz Jacobiana desacoplada. Todavia, duas delas são mais notáveis, as versões XB e BX , que consistem basicamente em ignorar as resistências série das linhas de transmissão na construção das matrizes H_{AA} e H_{RR} , respectivamente. Outras versões podem ser encontradas em (Garcia et al., 1979).

Quando comparados com a formulação completa do estimador de estado, os estimadores desacoplados rápidos possuem as seguintes vantagens: requerem menor quantidade de memória e são mais rápidos, já que trabalham com matrizes de dimensões menores e constantes. Entretanto, podem existir casos onde as condições de operação da rede e seus parâmetros elétricos não permitem a aplicação das hipóteses simplificadoras, levando a soluções não confiáveis. Outro aspecto importante é que a inclusão de medidas que não possuem desacoplamento entre as variáveis de estado, tais como medidas de corrente, não permitem a utilização desse tipo de estimador.

Utilizando as hipóteses simplificadoras, a solução do estimador desacoplado rápido regularizado na versão BX pode ser obtida através das seguintes equações recursivas:

$$[H_{AA}^T R_A^{-1} H_{AA} + S_A] \Delta \theta^\nu = H_{AA}^T R_A^{-1} \Delta z_A(\theta^\nu, V^\nu) + S_A [u_A - \theta^\nu] \quad (5.10)$$

$$\theta^{\nu+1} = \theta^\nu + \Delta \theta^\nu$$

$$[H_{RR}^T R_R^{-1} H_{RR} + S_R] \Delta V^\nu = H_{RR}^T R_R^{-1} \Delta z_R(\theta^{\nu+1}, V^\nu) + S_R [u_R - V^\nu] \quad (5.11)$$

$$V^{\nu+1} = V^\nu + \Delta V^\nu$$

onde, $\Delta z_A(\theta^\nu, V^\nu) = \bar{z}_A - \bar{h}_A(\theta^\nu, V^\nu)$ e $\Delta z_R(\theta^{\nu+1}, V^\nu) = \bar{z}_R - \bar{h}_R(\theta^{\nu+1}, V^\nu)$.

5.3.1 Exemplo

Considere novamente a rede não observável ilustrada na figura 5.1 e os valores de medidas e variâncias das tabelas 4.1 e 4.2. As pseudomedidas foram especificadas como o perfil plano de tensões e a ponderação adotada para elas foi $\lambda^2 = 10^{-3}$. Adotando uma tolerância de 10^{-4} para a convergência dos vetores de correção $\Delta\theta$ e ΔV , o estimador desacoplado rápido regularizado convergiu após 6 iterações ativas e 5 iterações reativas. O mesmo caso foi solucionado com o estimador regularizado completo. A tabela 5.8 apresenta ambos os estados estimados.

Tabela 5.8: Estados da rede de 14 barras não observável 1 - Método desacoplado.

Barra	Estado desacoplado		Estado completo	
	V	θ	V	θ
1	1,0550	0,0000	1,0543	0,0000
2	1,0423	-5,4459	1,0416	-5,4525
3	1,0133	-13,6824	1,0127	-13,7002
4	1,0169	-11,2222	1,0163	-11,2305
5	1,0189	-9,5771	1,0183	-9,5832
6	1,0573	-16,1349	1,0567	-16,1623
7	1,0000	0,0000	1,0000	0,0000
8	1,0000	0,0000	1,0000	0,0000
9	1,0681	3,7699	1,0676	3,4034
10	1,0772	-0,0083	1,0772	-0,0088
11	1,0897	0,0083	1,0897	0,0088
12	1,0532	0,0000	1,0532	0,0000
13	1,0453	-17,3092	1,0447	-17,3383
14	1,0168	-6,6511	1,0168	-6,8727

Nota: Valores de tensão em pu e ângulos em graus

A partir dos resultados das tabelas 5.8 e 5.9 é possível fazer uma comparação entre os resultados dos estimadores. Note que o estimador desacoplado não encontrou um mínimo da função objetivo, tendo em vista que o valor da norma do gradiente ainda é significativo. Como era de se esperar, o resultado do estimador completo é melhor. Contudo, em termos práticos, pode-se dizer que os resultados do estimador desacoplado são suficientemente precisos.

5.4 Conclusões

Neste capítulo foram desenvolvidas as versões completa e desacoplada do estimador de estado regularizado. Nas simulações realizadas foi possível observar o comportamento

Tabela 5.9: Resultados do processo iterativo de estimação utilizando o método desacoplado.

Estimador	$J(\hat{x})$	$F(\hat{x})$	$\ \nabla F(\hat{x})\ _2$	$\ x^{real} - \hat{x}\ _2$
Desacoplado	6,1328	6,1331	83,2533	0,7173
Completo	6,1259	6,1262	$5,5502 \times 10^{-4}$	0,7134

do estimador de estado regularizado em função de um conjunto particular de medidas e pseudomedidas. Conclui-se que, com a ponderação adequada para as pseudomedidas de variáveis de estado, o estado fornecido pelo estimador regularizado é muito próximo ao estado que seria obtido por um estimador de mínimos quadrados convencional.

Dentre as metodologias expostas para definição do parâmetro de regularização, as baseadas no critério da curva L ou na distribuição χ^2 do índice $F(x)$ apresentam resultados bastante promissores. Contudo, a implementação dessas metodologias não é atrativa pela complexidade associada e indicam a necessidade de se criar uma abordagem mais simples para a escolha do parâmetro λ .

Os resultados apresentados na figura 5.5 indicam ainda que o uso de valores menores para λ resultam em menor influência das pseudomedidas no estado das porções observáveis da rede. Em princípio, essa ideia pode ser generalizada independente das medidas e das pseudomedidas utilizadas. Testes realizados em redes distintas indicaram que para valores de λ^2 da ordem de 10^{-3} não houve prejuízo aos estados observáveis da rede. No entanto, como será visto no capítulo seguinte, em algumas situações particulares, problemas numéricos podem gerar dificuldades de convergência do estimador regularizado se pseudomedidas de baixa qualidade forem associadas a baixos valores do parâmetro de regularização.

Por fim, note que o uso do perfil plano de tensões como pseudomedidas não é o contexto adequado para a aplicação do estimador regularizado, pois nesse contexto, o estado estimado para as porções não observáveis da rede será, geralmente, pouco confiável. Portanto, como já foi estabelecido, o estimador regularizado é recomendado para casos onde a observabilidade é perdida temporariamente e um estado anterior de boa qualidade é conhecido.

Capítulo 6

Análise da Robustez dos Estimadores Regularizados

6.1 Introdução

No capítulo 5, o estimador regularizado originalmente proposto em (de Almeida et al., 2012) foi apresentado e discutido. A regularização utilizada consiste em introduzir pseudomedidas de variáveis de estado no problema original, a fim de assegurar a factibilidade do processo de estimação. Essa regularização foi estendida e uma versão desacoplada rápida do estimador foi desenvolvida. Foram apresentados estudos com os estimadores regularizados e foi dado destaque à escolha da ponderação para as pseudomedidas. Concluiu-se que uma maneira bastante simples e pragmática de ponderar as pseudomedidas é atribuindo pesos relativamente pequenos para elas.

Neste capítulo será mostrado que a adoção de pesos muito pequenos para pseudomedidas de baixa qualidade no estimador regularizado pode criar problemas numéricos em casos particulares. Esses problemas numéricos estão associados a elevados números de condição das matrizes envolvidas. O condicionamento numérico ruim acaba dificultando e até impedindo a convergência da versão completa do estimador de estado regularizado solucionado pelo método de Gauss-Newton.

Adicionalmente, neste capítulo serão investigadas as origens dos problemas numéricos citados e possíveis mecanismos de solução. Serão propostos dois algoritmos que utilizam as versões completa e desacoplada do estimador de estado regularizado. Ao final do capítulo são apresentados resultados de testes com as redes de 14 e 118 barras do IEEE. Nestes testes são discutidas as características dos algoritmos propostos.

6.2 Análise de Casos Críticos

Considere a rede elétrica não observável ilustrada na figura 6.1. Exceto pelas medidas de injeção de potência ativa e reativa na barra 6, esta rede é a mesma estudada na seção 5.2.4. A partir dos valores mostrados na tabela 4.1 é possível verificar que $P_6 = -0,0805$ e $Q_6 = 0,0675$. Note que as medidas de injeção são irrelevantes para a estimação do estado das ilhas observáveis, pois elas são adjacentes a ramos não observáveis. O perfil plano de tensões foi utilizado como pseudomedidas de estado e como ponto de inicialização do algoritmo. A ponderação de 10^{-3} foi adotada para todas as pseudomedidas.

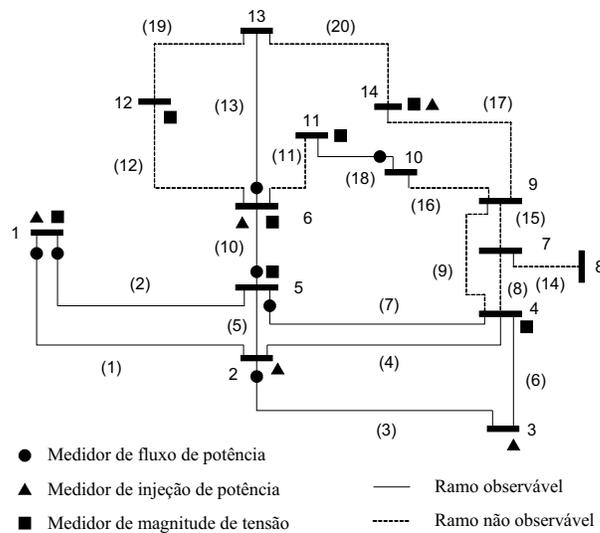


Figura 6.1: Rede de 14 barras do IEEE não observável 2.

Como pode ser visto na figura 6.2, após 10 iterações o processo iterativo é encerrado. Tendo em vista que na décima iteração a função objetivo e a norma do seu gradiente apresentam valores elevados, o estado estimado não é confiável. De fato, se o número de iterações não fosse limitado, o estimador iria seguir oscilando em torno de um estado não convergido.

Para o ponto de inicialização do algoritmo o $cond(G_r) = 1,9960 \times 10^9$. Esse número de condição da matriz ganho regularizada é bastante elevado e, em princípio, por si só poderia explicar a dificuldade de convergência do algoritmo. Todavia, como foi visto na seção 5.2.4, na ausência das medidas de injeção na barra 6, mesmo com valores pequenos de ponderação para as pseudomedidas, fato que também piora o condicionamento das matrizes envolvidas, não foram observados problemas de convergência. Portanto, o condicionamento numérico ruim das matrizes envolvidas e a presença de medidas de injeção nas fronteiras das ilhas observáveis, as chamadas injeções irrelevantes, podem dificultar a convergência do estimador regularizado. Cabe ressaltar que nem todas as injeções irrelevantes criam dificuldades de

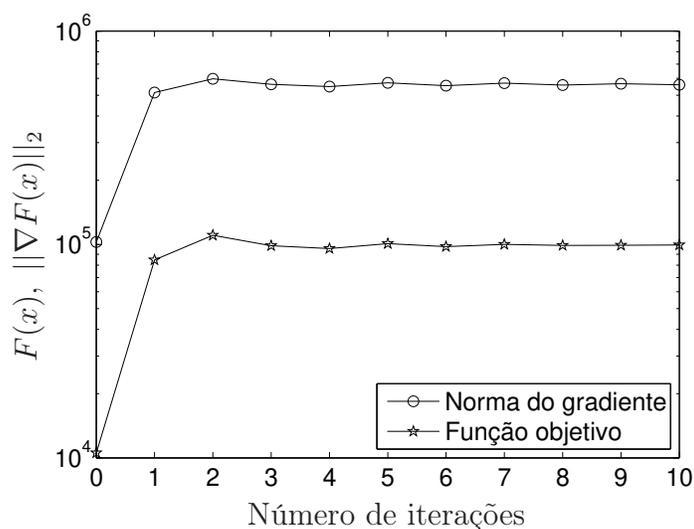


Figura 6.2: Comportamento de um caso divergente.

convergência para o estimador regularizado.

6.2.1 O efeito da ponderação das pseudomedidas

Conforme discutido na seção 5.2.4, dependendo da qualidade das pseudomedidas alocadas, se as suas ponderações forem adequadamente definidas os problemas de convergência podem ser contornados. Contudo, definir os valores adequados para as ponderações pode ser uma tarefa muito difícil. Para ilustrar esse fato considere dois casos distintos, no primeiro caso as pseudomedidas de estado foram especificadas como o perfil plano de tensões, no segundo caso adotou-se o estado real da rede como pseudomedidas. Para ambos os casos o estado da rede foi estimado utilizando diferentes ponderações para as pseudomedidas. As figuras 6.3 e 6.4 ilustram o comportamento do índice J e do número de iterações em função das ponderações utilizadas.

Na figura 6.3 é possível observar que para fatores de ponderação variando de 10^{-2} até 10^1 a influência das pseudomedidas nos estados observáveis é pouco significativa. Na figura 6.4 observa-se que para os fatores de ponderação em torno de 10^1 o número de iterações requeridas para a convergência é demasiadamente elevado.

Retomando o caso inicial, onde as pseudomedidas assumem o perfil plano de tensões com ponderação iguais a 10^{-3} , na tabela 6.1 são mostrados os valores das correções nas variáveis de estado das barras 5, 11, 12 e 13 em todas as iterações. Essas barras são vizinhas à barra 6 e, portanto, são diretamente influenciadas pelas injeções irrelevantes. A partir dos valores dispostos na tabela 6.1 é possível notar um comportamento oscilatório nas correções dos ângulos de fase das barras 11 e 12. Essas oscilações surgem devido à dificuldade da

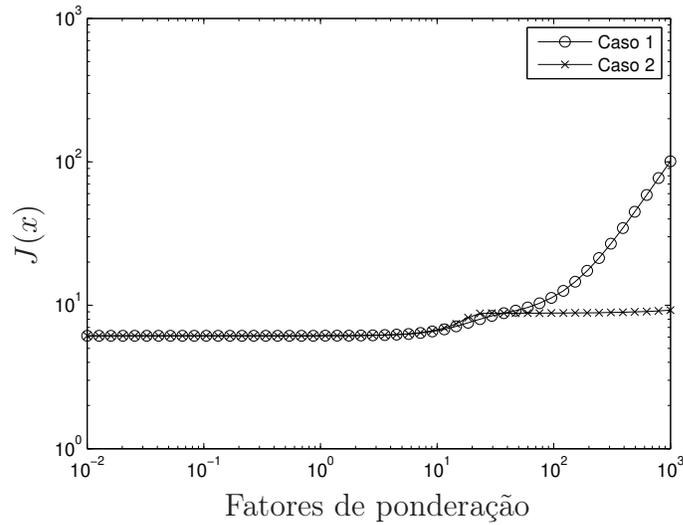


Figura 6.3: Variação do índice $J(x)$ para ambos os casos.

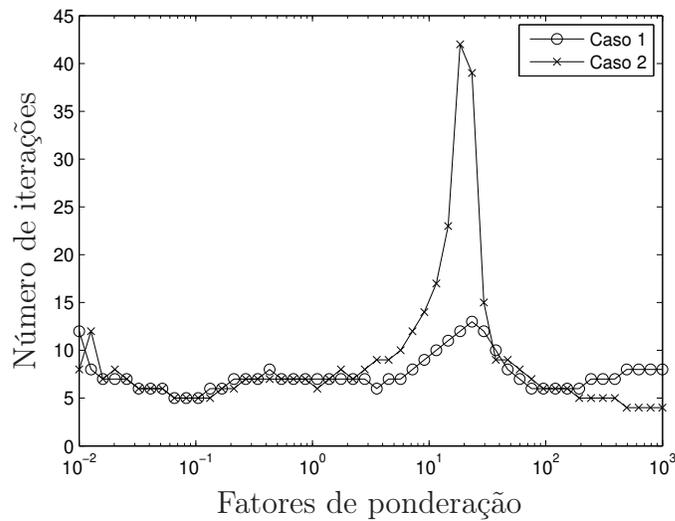


Figura 6.4: Número de iterações para ambos os casos.

injeção irrelevante na barra 6 aderir ao estado estimado.

A tabela 6.2 apresenta uma comparação entre o estado obtido após as 10 iterações realizadas e o estado real da rede. No ponto de parada, os resíduos das medidas de potência ativa e reativa na barra 6 são $-8,1929$ e $-7,5602$, respectivamente. Desconsiderando os resíduos das medidas da barra 6, o valor da função objetivo, $F(\hat{x})$ seria aproximadamente igual a $6,1259$, o que é um valor aceitável. Portanto, se as injeções da barra 6 fossem removidas o estado estimado para as ilhas observáveis seria confiável. Por fim, observe que embora os ângulos de fase das barras 10 e 11 sejam elevados, a abertura angular entre essas barras possui ordem de grandeza compatível com os resultados mostrados na tabela 5.8. Além disso, os

Tabela 6.1: Vetores de correção - Caso divergente.

n.i	Barra 5		Barra 11		Barra 12		Barra 13	
	ΔV	$\Delta\theta$	ΔV	$\Delta\theta$	ΔV	$\Delta\theta$	ΔV	$\Delta\theta$
1	0,0264	-0,1797	0,0627	-17,9486	0,0322	22,3805	0,0494	-0,3247
2	-0,0081	0,0120	0,0270	-2,3683	0,0209	-0,6130	-0,0048	0,0214
3	0,0001	0,0005	$< 10^{-4}$	2,2134	$< 10^{-4}$	-0,1461	0,0002	0,0007
4	-0,0001	$< 10^{-4}$	$< 10^{-4}$	-2,2016	$< 10^{-4}$	-0,2508	-0,0001	$< 10^{-4}$
5	0,0001	$< 10^{-4}$	$< 10^{-4}$	2,2449	$< 10^{-4}$	0,3153	0,0001	$< 10^{-4}$
6	-0,0001	$< 10^{-4}$	$< 10^{-4}$	-2,2083	$< 10^{-4}$	-0,1795	-0,0001	$< 10^{-4}$
7	0,0001	$< 10^{-4}$	$< 10^{-4}$	2,2356	$< 10^{-4}$	0,2873	0,0001	$< 10^{-4}$
8	-0,0001	$< 10^{-4}$	$< 10^{-4}$	-2,2190	$< 10^{-4}$	-0,2064	-0,0001	$< 10^{-4}$
9	0,0001	$< 10^{-4}$	$< 10^{-4}$	2,2246	$< 10^{-4}$	0,2511	0,0001	$< 10^{-4}$
10	-0,0001	$< 10^{-4}$	$< 10^{-4}$	-2,2254	$< 10^{-4}$	-0,2380	-0,0001	$< 10^{-4}$

Notas:

Valores de tensão em pu e ângulos em radianos

n.i - número de iterações

fluxos calculados para os ramos observáveis usando o estado estimado estão corretos, o que indica que o estado estimado para as ilhas observáveis é de boa qualidade.

Tabela 6.2: Estados da rede de 14 barras não observável 2 - Caso divergente.

Barra	Estado real		Estado estimado	
	V	θ	V	θ
1	1,0600	0,0000	1,0543	0,0000
2	1,0450	-5,3206	1,0416	-5,4528
3	1,0100	-13,4885	1,0127	-13,7010
4	1,0162	-10,9142	1,0163	-11,2312
5	1,0181	-9,2928	1,0183	-9,5838
6	1,0700	-15,0394	1,0567	-16,1631
7	1,0600	-14,1153	1,0000	0,0000
8	1,0900	-14,1153	1,0000	0,0000
9	1,0536	-15,7754	1,0676	3,4032
10	1,0487	-15,9453	1,0772	-1160,4169
11	1,0556	-15,6294	1,0897	-1160,3995
12	1,0543	-15,9375	1,0532	1237,6101
13	1,0491	-16,0197	1,0447	-17,3393
14	1,0327	-16,9365	1,0168	-6,8729

Nota: Valores de tensão em pu e ângulos em graus

6.2.2 Aplicação do estimador desacoplado regularizado

Considere novamente a rede não observável ilustrada na figura 6.1. Assim como nos estudos anteriores as pseudomedidas foram especificadas como o perfil plano de tensões e a ponderação adotada para as mesmas foi $\lambda^2 = 10^{-3}$. Adotando uma tolerância de 10^{-4} para os vetores de correção $\Delta\theta$ e ΔV , o estimador desacoplado rápido regularizado convergiu em 7 iterações ativas e 6 iterações reativas. A figura 6.5 ilustra os valores da função objetivo durante o processo iterativo. No ponto de convergência essa variável assume o valor de 8,6028 e a norma do seu gradiente é igual a 133,9316. Essas características mostram que, para o caso em estudo, o estimador desacoplado rápido regularizado não apresentou os problemas de convergência observados na versão completa do estimador regularizado. Para o caso em estudo a função objetivo $F(x)$ tem 24 graus de liberdade. Do ponto de vista da distribuição χ^2 , a probabilidade de obtermos um estado cujo valor de $F(x)$ seja menor que o valor obtido pelo estimador desacoplado é $P(F(x) \leq 8,6028) = 0,1671\%$. Isso indica que o estado estimado é estatisticamente aceitável. Além disso, $\|x^{real} - \hat{x}\|_2 = 0,5430$, o qual é compatível com os valores apresentados nas seções seguintes (vide tabela 6.5).

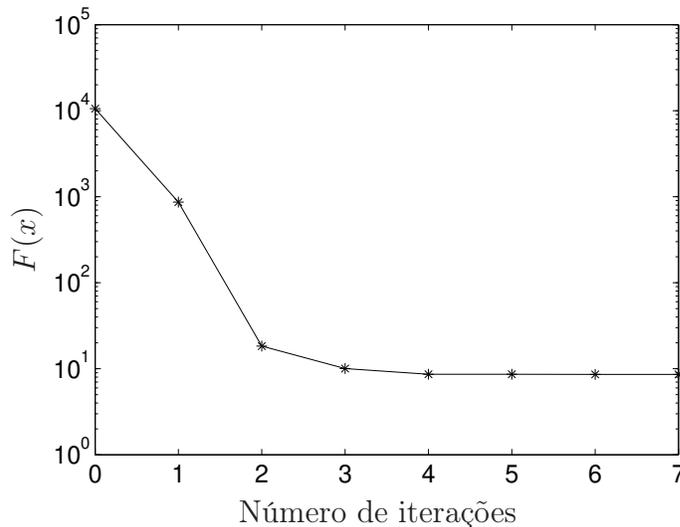


Figura 6.5: Comportamento da função objetivo no estimador desacoplado.

6.2.3 Alternativas preliminares

Na seção anterior foram mostradas as dificuldades numéricas que o estimador de estado regularizado completo encontra ao utilizar pseudomedidas de baixa qualidade em conjunto com valores muito pequenos para ponderar as pseudomedidas. Também foi possível verificar que os problemas de convergência ocorrem quando existem injeções irrelevantes particulares

no plano de medição. Com base nas observações realizadas é possível propor algumas alternativas para melhorar as características de convergência do estimador regularizado completo.

Alternativa 1 - Remoção das injeções irrelevantes

Uma maneira simples e robusta de evitar os problemas de convergência observados é remover as injeções irrelevantes do processo de estimação de estado. Para isso, é necessário realizar a análise de observabilidade e identificar as injeções irrelevantes antes de estimar o estado da rede. Note que a restauração da observabilidade não é requerida. Em termos práticos esse procedimento não traz esforço computacional adicional, já que a análise de observabilidade tem que ser realizada para que os estados estimados confiáveis sejam identificados.

A remoção das injeções irrelevantes não altera o estado estimado para as ilhas observáveis, entretanto, os eventuais ganhos na qualidade do estado estimado das porções não observáveis da rede são perdidos. Os resultados desta alternativa não são apresentados, todavia é intuitivo que eles sejam similares aos apresentados na seção 5.2.4.

Alternativa 2 - Alteração do critério de convergência

Apesar das oscilações observadas nos ângulos das tensões de barras adjacentes às injeções irrelevantes, constatou-se nos testes que os estados das barras observáveis apresentaram características de convergência adequadas. Assim, em princípio, é possível utilizar critérios de convergência locais baseados no estado das ilhas observáveis.

Nesse contexto, a convergência pode ser verificada, por exemplo, a partir dos fluxos nos ramos observáveis. Se os fluxos calculados nos ramos observáveis são aproximadamente constantes entre duas iterações consecutivas do processo de estimação, significa que o estado estimado para as ilhas observáveis está convergido. Entretanto, para que essa estratégia seja implementada é necessário realizar a análise de observabilidade e identificar os ramos observáveis. Alternativamente é possível monitorar apenas os fluxos nos ramos que possuem medidores, já que esses ramos são observáveis. Nesse caso, a análise de observabilidade não precisa ser realizada antes da estimação do estado. A variação nos fluxos de potência pode ser computada através da seguinte relação:

$$\Delta h_{km}^\nu = h(x)_{km}^\nu - h(x)_{km}^{\nu-1}$$

note que o cálculo de Δh_{km}^ν envolve apenas os valores do vetor $h(x)$ calculados na iteração anterior e na iteração corrente.

sejam de boa qualidade, é possível associar variâncias baixas, ou seja, pesos mais elevados às medidas providas pelas PMUs.

Na aplicação dessa estratégia, a rede da figura 6.1 teve as medidas de magnitude de tensão nas barras 1 e 11 substituídas por medidas fasoriais de tensão (módulo e ângulo) providas por PMUs, com variâncias iguais a 10^{-6} . Foram adotados os valores reais das variáveis de estado como valores medidos. Nesse caso, foram necessárias 5 iterações para o estimador convergir com uma tolerância de 10^{-5} para o valor máximo absoluto do vetor Δx . Resultados similares foram obtidos alocando medidas fasoriais de tensão em quaisquer barras das ilhas observáveis. Note que mesmo após a alocação das medidas fasoriais a rede ainda é não observável. No ponto de convergência do estimador a função objetivo atingiu o valor de 6,0945 comprovando que o estado aderiu bem às medidas.

Nesta terceira alternativa são requeridas apenas medidas fasoriais de tensão. Entretanto, sabe-se que as PMUs também fornecem medidas fasoriais de corrente. Certamente essas medidas podem contribuir para a melhoria do nível de redundância das medições, além de, dependendo da barra onde as PMUs forem alocadas, aumentar as ilhas observáveis. Destaca-se que as medições fasoriais de corrente também podem ser convertidas em tensões e incluídas no estimador regularizado.

Alternativa 4 - Controle do tamanho dos incrementos

A partir dos dados mostrados na tabela 6.1 é possível notar que os incrementos $\Delta\theta$ nos ângulos das barras 11 e 12 apresentam valores bastante elevados. Nesse sentido, uma metodologia para limitar os incrementos nas variáveis de estado durante o processo iterativo torna-se interessante. Talvez a maneira mais simples de se fazer isso, é simplesmente limitar o tamanho dos incrementos. Assim, durante o processo iterativo todos os valores que ultrapassam um limiar especificado assumem o valor do limiar. Em princípio, esses limiares podem ser ajustados empiricamente a partir de testes. Alternativamente, o controle dos incrementos pode ser feito através de um escalar α que deve ser multiplicado pelo passo Δx . Assim, em cada iteração o vetor de estado é atualizado da seguinte forma: $x^{\nu+1} = x^{\nu} + \alpha^{\nu} \Delta x^{\nu}$. Uma maneira de obter o multiplicador α é através da busca unidimensional, apresentada na seção 2.3.2.

A busca unidimensional inexata com interpolação quadrática foi aplicada ao problema em estudo. No caso em estudo o estimador de estado regularizado convergiu em 6 iterações com uma tolerância de 10^{-5} para o maior valor absoluto do vetor Δx . No ponto de convergência a função objetivo assume o valor de 6,1269. Utilizando a regra de Armijo, descrita na equação 2.25, a busca precisou ser realizada somente na primeira iteração e o passo calculado foi $\alpha^1 = 0,0111$. Nas demais iterações foi utilizado o passo completo, ou seja, $\alpha^{\nu} = 1$.

Na figura 6.7 está representada uma curva de superfície parametrizada pelo multiplicador α . Essa curva é definida pela seguinte equação: $\phi(\alpha) = F(x^0 + \alpha\Delta x^0)$. O eixo das ordenadas representa o valor da função objetivo $F(x)$ quando “caminhamos” ao longo da direção Δx^0 . A curva apresentada foi obtida na primeira iteração. Devido a presença de injeções irrelevantes e ao condicionamento numérico ruim da matriz ganho, o passo Δx^0 e, conseqüentemente, o intervalo $[x^0, x^0 + \Delta x^0]$ são demasiadamente grandes. Esses fatos associados às não linearidades do modelo $h(x)$ fazem com que a função $\phi(\alpha)$ seja multimodal, isto é, tenha vários máximos e mínimos locais. Como indica a figura 6.7 o menor valor de ϕ é obtido em $\alpha^* \approx 0,018$. Embora a busca realizada com interpolação quadrática tenha encontrado um ponto próximo de $\phi(\alpha^*)$, é evidente que a função interpoladora, $p(\alpha)$, não representa adequadamente a função $\phi(\alpha)$. Isso sugere que a busca unidimensional nem sempre é adequada ao problema.

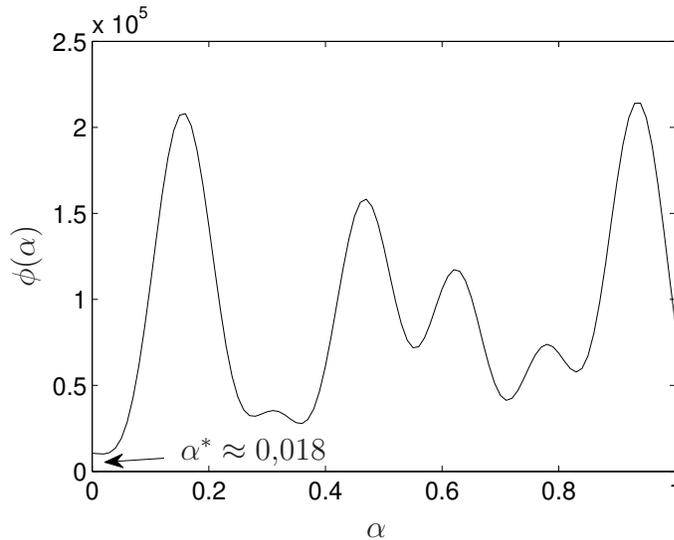


Figura 6.7: Múltiplos mínimos e máximos na busca unidimensional.

Alternativa 5 - Inicialização com estado de boa qualidade

Em todas as alternativas anteriores, o estimador de estado regularizado foi inicializado com o perfil plano de tensões, entretanto, é sabido que o estado estimado a partir do algoritmo de Gauss-Newton é fortemente dependente do ponto de inicialização do problema, principalmente em casos numericamente mal condicionados. Assim, nesta seção o estimador de estado regularizado foi inicializado com o estado estimado no momento em que a rede era observável. Esse estado encontra-se ilustrado nas figuras 4.2 e 4.3. A tabela 6.3 apresenta os resultados obtidos utilizando o algoritmo de Gauss-Newton convencional e com a busca unidimensional.

Tabela 6.3: Estado estimado para um ponto de inicialização alternativo.

Barra	Estado estimado			
	Sem busca		Com busca	
	V	θ	V	θ
1	1,0543	0,0000	1,0543	0,0000
2	1,0416	-5,4526	1,0416	-5,4525
3	1,0127	-13,7005	1,0127	-13,7002
4	1,0163	-11,2308	1,0163	-11,2305
5	1,0183	-9,5834	1,0183	-9,5832
6	1,0567	-16,1627	1,0567	-16,1622
7	1,0000	0,0000	1,0000	0,0000
8	1,0000	0,0000	1,0000	0,0000
9	1,0676	3,4034	1,0676	3,4034
10	1,0772	-362,1617	1,0772	-32,2067
11	1,0897	-362,1442	1,0897	-32,1892
12	1,0532	322,4860	1,0532	2,8885
13	1,0447	-17,3387	1,0447	-17,3382
14	1,0168	-6,8727	1,0168	-6,8726

Nota: Valores de tensão em pu e ângulos em graus

Ambos os estados apresentados na tabela 6.3 correspondem a mínimos locais de $F(x)$. Note a diferença entre a magnitude dos ângulos de fase das barras 10, 11 e 12. É importante salientar que mesmo com um ponto de inicialização melhor, ainda assim foi necessário utilizar a busca unidimensional para obter um estado mais próximo do ponto de operação real da rede. Destaca-se que nesta seção também foram utilizados pesos iguais a 10^{-3} para as pseudomedidas, entretanto, caso o ponto de inicialização fosse utilizado como pseudomedidas, pesos maiores poderiam ter sido adotados já que as pseudomedidas seriam de boa qualidade. Isso certamente melhoraria as características de convergência do estimador.

6.3 Os Métodos Propostos

Conforme discutido, a falta de uma metodologia eficiente para se obter as ponderações adequadas para as pseudomedidas de estado, nos obriga a ser pragmáticos e simplesmente adotar valores pequenos para essas ponderações. Como foi exposto na seção anterior, sob essas circunstâncias, o estimador de estado regularizado pode apresentar problemas de convergência. Nesse contexto, foram discutidas cinco alternativas preliminares para lidar com esse problema. Baseando-se nas discussões realizadas ao longo do trabalho e buscando apro-

veitar as ferramentas disponíveis em um estimador de estado convencional, nesta seção serão propostos dois algoritmos que buscam aumentar a confiabilidade do estimador de estado regularizado.

6.3.1 Versão 1 - Inicialização com o estimador desacoplado

Diferentemente da versão completa, o estimador de estado regularizado desacoplado rápido geralmente não apresenta problemas de convergência, mesmo na presença de injeções irrelevantes. Logo, a versão desacoplada pode ser utilizada para inicializar a versão completa. Assim, a solução do estimador desacoplado rápido é usada como ponto de inicialização do estimador completo e na definição dos valores das pseudomedidas na versão completa.

Considere novamente a rede estudada na seção 6.2. Inicialmente as pseudomedidas de estado foram ajustadas como o perfil plano de tensões. O estimador desacoplado rápido também foi inicializado com essa condição. Após 4 iterações ativas e 3 iterações reativas o estimador desacoplado convergiu com uma tolerância de 10^{-2} . Em seguida, a solução desacoplada foi utilizada para reajustar as pseudomedidas e inicializar a versão completa. Realizadas 6 iterações o estimador regularizado convergiu com a tolerância de 10^{-5} . Ao término do processo de estimação a função objetivo $F(x)$ assume o valor de 6,1261.

A figura 6.8 ilustra o comportamento da função objetivo $F(x)$ e da norma do seu gradiente durante o processo iterativo realizado pelo estimador regularizado completo. Perceba que na primeira iteração ambas as curvas apresentam um incremento. Esse comportamento vai na contramão dos métodos de descida, que buscam, a cada iteração, reduzir o valor da função objetivo. Por outro lado, nas iterações seguintes existe um claro decaimento da função objetivo e por conseguinte da norma do seu gradiente.

Destaca-se que não há comprovação matemática de que essa abordagem possa sempre tornar casos divergentes em convergentes.

Todo estimador de estado deve, depois que o processo de estimação é finalizado, verificar se o estado obtido adere bem às medidas disponíveis. Uma condição necessária para isso é que o gradiente da função objetivo tenda a zero, isto é, $\|\nabla J(\hat{x})\|_2 \leq \epsilon$. Entretanto, apenas essa condição não garante que o estado estimado é confiável. De fato, em um ponto confiável o índice $F(\hat{x})$ assume um valor da mesma ordem de grandeza do seu número de graus de liberdade, ou equivalentemente, os resíduos de estimação são pequenos. Por isso faz-se necessário realizar o processamento de erros grosseiros após a estimação.

Nos testes realizados verificou-se que nos casos divergentes, os resíduos normalizados de algumas medidas de injeção irrelevantes são elevados. Esse fato sugere que o processamento de erros grosseiros, utilizando o máximo resíduo normalizado, pode ser utilizado em conjunto com o estimador de estado regularizado. A ideia principal consiste em remover do plano de

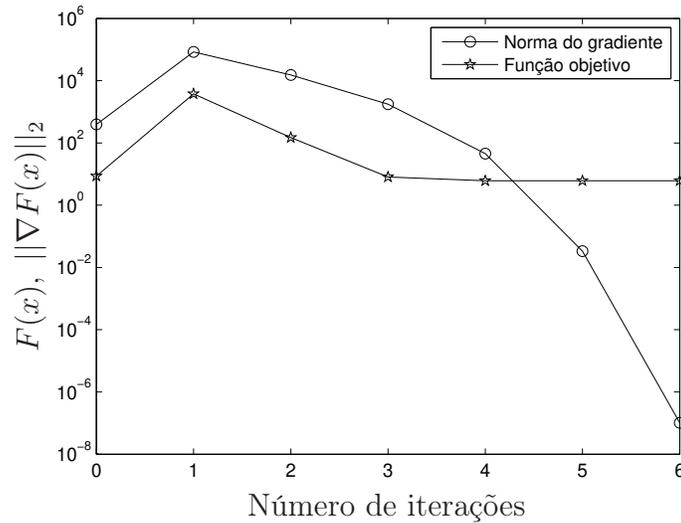


Figura 6.8: Comportamento de um caso convergente.

medição não só as medidas portadoras de erros grosseiros, mas também as injeções irrelevantes que dificultam a convergência do estimador. Nesse sentido, propõe-se o seguinte algoritmo:

Algoritmo 6.1:

- i* Defina uma ponderação pequena para as pseudomedidas. Esse valor pode ser, por exemplo, o inverso das ponderação das medidas;
- ii* Ajuste u_A e u_R com a melhor informação *a priori* disponível (perfil plano, última estimativa, previsor de carga, etc). Resolva as equações 5.10 e 5.11 até que $\Delta\theta^\nu \leq \epsilon$ e $\Delta V^\nu \leq \epsilon$;
- iii* Ajuste o vetor u com a solução obtida em *ii*;
- iv* Inicialize o estimador regularizado completo com a solução obtida em *ii*. Resolva a equação 5.4 até que o critério de parada seja atendido;
- v* Calcule os resíduos normalizados de acordo com 4.5. Se o maior resíduo normalizado for maior que 3 remova a medida associada e retorne ao passo *iv*. Caso contrário, um estado confiável foi obtido.

6.3.2 Versão 2 - Redução no conjunto de pseudomedidas

Na seção 5.2 definiu-se o modo de formação da matriz L . A matriz L define as barras que conterão as pseudomedidas de variáveis de estado. Até aqui foram alocadas pseudomedidas em todas as barras que não contêm medição direta das variáveis de estado. Dessa

forma, algumas pseudomedidas serão redundantes às medidas disponíveis no plano de medição. Daí a necessidade de se utilizar fatores de ponderação pequenos, evitando, pois, a deterioração do estado das porções observáveis da rede. Nesse sentido, se fossem alocadas apenas pseudomedidas críticas, as interferências dessas pseudomedidas no estado das ilhas observáveis seriam reduzidas e as ponderações das mesmas poderiam ser da mesma ordem de grandeza das ponderações das medidas. A seguir será apresentada uma maneira de incluir no estimador de estado regularizado um número menor de pseudomedidas.

Quando uma rede elétrica não é observável a sua matriz ganho, G , é simétrica e positiva semidefinida, isto é, ela possui deficiência de posto. Como mostrado em (Monticelli and Wu, 1985a), durante a fatoração de uma matriz que possua essas características se um pivô nulo é encontrado, a respectiva matriz possui a forma ilustrada na figura 6.9.

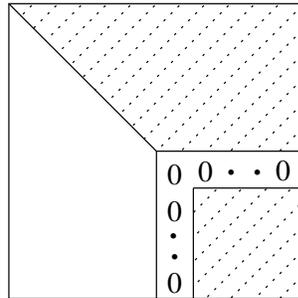


Figura 6.9: Pivô nulo na matriz ganho fatorada.

Fonte: (Monticelli and Wu, 1985a)

Uma maneira de aumentar o posto da matriz ilustrada na figura 6.9 é inserindo o valor 1 na posição onde o pivô nulo foi encontrado. Na estimação de estado isso corresponde a alocar uma pseudomedida de variável de estado na barra correspondente. Além disso, a pseudomedida alocada possui a característica de ser crítica (Monticelli and Wu, 1985a).

Em princípio, ao resolvermos a equação 4.2 do estimador convencional, durante a fatoração da matriz ganho, $G(x^\nu) = H(x^\nu)^T R_z^{-1} H(x^\nu)$, as pseudomedidas poderiam ser alocadas e o problema de deficiência de posto seria contornado. Todavia, foram verificados problemas numéricos com essa abordagem, mesmo utilizando fatores de ponderação mais elevados para as pseudomedidas. As origens desses problemas numéricos não foram exploradas. Assim, a alocação de pseudomedidas adotada foi baseada nos modelos desacoplados $P - \theta$ e $Q - V$. Da análise do modelo $P - \theta$ são alocadas as pseudomedidas de ângulo e da análise do modelo $Q - V$ são alocadas as pseudomedidas de módulo das tensões, como segue.

A versão desacoplada de um estimador de estado convencional é descrita pelas seguintes equações:

$$[H_{AA}^T R_A^{-1} H_{AA}] \Delta \theta^\nu = H_{AA}^T R_A^{-1} \Delta z_A(\theta^\nu, V^\nu) \quad (6.1)$$

$$\theta^{\nu+1} = \theta^\nu + \Delta \theta^\nu$$

$$[H_{RR}^T R_R^{-1} H_{RR}] \Delta V^\nu = H_{RR}^T R_R^{-1} \Delta z_R(\theta^{\nu+1}, V^\nu) \quad (6.2)$$

$$V^{\nu+1} = V^\nu + \Delta V^\nu$$

Considerando uma rede elétrica não observável, a solução do estimador de estado desacoplado poderia ser obtida da seguinte maneira: Durante a fatoração da matriz $G_{P\theta} = H_{AA}^T R_A^{-1} H_{AA}$, aloque pseudomedidas de ângulo de fase sempre que um pivô nulo for encontrado, e durante a fatoração da matriz $G_{QV} = H_{RR}^T R_R^{-1} H_{RR}$, aloque pseudomedidas de magnitude de tensão sempre que um pivô nulo aparecer. Cabe ressaltar que as pseudomedidas alocadas nos modelos desacoplados podem, eventualmente, não ser exatamente críticas para o modelo completo.

Considerando os aspectos discutidos acima, propõe-se o seguinte algoritmo:

Algoritmo 6.2:

- i* Defina uma ponderação para as pseudomedidas. Esse valor pode ser, por exemplo, da mesma ordem de grandeza das ponderações das medidas;
- ii* Resolva o estimador desacoplado das equações 6.1 e 6.2 alocando pseudomedidas conforme estabelecido;
- iii* Ajuste o vetor u com a melhor informação *a priori* disponível (perfil plano, última estimativa, previsor de carga, etc);
- iv* Ajuste a matriz L considerando somente as pseudomedidas críticas alocadas no estimador desacoplado do item *ii*;
- v* Inicialize o estimador regularizado completo com a solução obtida em *ii*. Resolva a equação 5.4 do estimador regularizado completo até que o critério de parada seja atendido.
- vi* Calcule os resíduos normalizados de acordo com 4.5. Se o maior resíduo normalizado for maior que 3 remova a medida associada e retorne ao passo *v*. Caso contrário, um estado confiável foi obtido.

6.4 Testes Finais

Com o objetivo de avaliar o desempenho dos algoritmos 1 e 2 propostos, foram realizadas simulações cujos resultados e análises mais importantes são mostrados nesta seção. Foram utilizados os sistemas testes de 14 e 118 barras do IEEE. Em todos os testes foram usadas somente medidas convencionais de magnitude de tensão, fluxos e injeções de potências, ou seja, medidas fasoriais não foram consideradas. As medidas de potência foram consideradas aos pares e os valores para as variâncias das medidas são mostrados na tabela 4.2. O critério de convergência adotado foi $\max|\Delta x| \leq 10^{-5}$, além disso limitou-se os processos de cálculos em 10 iterações.

6.4.1 Sistema teste IEEE 14 barras

Considere mais uma vez a rede elétrica mostrada na figura 6.1. Lembre-se que a rede possui duas ilhas observáveis, sendo que a ilha 1 é formada pelas barras $\{1,2,3,4,5,6,13\}$ e a ilha 2 pelas barras $\{10,11\}$, as demais barras estão isoladas. A ilha 1 é composta pelos ramos observáveis $\{1,2,3,4,5,6,7,10,13\}$ e a ilha 2 pelo ramo $\{18\}$. No plano de medição é possível verificar que as medidas de injeção nas barras 6 e 14 são irrelevantes. Nos testes a seguir foram utilizados os valores medidos dispostos na tabela 6.4.

Tabela 6.4: Valores reais e medidos para a rede de 14 barras não observável 2.

Tipo	Real	Medido	Tipo	Real	Medido	Tipo	Real	Medido
P_1	2,4692	2,4932	Q_{1-2}	-0,2276	-0,2573	P_{5-4}	0,6480	0,6705
Q_1	-0,1862	-0,1993	P_{1-5}	0,7981	0,7509	Q_{5-4}	-0,1488	-0,2013
P_2	0,1722	0,1611	Q_{1-5}	0,0414	0,0095	V_1	1,0600	1,0774
Q_2	0,3517	0,3439	P_{2-3}	0,7707	0,7696	V_4	1,0162	0,9475
P_3	-0,9891	-0,9757	Q_{2-3}	0,0320	0,0156	V_5	1,0181	1,0227
Q_3	0,0885	0,0696	P_{5-6}	0,4644	0,4731	V_6	1,0700	1,0912
P_6	-0,1176	-0,1973	Q_{5-6}	0,1205	0,1352	V_{11}	1,0556	1,0649
Q_6	0,0733	0,0641	P_{6-13}	0,1869	0,2043	V_{12}	1,0543	1,0349
P_{14}	-0,1565	-0,1978	Q_{6-13}	0,0777	0,0709	V_{14}	1,0327	1,0323
Q_{14}	-0,0525	-0,0912	P_{10-11}	-0,0404	-0,0178	—	—	—
P_{1-2}	1,6711	1,7029	Q_{10-11}	-0,0201	-0,0527	—	—	—

Nota: Valores dados em pu

Teste 1

No primeiro teste admitiu-se que não existe nenhuma informação *a priori* sobre o estado da rede, portanto, utilizou-se o perfil plano de tensões como pseudomedidas de variáveis de estado. No teste do primeiro algoritmo utilizou-se a ponderação de 10^{-3} para as pseudomedidas e no testes do segundo algoritmo a ponderação utilizada foi 10^3 .

No primeiro algoritmo, o estimador de estado desacoplado rápido realizou 3 iterações ativas e 2 iterações reativas com uma tolerância de 10^{-2} . Em seguida, o estimador regularizado completo foi encerrado após 10 iterações. No ponto de parada a norma do gradiente da função objetivo é igual a 240,9850, indicando, pois, que o estimador não encontrou um ponto convergente. Os resíduos normalizados das medidas foram calculados e a medida de injeção de potência ativa na barra 6 foi apontada como medida espúria. Na figura 6.10 é possível verificar que os resíduos normalizados 11 e 23, associados às medidas de injeção ativa e reativa na barra 6, assumem os valores mais elevados. Não obstante, o elevado resíduo normalizado não significa que essas medidas contêm erros grosseiros. De fato, devido a problemas de condicionamento numérico, essas medidas não conseguem se ajustar às medidas e pseudomedidas adjacentes. Como essas injeções são irrelevantes, sua remoção não provoca distúrbios no estado estimado das porções observáveis. Como sugere o algoritmo 6.1, depois da exclusão da medida P_6 o estimador regularizado completo foi executado novamente e a convergência foi alcançada após 3 iterações.

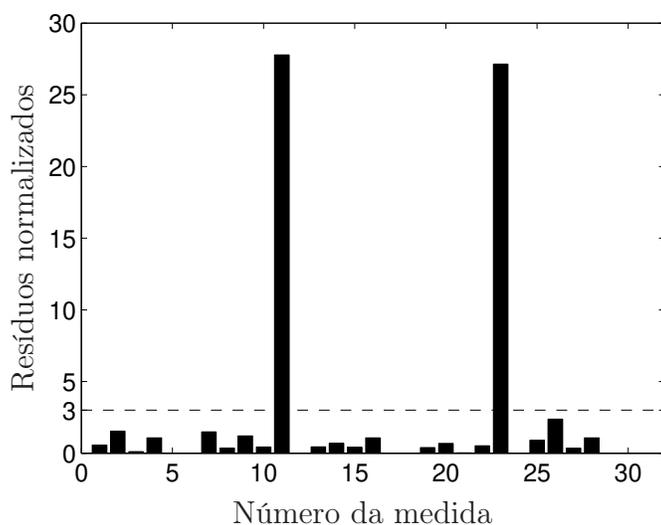


Figura 6.10: Resíduos de estimação normalizados no teste 1 - Versão 1.

No segundo algoritmo, o estágio desacoplado convergiu em 4 iterações ativas e 3 iterações reativas com uma tolerância de 10^{-2} . Durante esse primeiro estágio verificou-se a necessidade de alocar pseudomedidas de ângulo de fase nas barras 7, 8, 12 e 14, e medi-

das de magnitude de tensão nas barras 7 e 8. O segundo estágio convergiu em 4 iterações. Constatou-se que a medida de ângulo na barra 12 não é crítica já que o seu resíduo assume o valor de $7,1718 \times 10^{-3}$.

A figura 6.11 ilustra as magnitudes de tensão estimadas pelos dois algoritmos em comparação com os valores reais dessas grandezas. Da mesma forma na figura 6.12 estão dispostos os valores estimados e reais das aberturas angulares dos ramos. Perceba que para as barras correspondentes a ilha observável 1 os resultados obtidos para as magnitudes das tensões são coerentes com os valores reais das variáveis, já para a ilha 2 os resultados da versão 1 são superiores aos da versão 2. As aberturas angulares estimadas para os ramos observáveis das duas ilhas são de boa qualidade, independentemente da versão do estimador. Por outro lado, como era esperado, para as barras e ramos não observáveis ambos estados estimados são ruins, já que as pseudomedidas incluídas são de baixa qualidade.

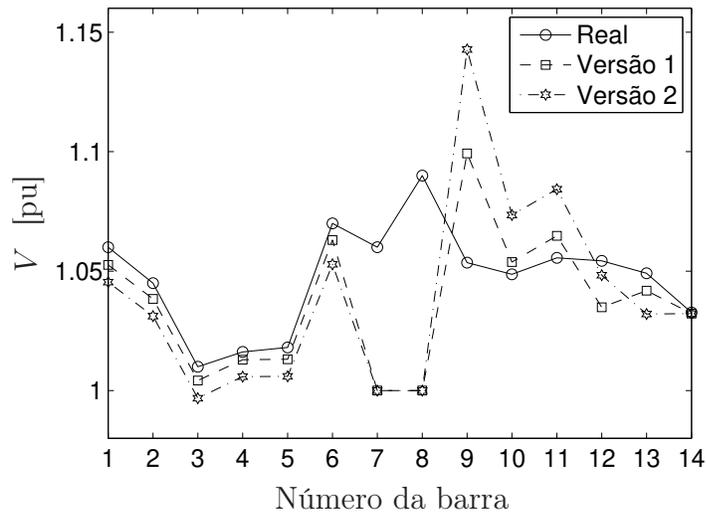


Figura 6.11: Magnitudes das tensões estimadas - Teste 1.

A tabela 6.5 apresenta uma comparação dos algoritmos para diferentes valores de ponderação das pseudomedidas. No caso do primeiro algoritmo, o uso de ponderações elevadas para as pseudomedidas pode fazer com que o estimador descarte medidas, tais como o fluxo P_{5-6} , que não são irrelevantes nem portadoras de erros grosseiros. Por outro lado, no caso do segundo algoritmo, ponderações baixas para as pseudomedidas fazem com que o estimador também tenha que descartar medidas de injeção irrelevantes para que a convergência possa ser obtida.

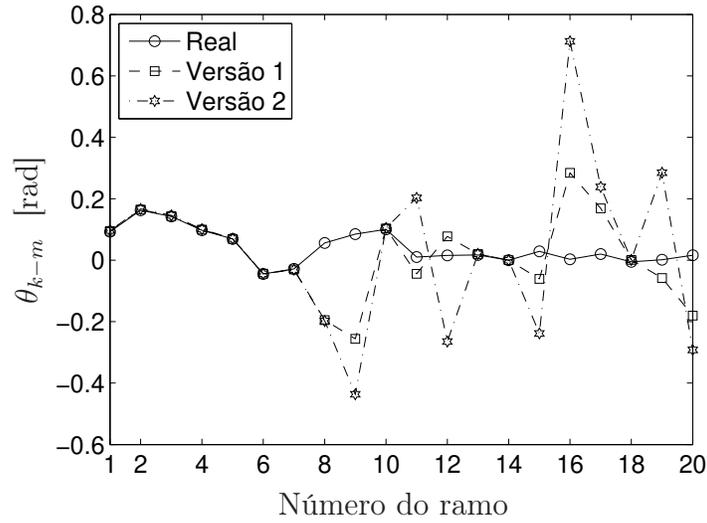


Figura 6.12: Aberturas angulares estimadas - Teste 1.

Tabela 6.5: Comparação entre os resultados do teste 1.

Versão	λ^2	$J(\hat{x})$	$F(\hat{x})$	$\ \nabla F(\hat{x})\ _2$	$\ x^{real} - \hat{x}\ _2$	M. d.	n° iter.		
							A	R	C
1	10^{-3}	9,2440	9,2440	$8,4665 \times 10^{-6}$	0,5426	P_6	3	2	13
	10^{-2}	9,2440	9,2440	$8,4671 \times 10^{-6}$	0,5426	P_6	3	2	13
	10^{-1}	9,2442	9,2447	$3,1603 \times 10^{-1}$	0,5344	—	3	2	10
	10^0	9,2454	9,2465	$1,5553 \times 10^{-5}$	0,5378	—	3	2	7
	10^1	9,2479	9,2486	$7,0012 \times 10^{-6}$	0,5413	—	3	2	4
	10^2	9,2670	9,3947	$2,5888 \times 10^{-5}$	0,5405	—	3	2	3
	10^3	10,8184	11,7820	$7,1975 \times 10^{-5}$	0,5861	P_{5-6}	3	2	3
2	10^{-3}	9,2440	9,2440	$1,2984 \times 10^{-4}$	0,7962	Q_6	4	3	13
	10^{-2}	9,2475	9,2482	$7,4287 \times 10^1$	0,7022	—	4	3	10
	10^{-1}	9,2446	9,2504	$8,0409 \times 10^{-3}$	0,7038	—	4	3	10
	10^0	9,2526	9,2906	$7,4373 \times 10^{-5}$	0,7113	—	4	3	10
	10^1	9,3393	9,4951	$2,2764 \times 10^{-5}$	0,7354	—	4	3	9
	10^2	9,7615	9,9650	$1,1553 \times 10^{-4}$	0,7791	—	4	3	6
	10^3	10,1995	10,2509	$2,2764 \times 10^{-5}$	0,8052	—	4	3	4

Notas:

M. d. - Medidas descartadas

A - Meia iteração ativa

R - Meia iteração reativa

C - Iteração completa

Teste 2

Neste teste admitiu-se uma perda temporária da observabilidade da rede e, portanto, o estado estimado da rede em um instante anterior é conhecido. Admitiu-se ainda que no

instante atual a carga é 5 % superior a observada no instante anterior. Nos estudos a seguir o estado estimado no instante anterior é utilizado como informação *a priori*. Note que nesse caso as pseudomedidas são mais precisas. As medidas utilizadas são as mesmas que estão dispostas na tabela 6.4.

No primeiro estágio do primeiro algoritmo, o estimador desacoplado rápido executou 3 iterações ativas e 2 reativas. Da mesma forma que no teste anterior, na primeira execução do estimador completo o estado estimado não é confiável. De fato, após a execução de 10 iterações do estimador completo, a medida de injeção reativa na barra 6, Q_6 , é apontada no teste do máximo resíduo normalizado, já que o seu resíduo normalizado é igual a $-80,4805$. Após a retirada dessa medida o estimador regularizado completo convergiu em 3 iterações.

No segundo algoritmo, o estágio desacoplado convergiu em 4 iterações ativas e 3 iterações reativas. Assim como no teste anterior, durante esse primeiro estágio verificou-se a necessidade de alocar pseudomedidas de ângulo de fase nas barras 7, 8, 12 e 14, e medidas de magnitude de tensão nas barras 7 e 8. O segundo estágio convergiu em apenas 2 iterações.

As figuras 6.13 e 6.14 ilustram os valores reais e estimados para as magnitudes de tensão nas barras e para as aberturas angulares dos ramos, respectivamente. Pode-se notar que os resultados são praticamente os mesmos, muito embora segundo os valores dispostos na tabela 6.6 seja possível notar uma ligeira superioridade do primeiro algoritmo. O estado estimado é de boa qualidade para todas as barras do sistema. Destaca-se que em ambos os testes o tempo de execução do segundo algoritmo é menor, já que ele requer menos iterações.

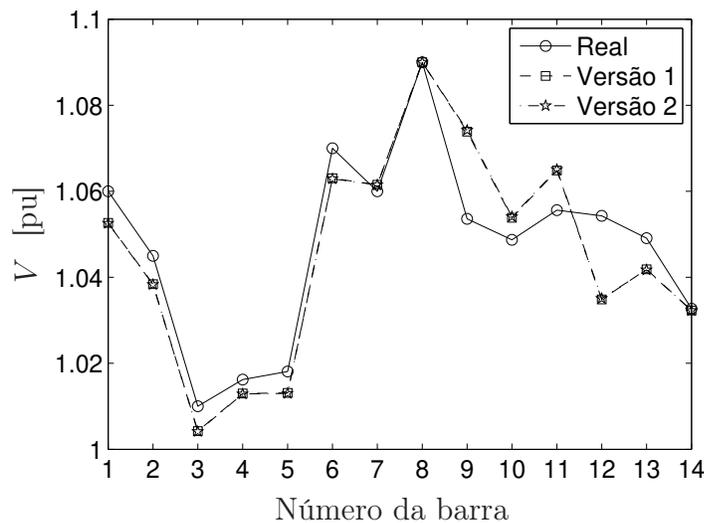


Figura 6.13: Magnitudes das tensões estimadas - Teste 2.

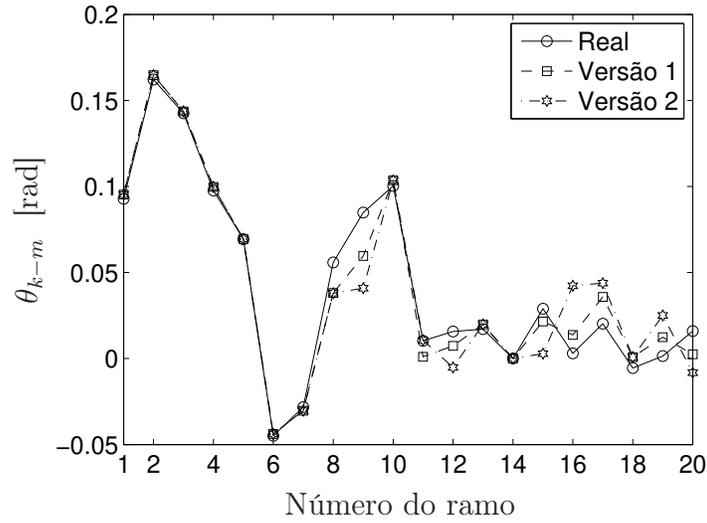


Figura 6.14: Aberturas angulares estimadas - Teste 2.

Tabela 6.6: Comparação entre os resultados do teste 2.

Versão	λ^2	$J(\hat{x})$	$F(\hat{x})$	$\ \nabla F(\hat{x})\ _2$	$\ x^{real} - \hat{x}\ _2$	M. d.	n° iter.		
							A	R	C
1	10^{-3}	9,2440	9,2440	$4,2726 \times 10^{-6}$	0,0473	Q_6	3	2	13
	10^{-2}	9,2502	9,2502	$9,9888 \times 10^1$	0,0471	—	3	2	10
	10^{-1}	9,2446	9,2446	$2,8212 \times 10^1$	0,0494	—	3	2	10
	10^0	9,2441	9,2441	$6,9382 \times 10^{-7}$	0,0474	—	3	2	4
	10^1	9,2441	9,2442	$1,6209 \times 10^{-3}$	0,0474	—	3	2	2
	10^2	9,2448	9,2473	$1,5363 \times 10^{-3}$	0,0467	—	3	2	2
	10^3	9,3264	9,3880	$1,5179 \times 10^{-6}$	0,0427	—	3	2	3
2	10^{-3}	9,2440	9,2440	$1,4840 \times 10^{-3}$	0,0613	P_6	4	3	12
	10^{-2}	9,2440	9,2440	$1,4840 \times 10^{-3}$	0,0613	P_6	4	3	12
	10^{-1}	9,2481	9,2481	$7,9588 \times 10^1$	0,0661	—	4	3	10
	10^0	9,2441	9,2441	$3,8947 \times 10^{-6}$	0,0606	—	4	3	6
	10^1	9,2441	9,2441	$1,1444 \times 10^{-3}$	0,0612	—	4	3	2
	10^2	9,2441	9,2441	$1,1525 \times 10^{-3}$	0,0613	—	4	3	2
	10^3	9,2441	9,2441	$1,1536 \times 10^{-3}$	0,0613	—	4	3	2

Notas:

M. d. - Medidas descartadas

A - Meia iteração ativa

R - Meia iteração reativa

C - Iteração completa

6.4.2 Sistema teste IEEE 118 barras

Considere a rede de 118 barras do IEEE mostrada na figura 6.15. Os parâmetros elétricos dessa rede podem ser encontrados em (PSTCA, 2013). A barra 69 foi adotada como

barra de referência angular. Nesse sistema existem 358 medidas, compostas por 54 medidas de magnitude de tensão, 42 pares de medidas de injeção ativa e reativa e 110 pares de medidas de fluxo de potência ativa e reativa. As variâncias dessas medidas estão dispostas na tabela 4.2. Note que esse conjunto de medidas não torna a rede completamente observável. De fato, esse sistema é composto por 7 ilhas observáveis e uma barra isolada. Na figura 6.15 é possível identificar os 23 ramos não observáveis. Para essa configuração do sistema de medição existem 3 pares de medidas de injeção irrelevantes. Essas injeções irrelevantes encontram-se nas barras 24, 49 e 65.

Conforme comentado em seções anteriores, durante a operação real de uma rede é possível obter informações *a priori* mais apuradas que o perfil plano de tensões. Para ilustrar esse fato, utilizou-se um conjunto de pseudomedidas de estado que representam um ponto com carga 5 % inferior ao ponto de operação em que as medidas foram obtidas. Ao aplicarmos os algoritmos propostos, no primeiro algoritmo foi utilizada a ponderação de 10^{-3} para as pseudomedidas, enquanto que no segundo algoritmo a ponderação utilizada foi 10^3 .

No primeiro estágio do primeiro algoritmo, o estimador desacoplado rápido executou 3 iterações ativas e 2 reativas. Após o término do primeiro ciclo de estimação o estimador regularizado completo não consegue obter um estado confiável. De fato, ao término do processo iterativo a medida de injeção ativa na barra 49 é apontada no teste do máximo resíduo normalizado, já que o seu resíduo normalizado é igual a $-1,8288 \times 10^5$. Após a retirada dessa medida o estimador regularizado completo converge em 8 iterações.

Ao aplicar o segundo algoritmo, o estágio desacoplado convergiu em 3 iterações ativas e 2 iterações reativas. Durante esse primeiro estágio foram alocadas pseudomedidas de ângulo de fase nas barras 66, 67, 116 e 117. O segundo estágio convergiu em 6 iterações. Enquanto a pseudomedida 67 apresenta um resíduo de estimação muito baixo, os resíduos das pseudomedidas 66, 116 e 117 são iguais a 0,0057, $-0,0062$ e $-0,0015$, respectivamente.

A tabela 6.7 apresenta um resumo dos resultados obtidos pelos algoritmos propostos e pela versão inicial do estimador regularizado (vide equações 5.4). Quando comparamos os resultados levando em conta somente os valores das funções objetivo $J(\hat{x})$ e $F(\hat{x})$, verifica-se que o primeiro algoritmo tem um desempenho ligeiramente superior ao segundo. Note que os valores obtidos para $F(\hat{x})$ são menores que o seu número de graus de liberdade, isto é, 304. Basicamente, esses valores relativamente pequenos das funções objetivo indicam que o estado estimado para as porções observáveis conseguiu se adequar bem às medidas disponíveis. Não obstante, pode-se verificar que a norma do erro do estado estimado pelo primeiro algoritmo é bastante superior à norma do erro obtido pelo segundo algoritmo. Esse erro elevado está associado, principalmente, aos ângulos de fase estimados. Embora os ângulos de fase estimados pelo primeiro algoritmo tenham se adequado bem às medidas, as suas magnitudes diferem

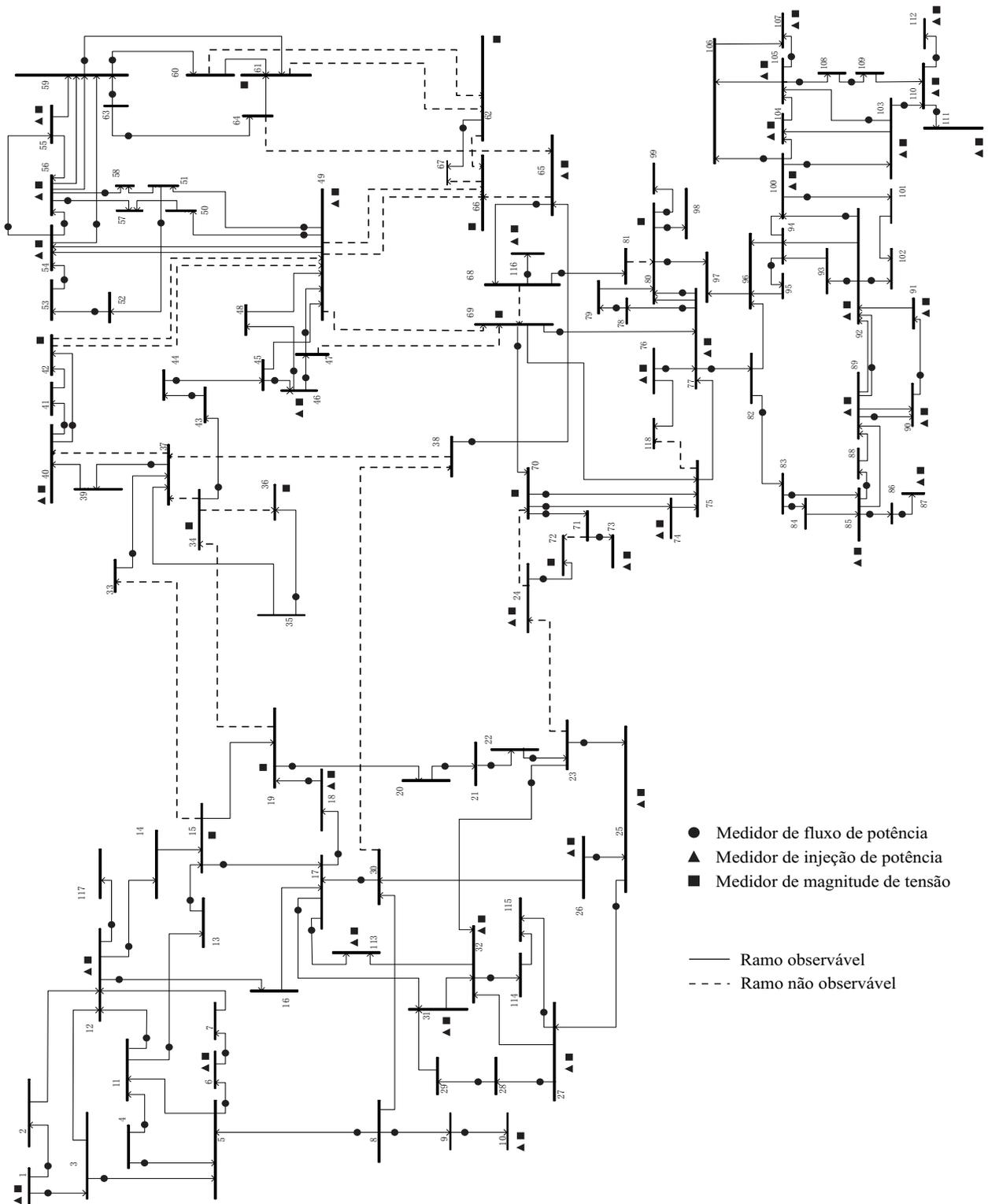


Figura 6.15: Rede de 118 barras do IEEE não observável.

bastante dos seus valores reais. Todavia, é sabido que as equações de fluxo de potência são funções das aberturas angulares entre barras adjacentes. A figura 6.16 apresenta as diferen-

ças entre as aberturas angulares reais e estimadas. Perceba que para os ramos observáveis esse erro é pequeno para ambas as versões. Por outro lado, nos ramos não observáveis o erro obtido pelo primeiro algoritmo é bastante grande atingindo valores em torno de 1 radiano. Para o segundo algoritmo o maior erro atinge valores em torno de 0,1 radianos.

Os resultados da versão inicial do estimador regularizado, expostos na tabela 6.7, indicam que, mesmo com um conjunto de informações *a priori* mais apurado, a utilização de ponderações pequenas impede a convergência do algoritmo. Além disso, o uso de ponderações elevadas pode fazer com que medidas que inicialmente eram aceitáveis sejam detectadas no processamento de erros grosseiros. De fato, para o caso onde as pseudomedidas foram ponderadas com o fator 10^3 foram necessários quatro ciclos de estimação, tendo em vista que os fluxos ativos nos ramos 119, 108 e 123 foram identificados como portadores de erros grosseiros. Note que na tabela 6.7, para esse caso, estão dispostos apenas os resultados para o primeiro ciclo de estimação.

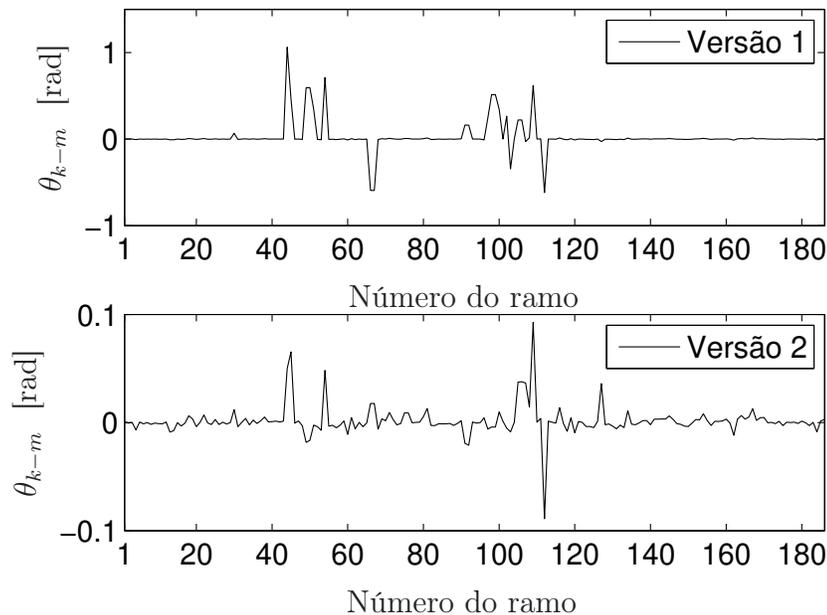


Figura 6.16: Erro das aberturas angulares estimadas para a rede de 118 barras.

6.5 Conclusões

Neste capítulo foi dada continuidade aos estudos acerca do estimador de estado regularizado, apresentado no capítulo anterior. Foram apresentados casos onde esse estimador apresenta dificuldades de convergência. Constatou-se que esses problemas estão relacionados aos fatores de ponderação reduzidos para as pseudomedidas e também à presença de injeções

Tabela 6.7: Comparação entre os resultados do teste com a rede de 118 barras.

Versão	λ^2	$J(\hat{x})$	$F(\hat{x})$	$\ \nabla F(\hat{x})\ _2$	$\ x^{real} - \hat{x}\ _2$	M. d.	n° iter.		
							A	R	C
1	10^{-3}	116,8667	116,8816	$2,3706 \times 10^{-5}$	4,3566	P_{49}	3	2	18
	10^{-2}	116,8711	117,0123	$3,5687 \times 10^{-7}$	4,2798	P_{49}	3	2	20
	10^{-1}	118,8117	119,4855	$7,6862 \times 10^{-5}$	3,1447	—	3	2	10
	10^0	120,2788	120,3717	$1,3426 \times 10^{-4}$	0,9337	—	3	2	5
2	10^0	118,5198	118,8081	$7,1723 \times 10^0$	4,1508	—	3	2	10
	10^1	119,5267	119,9158	$2,7542 \times 10^{-5}$	2,5740	—	3	2	10
	10^2	120,3534	120,4150	$7,5362 \times 10^{-5}$	0,8803	—	3	2	8
	10^3	120,4914	120,5655	$1,4052 \times 10^{-5}$	0,6094	—	3	2	6
Init	10^{-3}	—	—	—	—	—	—	—	div
	10^{-2}	—	—	—	—	—	—	—	div
	10^{-1}	118,8117	119,5015	$7,6163 \times 10^{-5}$	3,1447	—	—	—	8
	10^0	120,2790	120,5315	$1,3596 \times 10^{-4}$	0,9337	—	—	—	6
	10^1	120,4653	122,0408	$8,0160 \times 10^{-3}$	0,7001	—	—	—	4
	10^2	121,8639	134,8773	$6,0478 \times 10^{-3}$	0,6765	—	—	—	4
	10^3	155,7007	202,5720	$1,4104 \times 10^{-2}$	0,6908	*	—	—	4

Notas:

Init - Versão inicial

M. d. - Medidas descartadas

A - Meia iteração ativa

R - Meia iteração reativa

C - Iteração completa

div - Estimador divergiu

* - Foram detectados e identificados erros grosseiros

irrelevantes específicas no plano de medição. Nesse sentido, os estudos e análises apresentados dedicaram-se em buscar alternativas para contornar esses problemas numéricos. Durante o desenvolvimento do capítulo, as análises realizadas, baseadas em simulações com redes do IEEE, culminaram na proposta de dois algoritmos que visam aumentar a robustez do estimador regularizado. Esses algoritmos foram testados com as redes de 14 e 118 barras do IEEE e mostraram resultados satisfatórios, pois apresentaram padrão de respostas similares aos estimadores convencionais. Todavia, essas metodologias não podem ser consideradas consolidadas, já que para alguns casos apresentam elevado número de iterações. Mostrou-se que incluir pseudomedidas de variáveis de estado, da forma como o estimador regularizado propõe, pode ser uma alternativa muito interessante, tendo em vista que essas informações estão prontamente disponíveis durante a operação, principalmente nos casos de perda temporária da observabilidade.

Capítulo 7

Conclusões Gerais

No processo de estimação de estado em sistemas de energia elétrica, quando uma rede se torna não observável, a função de restauração da observabilidade é executada e pseudomedidas de injeção de potência críticas são alocadas para tornar a rede observável. Esse procedimento de restauração da observabilidade é iterativo e o número de iterações é igual ao número de pseudomedidas alocadas. Além disso, a qualidade do estado estimado para as porções não observáveis da rede dependerá da qualidade das pseudomedidas alocadas.

Como uma alternativa à execução do procedimento de restauração da observabilidade em redes em que a observabilidade foi temporariamente perdida, este trabalho apresenta e discute o estimador de estado regularizado. No estimador regularizado, pseudomedidas de estado são alocadas e o parâmetro de regularização, que define a ponderação das pseudomedidas, é especificado de modo que o estado das ilhas observáveis não seja deteriorado pelas pseudomedidas. Como utiliza pseudomedidas de estado, a formulação do estimador regularizado permite o uso de medidas clássicas obtidas via SCADA ou medidas fasoriais obtidas de PMUs.

Durante o desenvolvimento do trabalho foi constatado que é possível obter resultados promissores quando são utilizados métodos matemáticos adequados para a definição do parâmetro de regularização. Todavia, a complexidade e o custo computacional associado a tais métodos os tornam desinteressantes. Nesse contexto, a análise dos testes realizados revelou que o uso de parâmetros de regularização suficientemente pequenos é capaz de limitar a interferência das pseudomedidas no estado estimado, entretanto, em casos específicos, o uso de parâmetros de regularização muito pequenos pode causar problemas numéricos dificultando a convergência do estimador regularizado. Assim, foram propostos mecanismos de seleção das pseudomedidas de estado e foram definidas faixas para a escolha das ponderações a fim de contornar os problemas numéricos do estimador de estado regularizado.

O estimador regularizado foi desenvolvido nas versões completa e desacoplada. Essas

versões foram combinadas em dois algoritmos a fim de melhorar a qualidade do estado estimado e as características de convergência do estimador regularizado. Essas versões foram testadas com as redes do IEEE de 14 e 118 barras.

De modo geral, os resultados apresentados indicam que o estimador de estado regularizado apresenta padrão de respostas similar aos estimadores de estado convencionais. Isto é, ele sempre fornece um estado estimado que é compatível com a qualidade das medidas e pseudomedidas disponíveis. Entretanto, no estimador regularizado deve haver um bom compromisso entre a qualidade e as ponderações das pseudomedidas para que o estado estimado e as características de convergência sejam adequadas.

Apesar de os mecanismos propostos para tratamento das pseudomedidas e suas ponderações terem se mostrados adequados, existem casos que exigem um elevado número de iterações, portanto, novos estudos ainda precisam ser desenvolvidos para que o estimador regularizado seja considerado consolidado. Atualmente, estão sendo desenvolvidas novas versões do estimador regularizado que utilizam as variáveis de estado na forma retangular, além de versões que utilizam outras variáveis de estado. Essas variações do estimador regularizado serão apresentadas em trabalhos futuros.

Referências Bibliográficas

- A. G. Expósito A. Abur. *Power System State Estimation: Theory and Implementation*. CRC Press, 2004.
- A. Abur and A. Gomez-Exposito. Detecting multiple solutions in state estimation in the presence of current magnitude measurements. *Power Systems, IEEE Transactions on*, 12(1):370–375, 1997. ISSN 0885-8950. doi: 10.1109/59.575721.
- A.B. Alves, E.N. Asada, and A. Monticelli. Critical evaluation of direct and iterative methods for solving $ax=b$ systems in power flow calculations and contingency analysis. *Power Systems, IEEE Transactions on*, 14(2):702–708, may 1999.
- R. C. Aster, B. Borchers, and C. H. Thurber. *Parameter Estimation and Inverse Problems*. Elsevier Academic Press, 2005.
- D. Atanackovic and G. Dwernychuk. Bc hydro experiences with utilization of pseudo measurements in state estimation. In *Power and Energy Society General Meeting, 2011 IEEE*, pages 1–5, July 2011. doi: 10.1109/PES.2011.6039283.
- Y. Bard. *Nonlinear Parameter Estimation*. Academic Press, 1974.
- M. S. Bazaraa, H. D. Sherali, and C. M. Shetty. *Nonlinear Programming: Theory and Algorithms*. Jhon Wiley and Sons, Inc, 1993.
- A. Bjorck. *Numerical Methods for Least Squares Problems*. Society for Industrial and Applied Mathematics, 1996.
- M. Boddeti, D. Obadina, F. Garcia, N. D R Sarma, Yang Wu, and V. Kanduri. Ercot’s experiences in using pseudo measurements in state estimation. In *Power and Energy Society General Meeting, 2011 IEEE*, pages 1–5, July 2011. doi: 10.1109/PES.2011.6039374.
- F. Broussolle. State estimation in power systems: Detecting bad data through the sparse inverse matrix method. *Power Apparatus and Systems, IEEE Transactions on*, PAS-97(3): 678–682, May 1978. ISSN 0018-9510. doi: 10.1109/TPAS.1978.354538.
- M. Brown Do Coutto Filho and J.C. Stacchini de Souza. Forecasting-aided state estimation -;part i: Panorama. *Power Systems, IEEE Transactions on*, 24(4):1667–1677, Nov. 2009. ISSN 0885-8950. doi: 10.1109/TPWRS.2009.2030295.
- G. Casella and R. L. Berger. *Statistical Inference, 2nd ed.* Duxbury, 2002.

- A. S. Costa and A. Albuquerque. A two stage orthogonal estimator to incorporate phasor measurements into power system real time modeling. In *17th Power Systems Computation Conference*, 2011.
- M. C. de Almeida. Estimação de estado generalizada trifásica. Tese de doutorado, FEEC, Universidade Estadual de Campinas, Julho 2007.
- M.C. de Almeida, A.V. Garcia, and E.N. Asada. Regularized least squares power system state estimation. *Power Systems, IEEE Transactions on*, 27(1):290–297, Feb. 2012. ISSN 0885-8950. doi: 10.1109/TPWRS.2011.2169434.
- T.E. Dy Liacco. Real-time computer control of power systems. *Proceedings of the IEEE*, 62(7):884 – 891, july 1974. ISSN 0018-9219. doi: 10.1109/PROC.1974.9541.
- I. M. M. A. F. Freitas. A estimação de estado de sistemas eléctricos de energia. Dissertação de doutorado, DEEC, Faculdade de Engenharia da Universidade do Porto, 1995.
- A. R. Gallant. *Nonlinear Statistical Models*. John Wiley, 1987.
- A. Garcia, A. Monticelli, and P. Abreu. Fast decoupled state estimation and bad data processing. *Power Apparatus and Systems, IEEE Transactions on*, PAS-98(5):1645–1652, 1979. ISSN 0018-9510. doi: 10.1109/TPAS.1979.319482.
- P. C. Hansen. The truncated svd as a method for regularization. *BIT*, 1987.
- P. C. Hansen. The discrete picard condition for discrete ill-posed problems. *BIT*, 1990.
- P. C. Hansen. Analysis of discrete ill-posed problems by means of the l-curve. *SIAM Review*, 1992.
- P. C. Hansen. *Rank-Deficient and Discrete Ill-Posed Problems: Numerical Aspects of Linear Inversion*. Society for Industrial and Applied Mathematics, 1998.
- L. Holten, A. Gjelsvik, Sverre Aam, F.F. Wu, and W.-H.E. Liu. Comparison of different methods for state estimation. *Power Systems, IEEE Transactions on*, 1988.
- Jr J. E. Dennis and Robert B. Schnabel. *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*. Society for Industrial and Applied Mathematics, 1996.
- G.R. Krumpholz, K.A. Clements, and P.W. Davis. Power system observability: A practical algorithm using network topology. *Power Apparatus and Systems, IEEE Transactions on*, PAS-99(4):1534 –1542, july 1980. ISSN 0018-9510. doi: 10.1109/TPAS.1980.319578.
- R.E. Larson, W.F. Tinney, L.P. Hajdu, and D.S. Piercy. State estimation in power systems part ii: Implementation and applications. *Power Apparatus and Systems, IEEE Transactions on*, PAS-89(3):353 –363, march 1970a. ISSN 0018-9510. doi: 10.1109/TPAS.1970.292712.

- R.E. Larson, W.F. Tinney, and J. Peschon. State estimation in power systems part i: Theory and feasibility. *Power Apparatus and Systems, IEEE Transactions on*, PAS-89(3):345–352, march 1970b. ISSN 0018-9510. doi: 10.1109/TPAS.1970.292711.
- J. L. Mead. Parameter estimation: A new approach to weighting a priori information. *Journal of Inverse and Ill-Posed Problems*, 2008.
- A. Monticelli and A. Garcia. Reliable bad data processing for real-time state estimation. *Power Apparatus and Systems, IEEE Transactions on*, May 1983.
- A. Monticelli and A. Garcia. Fast decoupled state estimators. *Power Systems, IEEE Transactions on*, 5(2):556–564, 1990. ISSN 0885-8950. doi: 10.1109/59.54566.
- A. Monticelli and F.F. Wu. Network observability: Theory. *Power Apparatus and Systems, IEEE Transactions on*, PAS-104(5):1042–1048, may 1985a. ISSN 0018-9510. doi: 10.1109/TPAS.1985.323454.
- A. Monticelli and F.F. Wu. Network observability: Identification of observable islands and measurement placement. *Power Apparatus and Systems, IEEE Transactions on*, PAS-104(5):1035–1041, may 1985b. ISSN 0018-9510. doi: 10.1109/TPAS.1985.323453.
- A. Monticelli, C.A.F. Murari, and F.F. Wu. A hybrid state estimator: Solving normal equations by orthogonal transformations. *Power Apparatus and Systems, IEEE Transactions on*, PAS-104(12):3460–3468, dec. 1985. ISSN 0018-9510. doi: 10.1109/TPAS.1985.318896.
- A. J. Monticelli. *State Estimation in Electric Power Systems: A Generalized Approach*. Kluwer Academic Publishers, 1999.
- V. A. Morozov. *Methods for solving incorrectly posed problems*. Springer, 1984.
- PSTCA. Power systems test case archive. <http://www.ee.washington.edu/research/pstca/>, 2013. URL {<http://www.ee.washington.edu/research/pstca/>}. Acessado em: 30 de maio de 2013.
- R. Ramlau. Regularization of nonlinear ill-posed operator equations: Methods and applications. Phd thesis, Center of Technomathematics, University of Bremem, January 2004.
- R. J. Santos. Problemas inversos: Métodos iterativos, regularização e validação cruzada generalizada. Tese de doutorado, IMECC, Universidade Estadual de Campinas, Janeiro 1995.
- F.C. Schweppe. Power system static-state estimation, part iii: Implementation. *Power Apparatus and Systems, IEEE Transactions on*, PAS-89(1):130–135, jan. 1970. ISSN 0018-9510. doi: 10.1109/TPAS.1970.292680.
- F.C. Schweppe and D.B. Rom. Power system static-state estimation, part ii: Approximate model. *Power Apparatus and Systems, IEEE Transactions on*, PAS-89(1):125–130, jan. 1970. ISSN 0018-9510. doi: 10.1109/TPAS.1970.292679.

- F.C. Schweppe and J. Wildes. Power system static-state estimation, part i: Exact model. *Power Apparatus and Systems, IEEE Transactions on*, PAS-89(1):120 –125, jan. 1970. ISSN 0018-9510. doi: 10.1109/TPAS.1970.292678.
- A. Simoes-Costa and V.H. Quintana. A robust numerical technique for power system state estimation. *Power Apparatus and Systems, IEEE Transactions on*, PAS-100(2):691 –698, feb. 1981a. ISSN 0018-9510. doi: 10.1109/TPAS.1981.316920.
- A. Simoes-Costa and V.H. Quintana. An orthogonal row processing algorithm for power system sequential state estimation. *Power Apparatus and Systems, IEEE Transactions on*, PAS-100(8):3791 –3800, aug. 1981b. ISSN 0018-9510. doi: 10.1109/TPAS.1981.317022.
- W. Sun and Y. X. Yuan. *Optimization Theory and Methods: Nonlinear Programming*. Springer, 2006.
- A. Tarantola. *Inverse Problem Theory and Methods for Model Parameter Estimation*. Society for Industrial and Applied Mathematics, 2005.
- Jun Zhu and A. Abur. Effect of phasor measurements on the choice of reference bus for state estimation. In *Power Engineering Society General Meeting, 2007. IEEE*, pages 1 –5, june 2007. doi: 10.1109/PES.2007.386175.

Apêndice A

Trabalho Publicado

1. SCHMIDT F. e de ALMEIDA M. C. , “Análise das Características de Convergência de um Estimador de Estado Regularizado para Sistemas de Energia Elétrica”. *XIX Congresso Brasileiro de Automática, (CBA2012)*, Campina Grande, Paraíba , pg. 1-8, 2 a 6 de Setembro de 2012.