

Mario Germán Sandoval

**Identificação de parâmetros em sistemas
mecânicos através da resolução do problema
inverso com inferência estatística Bayesiana**

Dissertação de Mestrado

Dissertação apresentada ao Programa de Pós-graduação em Engenharia Mecânica da PUC-Rio como requisito parcial para obtenção do grau de Mestre em Engenharia Mecânica.

Orientador: Prof. Rubens Sampaio Filho

Rio de Janeiro
Agosto 2013



Mario Germán Sandoval

**Identificação de parâmetros em sistemas
mecânicos através da resolução do problema
inverso com inferência estatística Bayesiana**

Dissertação apresentada ao Programa de Pós-graduação em Engenharia Mecânica da PUC-Rio como requisito parcial para obtenção do grau de Mestre em Engenharia Mecânica. Aprovada pela Comissão Examinadora abaixo assinada.

Prof. Rubens Sampaio Filho

Orientador

Departamento de Engenharia Mecânica – PUC-Rio

Prof. Arthur Martins Barbosa Braga

Departamento de Engenharia Mecânica – PUC-Rio

Prof. Hans Ingo Weber

Departamento de Engenharia Mecânica – PUC-Rio

Prof. Thiago Gamboa Ritto

Universidade Federal do Rio de Janeiro

Prof. José Eugênio Leal

Coordenador do Centro Técnico Científico
Pontifícia Universidade Católica do Rio de Janeiro

Rio de Janeiro, 30 de agosto de 2013

Todos os direitos reservados. É proibida a reprodução total ou parcial do trabalho sem autorização da universidade, do autor e do orientador.

Mario Germán Sandoval

Graduou-se em Física na Universidad Nacional del Sur (Bahia Blanca, Buenos Aires, Argentina), cursando Álgebra, Informática, assim como Física Teórica e Experimental.

Ficha Catalográfica

Sandoval, Mario Germán

Identificação de parâmetros em sistemas mecânicos através da resolução do problema inverso com inferência estatística Bayesiana / Mario Germán Sandoval; orientador: Rubens Sampaio Filho. — Rio de Janeiro : PUC-Rio, Departamento de Engenharia Mecânica, 2013.

v., 74 f: il. ; 29,7 cm

1. Dissertação (mestrado) - Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Engenharia Mecânica.

Inclui referências bibliográficas.

1. Engenharia Mecânica – Dissertação. 2. Problemas inversos;. 3. inferência estatística bayesiana;. 4. mínimos quadrados;. 5. filtro de Kalman.. I. Sampaio Filho, Rubens. II. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Engenharia Mecânica. III. Título.

CDD: 621

Agradecimentos

Ao meu orientador Professor Rubens Sampaio Filho pelo apoio, simpatia de sempre, e incentivo para a realização deste trabalho.

Ao CNPq, CAPES e à PUC-Rio, pelos auxílios concedidos, sem os quais este trabalho não poderia ter sido realizado.

Ao Departamento de Física da Universidad Nacional del Sur (Bahía Blanca, Buenos Aires, Argentina) e sua gente, tanto pela formação quanto pelo apoio.

Aos meus colegas da PUC-Rio, quem me fizeram adorar esse lugar.

Aos professores Hans Weber, Marta Rosales, Thiago G. Ritto, Mario Escalante, Fernando Buezas e Marcelo Piovan que me ofereceram a oportunidade desta cooperação.

Ao pessoal do laboratório, Marcelo, Cesar, Roberta, Americo, Bruno e Leonardo, pela ajuda de todos os dias.

A Wagner pela colaboração, dedicação e compromisso com o laboratório, sem ele, muita coisa não seria possível de fazer.

A meus afetos, em especial aqueles que me esperam na Argentina, com os quais compartilamos o esforço de não estar-os juntos, entendendo o esforço que significa tanto para quem se encontra longe da sua terra quanto para quem espera o voltar.

Resumo

Sandoval, Mario Germán; Sampaio Filho, Rubens. **Identificação de parâmetros em sistemas mecânicos através da resolução do problema inverso com inferência estatística Bayesiana.** Rio de Janeiro, 2013. 74p. Dissertação de Mestrado — Departamento de Engenharia Mecânica, Pontifícia Universidade Católica do Rio de Janeiro.

O problema de estimação pode ser entendido como um caso particular dos problemas inversos. Dadas observações da resposta de um sistema para certas causas, deseja-se estimar certas características do sistema. Essas características, em um sistema dinâmico, geralmente são representadas por parâmetros. Assim, para uma representação físico-matemática do sistema, dada uma excitação e observando a resposta, é possível obter uma estimação dos parâmetros. A estimação paramétrica é de grande importância e utilizada em diversas situações, desde experimentalistas, ao observar fenômenos no laboratório, até quem estuda o comportamento de setores sociais por amostras populacionais. A parte inicial desta dissertação apresenta uma breve introdução ao problema inverso do marco da estatística Bayesiana. Neste marco trata-se a estimação paramétrica como resultado da resolução de um problema inverso. Duas técnicas de estimação são deduzidas a partir da inferência estatística Bayesiana. A primeira delas, mínimos quadrados, coleta todos os dados e logo faz a estimação. A segunda, filtro de Kalman (e filtro de Kalman estendido), melhora o estado do conhecimento dos parâmetros a serem estimados a cada nova observação. Para a abordagem destas técnicas de estimação, de modo de poder compará-las, é apresentada a resolução analítica de um sistema harmônico de um e dois graus de liberdade. Por último, é apresentada uma modelagem de uma bancada experimental, em escala de laboratório, que emula uma coluna de perfuração acoplada a um motor. Esta bancada foi desenvolvida para estudos de dinâmica torcional, na dissertação de mestrado de Bruno C. Cayres A., de modo que aqui só é de interesse a caracterização da mesma. As técnicas de estimação paramétrica são usadas de forma teórica, simulando os dados a partir de soluções analíticas para diferentes parâmetros da modelagem do motor e da coluna. Também usa-se medições feitas na bancada para estimar os parâmetros da modelagem, obtendo assim um conhecimento melhorado dos parâmetros envolvidos no sistema coluna-motor.

Palavras-chave

Problemas inversos; inferência estatística bayesiana; mínimos quadrados; filtro de Kalman.

Abstract

Sandoval, Mario Germán; Sampaio Filho, Rubens (Advisor). **Identification of Mechanical Systems Parameters Through Inverse Problem's Resolution with Bayesian Statistical Inference**. Rio de Janeiro, 2013. 74p. MSc. Dissertation — Departamento de Engenharia Mecânica, Pontifícia Universidade Católica do Rio de Janeiro.

The estimation problem can be understood as a particular case of an inverse problem. Given observations of the response of a system, due to certain causes, one wants to estimate certain characteristics of the problem. These features, in a dynamic system, are usually represented by parameters. Thus, for a mathematical representation of the physical system, given an excitation and given the observing response, it is possible to give an estimation of the parameters. The parameter estimation is of great importance and used in countless situations, such as experimental observation of a phenomena in the laboratory or even by those who study the behaviors social sectors by population samples. The initial part of this dissertation presents a brief introduction to the inverse problem the framework of the Bayesian statistics. In this context, the parametric estimation is a result of the resolution of an inverse problem. Two estimation techniques are derived from the Bayesian statistical inference. The first of these, least squares, collects all the data and then makes the estimation. The second, Kalman filter (and extended filter Kalman), improves the state of knowledge of the parameters to be estimated, with each new observation. To address these estimation techniques, in order to be able to compare them, presents the analytical resolution of a harmonious system of one and two degrees of freedom. Finally, it is presented a model for an experimental setup, in laboratory scale, which emulates a drillstring coupled to a motor. This experimental setup was developed to study the dynamic torsional and by the author of the dissertation of Bruno C. Cayres A., the mode that is of interest here only the characterization of it. These techniques are used for parameter estimation in theoretical way, simulating data from the analytical solutions, for different parameters involved in the column-motor modeling. Also, we use measurements obtained from the experimental setup to estimate the parameters of the column-motor model. Thereby, we obtain an improved knowledge of the parameters involved in the column-motor.

Keywords

Inverse problems; bayesian statistical inference; least squares; Kalman filter.

Sumário

1	Introdução	11
1.1	O que é um problema inverso?	11
1.2	Problemas inversos e estimação paramétrica	12
1.3	Questões de interesse e objetivos do trabalho	13
1.4	Revisão bibliográfica	14
1.5	Organização da dissertação	16
2	Análise de Vibrações no Domínio do Tempo	17
2.1	Resposta impulso	17
2.2	Oscilador amortecido com um grau de liberdade	18
2.3	Oscilador amortecido com dois graus de liberdade	20
3	Estatística Bayesiana para Estimação Paramétrica	24
3.1	A regra de Bayes	24
3.2	Inferência estatística bayesiana	24
3.3	Estimação sujeita a um ruído gaussiano	25
3.4	Estimação do erro nas medidas	27
3.5	Estimação por mínimos quadrados	30
3.6	Estimação num sistema com 1 grau de liberdade	31
4	Estimação Paramétrica Recursiva	38
4.1	Filtro de Kalman sem ruído no modelo	38
4.2	Estimação linear sem ruído no modelo	40
4.3	Filtro de Kalman com ruído no modelo	42
4.4	Estimação linear de um parâmetro constante com ruído no modelo	43
4.5	Estimação linear de um parâmetro que evolui no tempo com ruído no modelo	47
4.6	Filtro de Kalman extendido	49
4.7	Estimação não linear com ruído no modelo	51
4.8	Filtro de Kalman extendido com correção de segunda ordem	52
4.9	Estimação não linear com correção de segunda ordem	52
5	Estimação de parâmetros de uma bancada experimental	55
5.1	Descrição da bancada experimental	55
5.2	Modelagem da bancada experimental	55
5.3	Estimação por mínimos quadrados	59
5.4	Estimação por filtro de Kalman extendido	64
6	Conclusões	70
6.1	Contribuições deste trabalho	70
6.2	Observações deste trabalho	71
6.3	Sugestões para trabalhos futuros	72
6.4	Publicações	72
	Referências Bibliográficas	73

Lista de Figuras

2.1	Sistema massa-mola-amortecedor de um grau de liberdade e excitação de base.	18
2.2	solução Resposta de um sistema massa-mola-amortecedor com 1 grau de liberdade.	20
2.3	Sistema massa-mola-amortecedor de dois graus de liberdade e excitação de base.	20
2.4	Resposta de um sistema massa-mola-amortecedor com 2 graus de liberdade.	23
3.1	Curvas gaussianas das PDFs diretas (esquerda) e dos histogramas dos valores medidos (direita) para $N = 5$.	26
3.2	Curvas gaussianas das PDFs diretas (esquerda) e dos histogramas dos valores medidos (direita) para $N = 100$.	27
3.3	Curvas gaussianas das PDFs posteriores dadas pela Eq.(3.5) para $N = 5$ (esquerda) e $N = 100$ (direita).	27
3.4	Probabilidade marginal $P(x, \sigma \mathbf{y})$ em função de σ e x para $N = 5$.	29
3.5	Probabilidade marginal $P(x, \sigma \mathbf{y})$ em função de σ e x para $N = 100$.	29
3.6	Probabilidades marginais $P(x \mathbf{y})$ (esquerda) e $P(\sigma \mathbf{y})$ (á direita), em função de σ e X , no casos em que $N = 5$.	30
3.7	Probabilidades marginais $P(x \mathbf{y})$ (esquerda) e $P(\sigma \mathbf{y})$ (á direita), em função de σ e X , no casos em que $N = 100$.	30
3.8	Deslocamento e função resposta ao impulso unitário do sistema com 1 grau de liberdade cujas propriedades estão sendo estimadas.	34
3.9	Deslocamentos, forçamento e convergências, para $\sigma_n = 0,002$.	35
3.10	Deslocamentos, forçamento e convergências, para $\sigma_n = 0,006$.	36
3.11	Deslocamentos, forçamento e convergências, para $\sigma_n = 0,006$, $N = 31$ e $N = 61$.	37
4.1	Curvas das soluções.	41
4.2	Curvas dos forçamentos original e reconstruído	41
4.3	Curvas de convergência do modelo, representada pelo valor melhorado do parâmetro a estimar (A_b) para cada nova medição Z_k .	42
4.4	Curvas de convergência.	44
4.5	Curvas de soluções.	45
4.6	Curvas de iterações com $\sigma_n = 0,004$.	46
4.7	Curvas de convergência para um tempo de 90 s e para $N = 451$ medições.	46
4.8	Função densidade de probabilidade priori de A_{b0} .	47
4.9	Função densidade de probabilidade priori de A_{b0} .	48
4.10	Curvas do forçamento, de solução e de convergência.	51
4.11	Curvas do forçamento, de solução e de convergência.	53
5.1	Esquema da bancada experimental, horizontal, como vista de cima.	56
5.2	Curvas reconstruídas e medições para $\sigma_n = 0.01$	60
5.3	Converência da função objetivo ($\varepsilon(\mathbf{x}; \boldsymbol{\eta})$) para $\sigma_n = 0.01$	61

5.4	Heurística para minimizar a função objetivo.	62
5.5	Valores medidos e curvas reconstruídas.	63
5.6	Convergência.	63
5.7	Descentralização do sistema de atrito posicionado em r_1 .	65
5.8	Medição da descentralização do dispositivo de atrito.	65
5.9	Esquemas das modelagens para o torque de atrito τ_1 .	66
5.10	Observações, solução analítica e curvas reconstruídas para o <i>Modelo 1</i> .	67
5.11	Torque e convergência para o <i>Modelo 1</i> .	68
5.12	Observações, solução analítica e curvas reconstruídas para o <i>Modelo 2</i> .	68
5.13	Torque e convergência para o <i>Modelo 2</i> .	69

Lista de Tabelas

3.1	Estimação de parâmetros referente à Figura 3.9.	35
3.2	Estimação de parâmetros referente à Figura 3.10	36
3.3	Estimação de parâmetros referente à Figura 3.11.	37
5.1	Parâmetros da bancada experimental.	57
5.2	Parâmetros estimados com valores iniciais de $c_m = 0,001$ e $k_e = 0,01$	61
5.3	Parâmetros estimados	64
5.4	Estimações por FKE para as modelagens <i>metodo 1</i> e <i>metodo 2</i> .	69

1

Introdução

Neste capítulo, será apresentada uma breve discussão do problema inverso de modo de introduzi-lo, e se dará uma breve introdução à estimação paramétrica no marco do problema inverso, destacando ao conhecido filtro de Kalman como ferramenta de estimação. Será descrito o objetivo do trabalho e será feita uma breve revisão bibliográfica sobre inferência bayesiana e técnicas de estimação de parâmetros.

1.1

O que é um problema inverso?

Segui um exemplo simples de modo de introduzir os conceitos de *problema direto* e *problema inverso*.

Considere um sistema linear de equações algébricas

$$[a] \mathbf{x} = \mathbf{b}, \quad (1.1)$$

onde $[a]$ é uma matriz quadrada, \mathbf{b} e \mathbf{x} são vetores.

Suponha que $[a]$ e \mathbf{x} são conhecidos e se deseja encontrar o valor de \mathbf{b} . Isso pode ser feito de maneira direta, com a ajuda da Eq.(1.1), utilizando somente uma operação de multiplicação de matrizes. Portanto, esse é um exemplo de um *problema direto*.

Por outro lado, suponha agora que as informações conhecidas sejam $[a]$ e \mathbf{b} , e que se deseje encontrar o vetor \mathbf{x} . Nesse caso, se $[a]$ for não singular o problema tem uma única solução dada por

$$\mathbf{x} = [a]^{-1} \mathbf{b}, \quad (1.2)$$

onde $[a]^{-1}$ é a matriz inversa de $[a]$. Nesse caso tem-se um exemplo simples de *problema inverso*.

Um exemplo mais complexo de problema inverso é a determinação da matriz $[a]$ dado que se conhece os vetores \mathbf{b} e \mathbf{x} .

Os problemas inversos podem ser classificados, do ponto de vista matemático, como *bem postos* ou *mal postos*. Se a solução do problema inverso existe, é única e é estável com respeito a qualquer perturbação, então é possível dizer que o problema é bem posto. Caso contrário, ele é dito mal posto.

Nos exemplos apresentados acima, o problema de calcular a inversa de uma matriz não singular é bem posto. Já o problema de determinação da matriz $[a]$ conhecendo os vetores \mathbf{b} e \mathbf{x} é mal posto.

Problemas inversos podem ser encontrados em diversas áreas de interesse prático [4, 5], como por exemplo: (i) a construção de uma imagem a partir da integral de linha da absorção ao longo de linhas através do corpo humano, num exame de tomografia; (ii) ajustar uma curva para a descrição de dados experimentais; (iii) identificação de imagens de satélite; (iv) tratamento de imagens desfocadas; (v) análise de ondas sísmicas; (vi) identificação de parâmetros em sistemas mecânicos; etc.

1.2

Problemas inversos e estimação paramétrica

Ao realizar medições no laboratório, o objetivo é obter informações para caracterizar algum sistema físico ou fenômeno de interesse. Entretanto, em muitos casos as informações necessárias para caracterização do sistema/fenômeno não são aquelas que podem ser efetivamente medidas no laboratório. Nesses casos é necessário buscar relações entre as informações disponíveis e aquelas que realmente são necessárias para completa descrição do sistema ou fenômeno físico. Esse tipo de situação, em que se busca uma causa para um efeito medido, é um problema inverso no contexto de estimação de parâmetros [16].

A estatística bayesiana desempenha um papel importante na resolução de problemas inversos, mais precisamente no campo da estimação de parâmetros. Assim, esta fornece uma teoria de inferência proporcionando uma relação entre as medições no laboratório e as previsões teóricas.

Na estatística bayesiana, a função densidade de probabilidade (PDF) posterior, $p(\mathbf{x}|\mathbf{y})$, fornece o estado do conhecimento sobre um parâmetro ou um vetor de parâmetros, \mathbf{x} a partir de dados coletados, \mathbf{y} .

A partir da inferência bayesiana pode-se deduzir o método de mínimos quadrados. Este usa todos os dados coletados e minimiza a distância entre estes dados e os dados pelo modelo teórico, variando valores de parâmetros desejados para consegui-lo (vease o capítulo 3). Este método é muito usado por experimentalistas para na estimação de parâmetros em problemas que apresentem tanto lineares quanto não lineares respeito aos mesmos. Este método requer a coleção prévia de todos dados para a estimação.

Além disso, a inferência bayesiana também fornece uma teoria que possibilita a dedução de métodos de estimação recursivos, i.e. que atualiza o estado de conhecimento da estimação a cada medição. Considerando linearidade com respeito aos parâmetros, o filtro de Kalman é um destes métodos. No entanto,

existem variações baseadas na aproximação por polinômio de Taylor de primeira e segunda ordem, permitindo estimações não lineares. Nesta dissertação serão abordados os filtros de Kalman estendido e de segunda ordem para problemas que apresentam não linearidade com respeito aos parâmetros.

1.3

Questões de interesse e objetivos do trabalho

Muitas vezes, os sistemas mecânicos podem apresentar propriedades, ou serem submetidos a forçamentos com caracteres aleatórios. Assim como o módulo de elasticidade ao longo do comprimento de uma coluna de perfuração pode ter mudanças, ou diversos atritos ou interações tanto na broca no extremo como no comprimento da coluna, para o qual uma simples modelagem determinística pode não ser suficiente para representar as respostas destes sistemas.

Desta forma, a probabilidade nos problemas inversos serve de motivação para este trabalho. Junto com a estatística bayesiana, a estrutura matemática permite uma forma de avaliar, ou quantificar, o grau de crença, ou confiabilidade, para a reconstrução do problema inverso. Assim, a densidade de probabilidade representa o estado do conhecimento sobre o espaço de possibilidades e pode-se formular uma teoria geral de problemas inversos como uma de inferência estatística. Dentro deste quadro, pode-se considerar desde problemas diretos não lineares até processos de ruído não aditivo, mas o número de problemas que podem ser completamente tratados desta forma pode ser bem pequeno. No entanto, estas ideias são bastante usadas na análise de dados experimentais.

O objetivo deste trabalho é dar uma introdução ao problema inverso do marco da estatística bayesiana. Neste marco Kaipio-2004 trata-se a estimação paramétrica como resultado da resolução do problema inverso, abordando-se duas técnicas de estimação.

- A primeira delas, mínimos quadrados, coleta todos os dados e logo faz a estimação (estimação não recursiva).
- A segunda, filtro de Kalman, melhora o estado do conhecimento dos parâmetros a serem estimados com cada nova observação (estimação recursiva).

Em todos os capítulos ao longo desta dissertação são realizados exemplos e apresentados os resultados obtidos através de rotinas desenvolvidas em MATLAB.

Por último, e com o intuito de aplicar os métodos estudados em um problema real, é apresentada uma análise de três graus de liberdade da dinâmica torsional de uma bancada experimental, desenhada para o trabalho de mestrado de Cayres B. ([2]). Tal análise tem como finalidade a caracterização paramétrica da mesma. De esta forma, a análise paramétrica consta de uma modelagem teórica pela qual são comprovadas de forma teórica as convergências dos métodos implementados, e logo, são mostrados resultados com valores experimentais.

1.4

Revisão bibliográfica

Nesta seção será apresentada uma breve revisão bibliográfica sobre as leituras de maior influência neste trabalho.

1.4.1

Problema inverso e a estatística bayesiana

Os problemas inversos tem sido abordados de diferentes pontos de vista, o que torna essa uma área interdisciplinar que compatibiliza a análise matemática de problemas aos dados experimentais. Assim surgiu uma quantidade imensa de aplicações relevantes desta metodologia à engenharia, medicina, geofísica, astrofísica, etc. Um exemplo conhecido pelos matemáticos é o de identificar a forma de um tambor pelo som que ele emite [3]. O problema direto é identificar o som de um tambor de forma conhecida. A solução ao problema inverso foi apresentada por [6].

Uma classe importante trata da reconstrução de imagens. Esta pode incluir áreas de importância como a de tomografia ou de sensoriamento remoto por imagens entre outras. Nas que, por estimativas do meio pode-se identificar irregularidades nas regiões observadas, seja para peças de maquinarias na indústria ou tecidos biológicos na medicina. Outra classe especial de problemas inversos é constituída pela estimação de certas propriedades que são utilizadas na modelagem matemática de sistemas físicos [5]. Na modelagem física, estas propriedades, em muitos casos, são descritas como parâmetros do modelo matemático que descreve, ou tenta descrever, ao problema real.

A inferência bayesiana proporciona uma ferramenta que liga o resultado das medições com as previsões teóricas. Então, dado que o problema direto diz que para certas causas (parâmetros) por meio do modelo físico tem-se certos efeitos (observações). Na estimação de parâmetros, o problema inverso diz que dado um certo conjunto de observações aos quais se deseja saber os parâmetros, que conformam ao modelo teórico, que as produzem [15].

Assim, dado que se conhece a distribuição das observações condicionada pelos parâmetros, se deseja encontrar a distribuição de probabilidade dos parâmetros condicionada pelas observações. Desta forma, acaba-se dando passo ao processo de caracterização que define a um problema inverso estocástico [1].

1.4.2

Estimação de parâmetros

A estimação paramétrica é o processo pelo qual atribui-se uma descrição paramétrica de um objeto, um processo físico ou um evento (abreviando, serão chamados de objeto físico), baseada em observações obtidas destes. Por exemplo, deseja-se estimar a posição de um barco com três sensores de distância montados em três faróis com posições conhecidas. Assim, considerando um movimento plano, a posição do barco é determinada por dois parâmetros e pode ser deduzida das três observações (triangulação). Porém, a situação tem uma maior complexidade já que as medições podem estar sujeitas a variações ou incertezas. Desta forma, além da estimação dos parâmetros, precisa-se fazer uma avaliação da incerteza envolvida. Para melhorar a precisão da estimação, pode-se aumentar o número de observações obtendo um sistema de equações sobre-determinado [17].

Para a estimação paramétrica, existem técnicas como a modelagem de dados (sem contexto probabilístico) e as enquadradas na teoria de Bayes (no contexto probabilístico). Na estimação bayesiana, o ponto de partida é o experimento probabilístico no qual o vetor de acontecimentos é $\mathbf{x} \in \mathbb{R}^N$ e a sua função densidade de probabilidade (PDF - *probability density function*) é $p(\mathbf{x})$. Estes são chamados de vetor de parâmetros e de PDF priori, respetivamente, quando \mathbf{x} representa propriedades do objeto físico [14].

A tarefa de um estimador de parâmetros é de reconstruir o vetor de parâmetros original \mathbf{x} dadas certas observações do objeto físico ($\mathbf{y} \in \mathbb{R}^{N_y}$). Este é definido como a função $\hat{\mathbf{X}}(\mathbf{y}) : \mathbb{R}^{N_x} \rightarrow \mathbb{R}^{N_y}$. Para um vetor \mathbf{x} fixo, a PDF condicional $p(\mathbf{y}|\mathbf{x})$ fornece a ligação entre o vetor de parâmetros e as medições. Em seguida, relaciona-se a PDF $p(\mathbf{y}|\mathbf{x})$ com a PDF posterior $p(\mathbf{x}|\mathbf{y})$ a partir da lei de Bayes. A PDF posterior representa o que se conhece de \mathbf{x} depois de as medições \mathbf{y} serem obtidas. Este conceito é base para métodos de estimação conhecidos como o método de mínimos quadrados [17] e o método de filtro de Kalman [7].

1.5

Organização da dissertação

Esta dissertação é composta de seis capítulos organizados como a seguir. No capítulo dois é destrita uma análise analítica no domínio do tempo de sistemas harmônicos de um e dois graus de liberdade. Esta é resolvida pela sulução de resposta impulso na integração de convolução com as interações envolvidas.

No capítulo três é definida a inferência da estatística bayesiana na estimação paramétrica. É definida estimação para medidas gaussianas e é deduzido o método de estimação por mínimos quadrados (MQ). Exemplos de um grau são resolvidos usando as soluções do capítulo dois. Nestes exemplos são estimados parâmetros, um linear e outro não linear. As amostras são geradas ao somar um ruído gaussiano independente e identicamente distribuido aos valores da solução analítica do problema.

No capítulo quatro é definido o método de filtro de Kalman (FK), desde a inferência estatísitica bayesiana, como algoritmo de estimação linear recursivo. Para o caso não linear é definido o filtro de Kalman extendido (FKE) e é dada uma solução para a perda de acurase pela aproximação de segundo ordem do polinômio de Taylor utilizada na dedução do método de FKE. Exemplos de um e dois graus de liberdade (resolvidos no capítulo dois) são feitos com o fim de caraterizar/comprender esta técnica.

No capítulo quinto, é descrita a modelagem de uma bancada experimental de dois grausde liberdade que simula uma coluna de perfuração a escala reduzida. Esta modelagem é feita como um sistema coluna-motor acoplados. Esta modelagem é resolvida numericamente e o resultado é utilizado para a geração de amostras. O método de mínimos quadrados é testado para caraterizar parâmetros envolvidos na dinâmica do sistema quando este é livre de atrito no extremo que simula a broca de perfuração. Uma vez caraterizado o método de estimação, medições experimentáis são usadas para estimar nove parâmetros envolvidos nesta dinâmica. Logo depois é usado o método de filtro de Kalman extendido com correção de segundo ordem para compreender os fenômenos envolvidos no dispositivo desenhado para fornecer o atrito.

No capítulo sexto são apresentadas conclusões e são discutidos trabalhos futuros.

2

Análise de Vibrações no Domínio do Tempo

Este capítulo faz uma revisão sobre a análise de sistemas mecânicos, com 1 ou 2 graus de liberdade, no domínio do tempo. Sistemas como esses serão investigados mais a frente neste trabalho, o que justifica essa revisão.

2.1

Resposta impulso

Considere um sistema linear invariante no tempo tendo uma excitação escalar $f(t)$ e resposta $x(t)$. Seja a função $h_x(t)$ a resposta do sistema a uma excitação do tipo $f(t) = \delta(t)$, onde $\delta(\cdot)$ é a função Delta de Dirac.

A função $h_x(t)$ é uma função adequada para caracterizar a resposta do sistema para qualquer excitação $f(t)$. Para demonstrar isto, pode-se escrever $f(t)$ como a superposição de funções delta de Dirac, e então, usar a integração de convolução

$$f(t) = \int_{-\infty}^{+\infty} f(s)\delta(t-s)ds \equiv \int_{-\infty}^{+\infty} f(t-r)\delta(r)dr, \quad (2.1)$$

onde $r = t - s$.

Segundo a definição de $h_x(t)$ observa-se o que a excitação dada por $\delta(t-s)$ induz a resposta $h_x(t-s)$, ou, o que é a mesma coisa, $\delta(r)$ induz $h_x(r)$. Multiplicando esses pulsos de excitação de função delta pela suas amplitudes, dadas na Eq.(2.1), e pela superposição das respostas, obtém-se

$$x(t) = \int_{-\infty}^{+\infty} f(s)h_x(t-s)ds \equiv \int_{-\infty}^{+\infty} f(t-r)h_x(r)dr, \quad (2.2)$$

onde a última integral é conhecida como integral de convolução de Duhamel para sistemas lineares, e a função $h_x(t)$ é chamada função de resposta impulso para a resposta x do sistema.

É fácil ver que para a excitação estática $f(t) = f_0$ (constante para todo tempo t), tem-se a resposta de estado estacionário $x(t) = f_0 h_{x,st}$

$$h_{x,st} \equiv \int_{-\infty}^{+\infty} h_x(r)dr, \quad (2.3)$$

e o sistema tem resposta x de estado estacionário estática se a resposta da excitação estática f_0 é finita.

A determinação da função resposta impulso envolve a procura das condições iniciais do problema. Ao se considerar o sistema diferencial de equações

$$\sum_{j=0}^n a_j \frac{d^j x(t)}{dt^j} = f(t), \quad (2.4)$$

no caso em que $f(t) = \delta(t)$, tem-se

$$\left[\frac{d^{n-1} h_x(t)}{dt^{n-1}} \right]_{t=0^+} = a_n^{-1}, \quad (2.5)$$

e

$$\left[\frac{d^j h_x(t)}{dt^j} \right]_{t=0^+} = 0 \quad \text{com } j \leq n-2. \quad (2.6)$$

como sendo as condições iniciais do problema.

2.2

Oscilador amortecido com um grau de liberdade

Suponha-se um sistema massa-mola-amortecedor de um grau de liberdade segundo é apresentado na Figura 2.1.

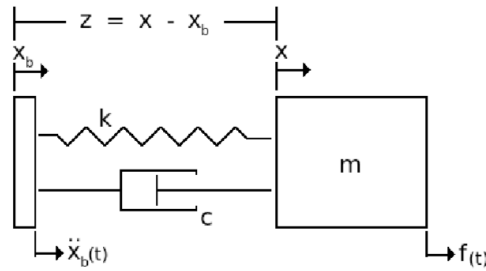


Figura 2.1: Sistema massa-mola-amortecedor de um grau de liberdade e excitação de base.

Nesse sistema k é constante da mola, c a constante de amortecimento, m a massa, x a posição da massa, x_b a posição da base, $f(t)$ o forçamento na mola e $\ddot{x}_b(t)$ a aceleração do movimento da base.

Sejam a energia cinética

$$E_c = \frac{1}{2} m \dot{x}^2, \quad (2.7)$$

a energia potencial elástica

$$E_e = \frac{1}{2} k (x - x_b)^2, \quad (2.8)$$

e a energia dissipada pelo amortecedor

$$E_p = \frac{1}{2} c (\dot{x} - \dot{x}_b)^2, \quad (2.9)$$

Assim, o lagrangeano é

$$\mathcal{L}(x, \dot{x}) = \frac{1}{2}m\dot{x}^2 - \frac{1}{2}k(x - x_b)^2, \quad (2.10)$$

e pela equação Euler-Lagrange

$$\frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{x}} - \frac{\partial \mathcal{L}}{\partial x} + \frac{\partial E_p}{\partial x} = f(t), \quad (2.11)$$

obtêm-se o equacionamento do sistema dado na Figura 2.1 o seguinte

$$m\ddot{x}(t) + c(\dot{x}(t) - \dot{x}_b(t)) + k(x(t) - x_b(t)) = f(t). \quad (2.12)$$

Fazendo a mudança de variáveis $z = x(t) - x_b(t)$ tem-se

$$\begin{aligned} m\ddot{x}(t) + c(\dot{x}(t) - \dot{x}_b(t)) + k(x(t) - x_b(t)) &= f(t), \\ m(\ddot{z}(t) + \ddot{x}_b(t)) + c\dot{x}_b(t) + kz(t) &= f(t), \\ m\ddot{z}(t) + c\dot{z}_b(t) + kz(t) &= f(t) - m\ddot{x}_b(t). \end{aligned} \quad (2.13)$$

Multiplicando por m^{-1} a direita em ambos os lados da Eq.(2.13), e fazendo $\omega_0^2 = \frac{k}{m}$ e $\zeta = \frac{c}{2\sqrt{mk}}$ obtêm-se

$$\ddot{z}(t) + 2\zeta\omega_0\dot{z}_b(t) + \omega_0^2 z(t) = \frac{f(t)}{m} - \ddot{x}_b(t) = (m^{-1})q_z(t), \quad (2.14)$$

onde $q_z(t) = f(t) - m\ddot{x}_b(t)$ é a força no sistema da variável z , relativa à base, $\omega_0 = \sqrt{k/m}$ é a frequência circular natural não amortecida, e $\zeta = \frac{c}{2\sqrt{km}}$ é a fração de amortecimento crítica.

Para achar a resposta impulso $h_x(t)$ faz-se $q_z(t) = \delta(t)$ na Eq.(3.26) e tem-se:

$$\ddot{h}_x(t) + 2\zeta\omega_0\dot{h}_x(t) + \omega_0^2 h_x(t) = 0 \quad \text{para } t > 0. \quad (2.15)$$

Das condições iniciais (2.5) e (2.6), tem-se:

$$\left[\frac{dh_x(t)}{dt} \right]_{t=0^+} = m^{-1}, \quad (2.16)$$

e

$$[h_x(t)]_{t=0^+} = 0. \quad (2.17)$$

Considerando que $|\zeta| < 1$, a solução homogênea da Eq.(2.15) é:

$$h(t) = e^{-\zeta\omega_0 t} [A \sin(\omega_a t) + B \cos(\omega_a t)], \quad (2.18)$$

onde $\omega_a = \omega_0 \sqrt{1 - \zeta^2}$ é a frequência circular natural amortecida.

Das condições (2.17) e (2.16) tem-se $A = 0$, e $B = \frac{1}{m\omega_a}$ respectivamente. Então a função resposta impulso é:

$$h_x(t) = \frac{e^{-\zeta\omega_0 t}}{m\omega_a} \sin(\omega_a t) U(t), \quad (2.19)$$

onde $U(t)$ é a função degrau unitário.

Assim, a resposta para o sistema (3.26), segundo a Eq.(2.2), é:

$$z(t) = \int_{-\infty}^t q_z(s) h_z(t-s) ds, \quad (2.20)$$

sendo $h_z(t)$ é a resposta impulso:

$$h_z(t) = \frac{e^{-\zeta\omega_0 t}}{m\omega_a} \sin(\omega_a t) U(t). \quad (2.21)$$

A Figura 2.2 mostra a solução determinística do sistema massa-mola-amortecedor (2.1).

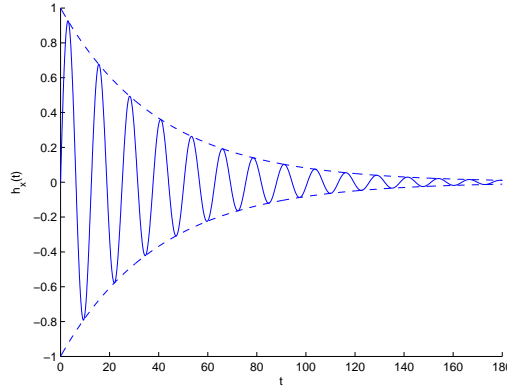


Figura 2.2: solução Resposta de um sistema massa-mola-amortecedor com 1 grau de liberdade.

Para achar a resposta de estado estacionário tem-se

$$h_{x,st} = \int_{-\infty}^{+\infty} \frac{e^{-\zeta\omega_0 r}}{m\omega_a} \sin(\omega_a r) U(r) dr = \int_0^{\infty} \frac{e^{-\zeta\omega_0 r}}{m\omega_a} \sin(\omega_a r) dr \quad (2.22)$$

de onde tem-se $h_{x,st} = k^{-1}$.

2.3

Oscilador amortecido com dois graus de liberdade

Considere agora um sistema massa-mola-amortecedor de dois graus de liberdade, segundo é apresentado na Figura 2.3.

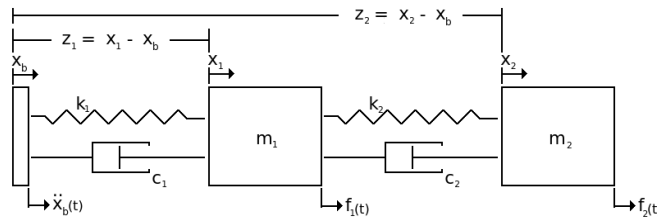


Figura 2.3: Sistema massa-mola-amortecedor de dois graus de liberdade e excitação de base.

Nesse sistema k_1 e k_2 são as constantes das molas, c_1 e c_2 as constantes de amortecimento, m_1 e m_2 as massas, x_1 e x_2 as posições das massa m_1 e m_2 respectivamente, x_b a posição da base, $f_1(t)$ e $f_2(t)$ os forçamentos nas massas e $\ddot{x}_b(t)$ a aceleração do movimento da base.

Sejam a energia cinética

$$E_c = \frac{1}{2}m_1\dot{x}_1^2 + \frac{1}{2}m_2\dot{x}_2^2, \quad (2.23)$$

a energia potencial elástica

$$E_e = \frac{1}{2}k_1(x_1 - x_b)^2 + \frac{1}{2}k_2(x_2 - x_1)^2, \quad (2.24)$$

e a energia dissipada pelo amortecedor

$$E_p = \frac{1}{2}c_1(\dot{x}_1 - \dot{x}_b)^2 + \frac{1}{2}c_2(\dot{x}_2 - \dot{x}_1)^2. \quad (2.25)$$

Assim, o lagrangeano é

$$\mathcal{L}(x, \dot{x}) = \frac{1}{2}m_1\dot{x}_1^2 + \frac{1}{2}m_2\dot{x}_2^2 - \frac{1}{2}k_1(x_1 - x_b)^2 - \frac{1}{2}k_2(x_2 - x_1)^2, \quad (2.26)$$

e pela equação Euler-Lagrange para cada coordenada x_i (com $i = 1, 2$)

$$\frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{x}_i} - \frac{\partial \mathcal{L}}{\partial x_i} + \frac{\partial E_p}{\partial x_i} = f_i(t), \quad \text{com } i = 1, 2, \quad (2.27)$$

obtem-se o equacionamento do sistema

$$m_1\ddot{x}_1 + c_1(\dot{x}_1 - \dot{x}_b) - c_2(\dot{x}_2 - \dot{x}_1) + k_1(x_1 - x_b) - k_2(x_2 - x_1) = f_1(t), \quad (2.28)$$

$$m_2\ddot{x}_2 + c_2(\dot{x}_2 - \dot{x}_1) + k_2(x_2 - x_1) = f_2. \quad (2.29)$$

Fazendo a troca de variáveis $z_1 = x_1 - x_b$ e $z_2 = x_2 - x_b$ para obter as equações relativas à base b como a seguir

$$m_1\ddot{z}_1 + (c_1 + c_2)\dot{z}_1 - c_2\dot{z}_2 + (k_1 + k_2)z_1 - k_2z_2 = f_1(t) - m_1\ddot{x}_b, \quad (2.30)$$

$$m_2\ddot{z}_2 - c_1\dot{z}_1 + c_2\dot{z}_2 - k_1z_1 + k_2z_2 = f_2(t) - m_2\ddot{x}_b. \quad (2.31)$$

De forma que é conveniente, escrever as Eqs.(2.30) e (2.31) em forma matricial como a seguir

$$\begin{pmatrix} m_1 & 0 \\ 0 & m_2 \end{pmatrix} \begin{pmatrix} \ddot{z}_1 \\ \ddot{z}_2 \end{pmatrix} + \begin{pmatrix} c_1 + c_2 & -c_2 \\ -c_2 & c_2 \end{pmatrix} \begin{pmatrix} \dot{z}_1 \\ \dot{z}_2 \end{pmatrix} + \begin{pmatrix} k_1 + k_2 & -k_2 \\ -k_2 & k_2 \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \begin{pmatrix} f_1(t) \\ f_2(t) \end{pmatrix} - \begin{pmatrix} m_1 \\ m_2 \end{pmatrix} \ddot{x}_b,$$

$$[m] \ddot{\mathbf{z}} + [c] \dot{\mathbf{z}} + [k] \mathbf{z} = \mathbf{f}(t) - \mathbf{m}x_b. \quad (2.32)$$

Seja o problema de autovalores e autovetores

$$([m]^{-1} [k] - \lambda_i [I]) \boldsymbol{\nu}_i, \quad (2.33)$$

onde λ_i são os autovalores e $\boldsymbol{\nu}_i$ seus correspondentes autovetores.

Sejam $[\lambda]$ e $[\nu]$ as matrizes de autovalores e de autovetores correspondentes ao problema (2.33). Fazendo a mudança de variáveis dada por

$$\mathbf{z} = [\nu] \boldsymbol{\xi}, \quad (2.34)$$

e multiplicando a esquerda ao sistema (2.32), tem-se

$$[\nu]^T [m] [\nu] \ddot{\boldsymbol{\xi}} + [\nu]^T [c] [\nu] \dot{\boldsymbol{\xi}} + [\nu]^T [k] [\nu] \boldsymbol{\xi} = [\nu]^T (\mathbf{f}(t) - \mathbf{m}x_b), \quad (2.35)$$

$$[\tilde{m}] \ddot{\boldsymbol{\xi}} + [\tilde{c}] \dot{\boldsymbol{\xi}} + [\tilde{k}] \boldsymbol{\xi} = [\nu]^T (\mathbf{f}(t) - \mathbf{m}x_b), \quad (2.36)$$

onde $[\tilde{m}]$, $[\tilde{c}]$ e $[\tilde{k}]$ são matrizes diagonais, sendo chamadas de matrizes generalizadas de massa, amortecimento e rigidez respectivamente.

Pode-se demonstrar facilmente, [8], que $[\tilde{k}] = [\tilde{m}] [\nu]$. Assim, multiplicando a Eq.(2.36) a esquerda por $[\tilde{m}]^{-1}$, obtém-se

$$\ddot{\boldsymbol{\xi}} + [\tilde{m}]^{-1} [\tilde{c}] \dot{\boldsymbol{\xi}} + [\nu] \boldsymbol{\xi} = [\tilde{m}]^{-1} [\nu]^T (\mathbf{f}(t) - \mathbf{m}x_b), \quad (2.37)$$

e escrevendo a Eq.(2.37) em termos da suas componentes

$$\ddot{\xi}_i + 2\zeta_i \omega_i \dot{\xi}_i + \omega_i^2 \xi_i = \frac{1}{\tilde{m}_{i,i}} \sum_{l=1}^2 (\nu_{l,i} (f_l(t) - m_l x_b)), \quad (2.38)$$

onde $\omega_i^2 = \nu_{i,i}$ e $2\zeta_i \omega_i = [\tilde{m}]^{-1} [\tilde{c}]$.

Segundo a Eq.(2.1) e fazendo o mesmo que no exemplo (2.3), obtém-se as respostas impulso como a seguir

$$\tilde{h}_i(t) = \frac{U(t)}{\tilde{m}_{i,i} \omega_{a_i}} \sin(\omega_{a_i} t) e^{-\zeta_i \omega_i t}. \quad (2.39)$$

Definindo a matriz $[\tilde{h}]$ como

$$[\tilde{h}] = \begin{pmatrix} \tilde{h}_1(t) & 0 \\ 0 & \tilde{h}_2(t) \end{pmatrix}, \quad (2.40)$$

e segundo a Eq.(2.2) em forma matricial

$$\boldsymbol{\xi}(t) = \int_{-\infty}^{+\infty} [\tilde{h}](t-s) [\nu]^T (\mathbf{f}(s) - \mathbf{m}\ddot{x}_b) ds, \quad (2.41)$$

pela Eq.(2.34) e pelo fato de que $[h] = [\nu] [\tilde{h}] [\nu]^T$, a solução ao sistema dado na Figura 2.3 é

$$\begin{aligned}
\mathbf{z}(t) &= [\nu] \boldsymbol{\xi}(t) \\
&= [\nu] \int_{-\infty}^t [\tilde{h}](t-s) [\nu]^T (\mathbf{f}(s) - \mathbf{m}\ddot{\mathbf{x}}_b) ds \\
&= \int_{-\infty}^t [h](t-s)(\mathbf{f}(s) - \mathbf{m}\ddot{\mathbf{x}}_b) ds
\end{aligned} \tag{2.42}$$

Supondo que não tem os forçamentos nas massas ($\{f(t)\} = 0$), então só tem-se o movimento de base dado pela equação:

$$q_z(s) = -m\ddot{x}_b(t) = -mA_b \sin(\Omega_b t). \tag{2.43}$$

Sejam $m_1 = \frac{3}{2}$, $m_2 = 100 \text{ kg}$, $k_1 = \frac{3}{2}$, $k_2 = 10000 \text{ N/m}$, $c_1 = \frac{3}{2}$, $c_2 = 10000 \text{ N}\cdot\text{s/m}$, e com o forçamento dado pelos valores $A_b = 1 \text{ m/s}^2$ e $\Omega_b = 0,5 \text{ rad/s}$. A solução determinística é mostrada na Figura 2.4.

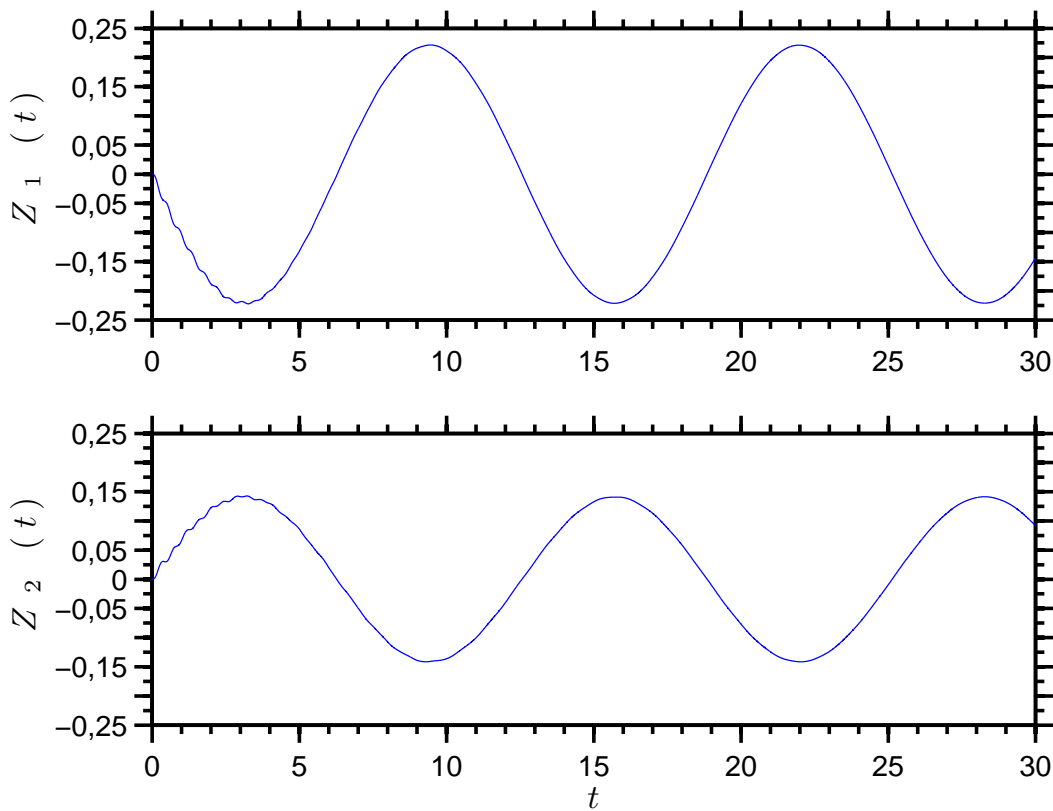


Figura 2.4: Resposta de um sistema massa-mola-amortecedor com 2 graus de liberdade.

3

Estatística Bayesiana para Estimação Paramétrica

Neste capítulo serão apresentada brevemente a estimação de parâmetros dada no marco da teoria bayesiana. Em este marco, será tratado a estimação por meio da função gaussiana e será definido o método de quadrados mínimos.

Por último, será resolvido um exemplo simples com a metodologia de mínimos quadrados.

3.1

A regra de Bayes

Na teoria da probabilidade, o teorema de Bayes é de grande interesse. Enunciado por Thomas Bayes em 1767, ele expressa a probabilidade condicional de um evento aleatório A , sabendo que um evento B ocorreu, em termos da probabilidade condicional de B , dado que A ocorreu, e das probabilidade marginais de A e B [13].

Teorema 3.1 (Teorema de Bayes) *Sejam A e B dois eventos aleatórios. Então tem-se que*

$$p(A|B) = \frac{p(B|A)p(A)}{p(B)},$$

onde $p(A|B)$ denota a probabilidade de A dado B , $p(B|A)$ representa a probabilidade de B dado A , $p(A)$ é a probabilidade marginal de A e $p(B)$ é a probabilidade marginal de B .

3.2

Inferência estatística bayesiana

No processo de estimação paramétrica deseja-se caracterizar os parâmetros de um sistema físico, representados aqui pelo vetor \mathbf{x} , através de um conjunto de observações (experimentos ou simulações), representadas pelo vetor \mathbf{y} .

Devido às imprecisões da medição, interferência de agentes externos, entre outras causas, as observações estão sujeitas a variabilidades (ruído), e por isso são modeladas como um vetor de valores aleatórios \mathbf{Y} . Por conseguinte, e sobre o ponto de vista bayesiano, os parâmetros estimados passam a ser representados pelo vetor aleatório \mathbf{X} , [14].

Esse processo de caracterização define um problema inverso estocástico, no qual se deseja encontrar a distribuição de probabilidade dos parâmetros condicionada pelas observações $p(\mathbf{x}|\mathbf{y})$ (PDF posterior), dado que se conhece a distribuição das observações condicionada pelos parâmetros $p(\mathbf{y}|\mathbf{x})$ (PDF direta) [1].

Nesse contexto a teoria de inferência estatística bayesiana se mostra de grande valia, pois permite relacionar os resultados das observações com as previsões, pois o teorema de Bayes diz que

$$p(\mathbf{x}|\mathbf{y}) = \frac{1}{p(\mathbf{y})} p(\mathbf{y}|\mathbf{x}) p(\mathbf{x}). \quad (3.1)$$

Esta equação diz que a PDF posterior $p(\mathbf{x}|\mathbf{y})$ pode ser calculada utilizando a PDF direta $p(\mathbf{y}|\mathbf{x})$, que é determinada por experimento ou simulação, a PDF a priori $p(\mathbf{x})$, que é calculada com base no que se conhece (ou se acredita conhecer) do parâmetro \mathbf{X} antes das observações, e da constante de normalização $1/p(\mathbf{y})$. A PDF direta $p(\mathbf{y}|\mathbf{x})$ também é chamada de função de verossimilhança.

3.3

Estimação sujeita a um ruído gaussiano

Para exemplificar a metodologia, suponha que se deseja estimar o comprimento de um canudo. Para isso, são realizadas medições da longitude x do canudo, sendo armazenadas no vetor \mathbf{y} . Essas medições estão sujeitas a erros, os quais são modelados como sendo um ruído que tem distribuição de probabilidade gaussiana, com média x e desvio padrão σ .

Com a finalidade de obter uma formulação matemática, supõe-se conhecido o parâmetro x . Então, para uma medição Y_i , dado que se conhece x , tem-se

$$p(y_i|x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(y_i-x)^2}{2\sigma^2}}. \quad (3.2)$$

Se são realizadas N medições independentes e igualmente distribuídas, agrupadas em $\mathbf{Y} = (Y_1, \dots, Y_N)^T$, a PDF direta de todas elas é dada por

$$p(\mathbf{y}|x) = \frac{1}{(2\pi\sigma^2)^{\frac{N}{2}}} e^{-\sum_{i=1}^N \frac{(y_i-x)^2}{2\sigma^2}}, \quad (3.3)$$

e a PDF posterior, segundo a Eq.(3.1), é

$$p(x|\mathbf{y}) = \frac{\aleph}{(2\pi\sigma^2)^{\frac{N}{2}}} e^{-\sum_{i=1}^N \frac{(y_i-x)^2}{2\sigma^2}} p(x), \quad (3.4)$$

onde $\aleph = 1/p(\mathbf{y})$ é a constante de normalização.

A equação acima pode ser desenvolvida de forma que

$$p(x|\mathbf{y}) = \frac{N}{(2\pi\sigma^2)^{\frac{N}{2}}} e^{-\left(\frac{N}{2\sigma^2}\right)\left(x - \frac{1}{N} \sum_{i=1}^N y_i\right)^2} p(x), \quad (3.5)$$

donde se conclui que a média gaussiana é

$$\bar{y} = \sum_{i=1}^N \frac{y_i}{N}, \quad (3.6)$$

e o desvio padrão gaussiano é

$$\bar{\sigma} = \frac{\sigma}{\sqrt{N}}. \quad (3.7)$$

A seguir são realizados dois experimentos numéricos para estimar o comprimento do canudo. Ambos consideram $x = 150 \text{ mm}$, mas o primeiro experimento usa $N = 5$ medições para estimar x , enquanto que o segundo usa $N = 100$. As medições foram geradas somando-se a x um ruído gaussiano de média zero e $\sigma = 0,05$. Esse ruído é gerado com a função `randn` do MATLAB.

Utilizando as Eq.(3.6) são calculadas as médias gaussianas, onde obtêm-se os valores de $\bar{y}_5 = 0,1296$ e de $\bar{y}_{100} = 0,1519$ respectivamente. Da mesma forma, com ajuda da Eq.(3.7), obtêm-se os desvios padrões gaussianos, cujos valores são $\bar{\sigma}_5 = 0,0224$ e $\bar{\sigma}_{100} = 0,005$.

As PDFs diretas e os histogramas das amostras geradas em MATLAB podem ser vistas nas Figuras 3.1 e 3.2 para $N = 5$ e $N = 100$ respectivamente.

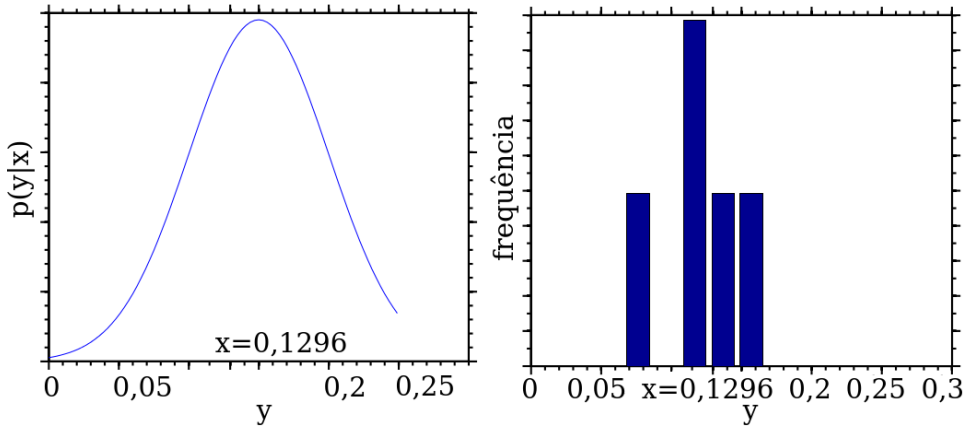


Figura 3.1: Curvas gaussianas das PDFs diretas (esquerda) e dos histogramas dos valores medidos (direita) para $N = 5$.

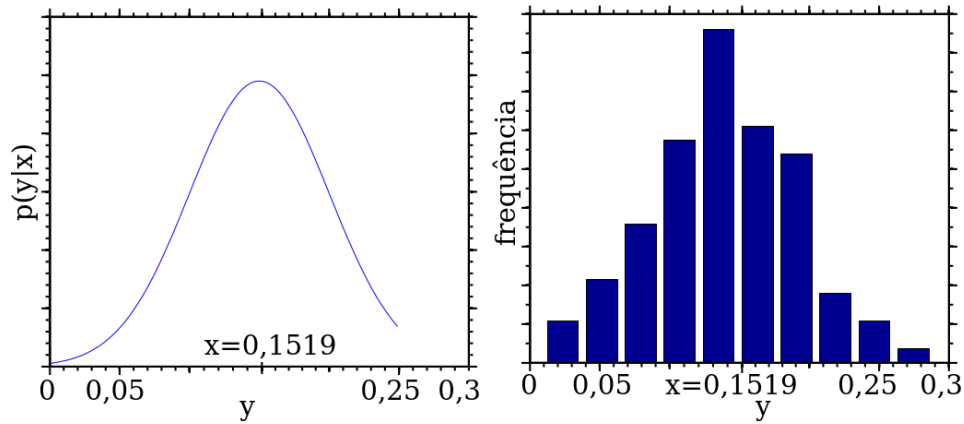


Figura 3.2: Curvas gaussianas das PDFs diretas (esquerda) e dos histogramas dos valores medidos (direita) para $N = 100$.

Na Figura 3.3 apresenta-se as curvas das PDFs posteriores para os dois casos estudados. Tais curvas são dadas pela Eq.(3.5), onde a PDF a priori é considerada uniforme. Assim esta última pode ser aproximada como sendo uma constante e ser absorvida pela constante de normalização. As situações em que $p(x)$ não é considerada como uniforme e a importância como tal, serão tratadas mais na frente neste trabalho. Voltando ao gráfico, é possível observar que os picos em ambas curvas ocorrem ao redor dos valores 0,1296 e 0,1519 respectivamente, que são valores bem próximos de \bar{y}_5 e \bar{y}_{100} .

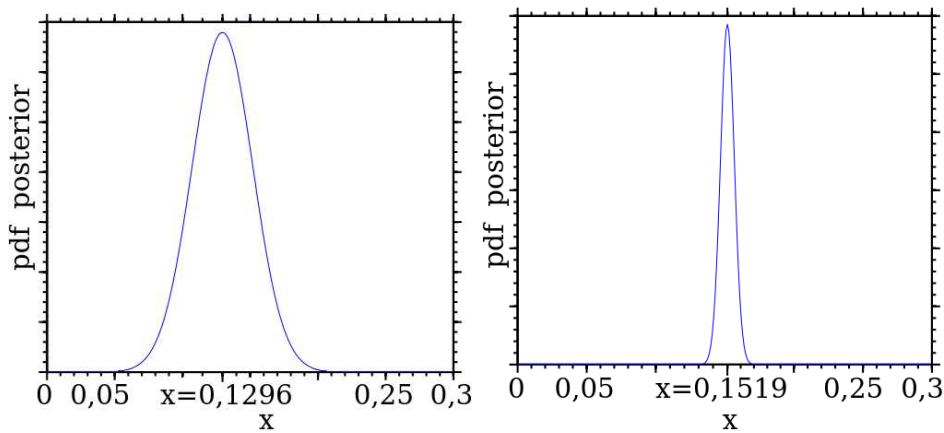


Figura 3.3: Curvas gaussianas das PDFs posteriores dadas pela Eq.(3.5) para $N = 5$ (esquerda) e $N = 100$ (direita).

3.4

Estimação do erro nas medidas

Ainda no problema do canudo, pode-se supor que o valor do desvio padrão σ não é conhecido e usar a técnica bayesiana para estima-lo. Então, seguindo

um raciocínio similar ao apresentado anteriormente, pode-se obter a formula de Bayes da forma

$$p(x, \sigma | \mathbf{y}) = \frac{\aleph}{(2\pi\sigma^2)^{\frac{N}{2}}} e^{-\frac{N}{2\sigma^2}[(x-\bar{y})^2 + S^2]} p(x, \sigma), \quad (3.8)$$

onde $S^2 = \sum_{i=1}^N \frac{(y_i^2 - \bar{y}^2)}{N}$. Nessa equação não se conhecem os valores de X e σ .

Se a PDF priori $p(x, \sigma)$ é uniforme, ela poder ser aproximada por uma constante, e levando-se em conta a integral

$$\int_0^\infty \frac{1}{\sigma^k} e^{-\frac{A}{\sigma^2}} d\sigma = \frac{\Gamma(\frac{k-1}{2})}{2\sqrt{A^{k-1}}}, \quad (3.9)$$

a normalização da PDF posterior pode ser encontrada de forma que

$$\int_0^\infty \int_{-\infty}^\infty p(x, \sigma | \mathbf{y}) dx d\sigma = 1. \quad (3.10)$$

Assim, obtem-se a seguinte PDF posterior conjunta

$$p(x, \sigma | \mathbf{y}) = \sqrt{\frac{8}{N\pi}} \left(\frac{NS^2}{2} \right)^{\frac{N}{2}} \frac{1}{S^2 \Gamma(\frac{1}{2}N-1) \sigma^N} e^{-\frac{N}{2\sigma^2}[(x-\bar{Y})^2 + S^2]}, \quad (3.11)$$

onde $\Gamma(n+1) = (n)!$ é a função gamma [11], e o símbolo $!$ é o operador fatorial.

Ao se integrar a PDF posterior conjunta dada pela Eq.(3.11), obtêm-se as PDF marginais dadas por

$$p(x | \mathbf{y}) = \int_0^\infty p(x, \sigma | \mathbf{y}) d\sigma = \frac{\Gamma(\frac{1}{2}N-\frac{1}{2})}{\Gamma(\frac{1}{2}N-1)} \frac{S^{N-2}}{\pi^{\frac{1}{2}} [(x-\bar{y})^2 + S^2]^{\frac{N-1}{2}}}, \quad (3.12)$$

e

$$p(\sigma | \mathbf{y}) = \int_{-\infty}^\infty p(x, \sigma | \mathbf{Y}) dx = \frac{2}{\Gamma(\frac{1}{2}N-1)} \left(\frac{NS^2}{2} \right)^{\frac{N}{2}-1} \frac{e^{-\frac{N}{2\sigma^2}S^2}}{\sigma^{N-1}}. \quad (3.13)$$

Agora considere os dois experimentos numéricos realizados anteriormente com os mesmos velores, mas com a diferencia de não conhecer σ . Usando as Eqs.(3.11), (3.12) e (3.13) obtêm-se as superfícies apresentadas nas Figuras 3.4 e 3.5, que representam as PDFs posteriores para $N = 5$ e $N = 100$ respectivamente.

Pode-se observar que os picos ocorrem quando X e σ estão perto de $\bar{y}_5 = 0,1296$ e $S_5 = 0,0294$ para $N = 5$, e perto de $\bar{y}_{100} = 0,1519$ e $S_{100} = 0,0492$ no caso em que $N = 100$.

Para uma melhor observação, em ambos os casos constroem-se as curvas das probabilidades marginais dadas pelas Eqs.(3.12) e (3.13), como pode ser visto nas Figuras 3.6 e 3.7.

Nestas figuras, em ambos casos ($N = 5$ e $N = 100$), pode-se observar que os máximos de cada uma ocorrem em $X = \bar{Y}$ para $P(x|\mathbf{y})$, mas para $P(\sigma|\mathbf{y})$ não ocorrem exatamente em $\sigma = S$ e sim um pouco mais deslocados para a frente, e, também, são ligeiramente assimétricos. Isso é possível de se observar ao calcular as médias das probabilidades marginais

$$E[x|\mathbf{y}] = \bar{y}, \quad (3.14)$$

$$E[\sigma|\mathbf{y}] = \sqrt{\frac{N}{2}} \frac{\Gamma(\frac{1}{2}N - \frac{3}{2})}{\Gamma(\frac{1}{2}N - 1)} S \sim \left(1 + \frac{7}{4N} + \frac{145}{32N^2} + \dots\right) S, \quad (3.15)$$

onde pode-se observar que para valores pequenos de N o valor esperado de σ é maior que S .

Calculando-se as covariâncias da probabilidade posterior e das probabilidades marginais tem-se

$$E[(x - E[x|\mathbf{y}])^2|\mathbf{y}] = E[x^2|\mathbf{y}] - (E[x|\mathbf{y}])^2 = \frac{S^2}{N - 4}, \quad (3.16)$$

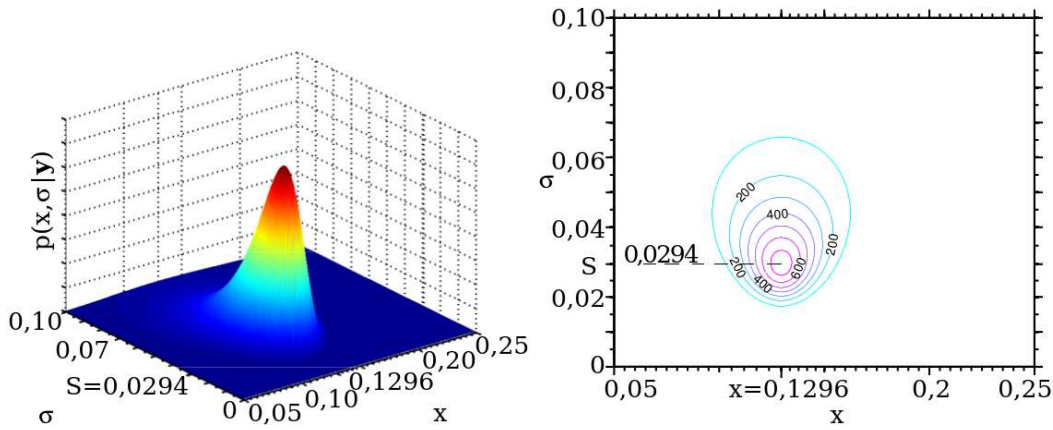


Figura 3.4: Probabilidade marginal $P(x, \sigma|\mathbf{y})$ em função de σ e x para $N = 5$.

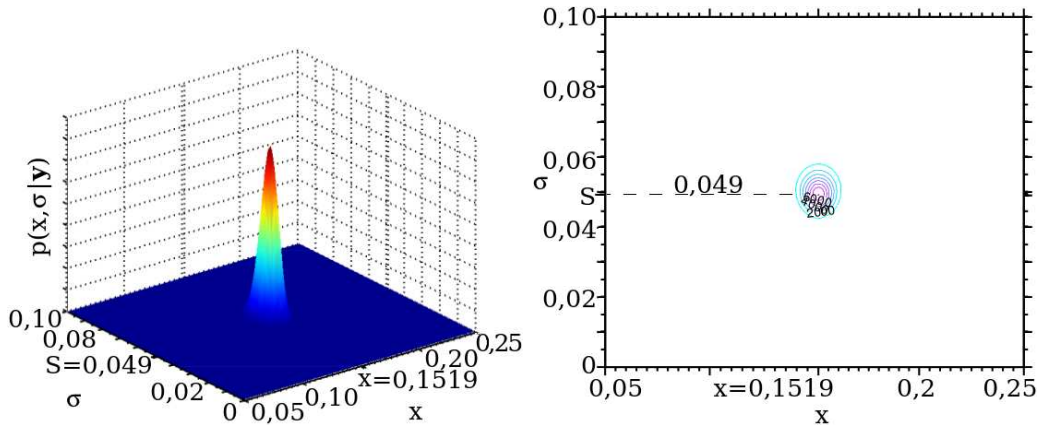


Figura 3.5: Probabilidade marginal $P(x, \sigma|\mathbf{y})$ em função de σ e x para $N = 100$.

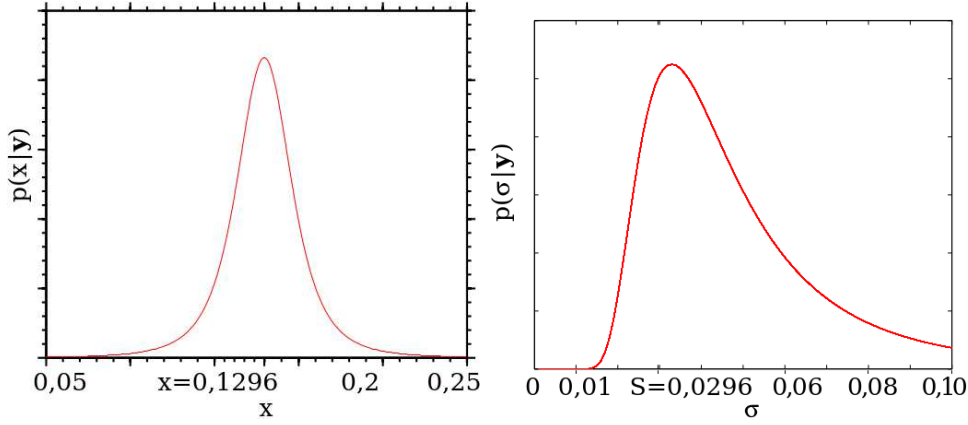


Figura 3.6: Probabilidades marginais $P(x|\mathbf{y})$ (esquerda) e $P(\sigma|\mathbf{y})$ (à direita), em função de σ e X , no casos em que $N = 5$.

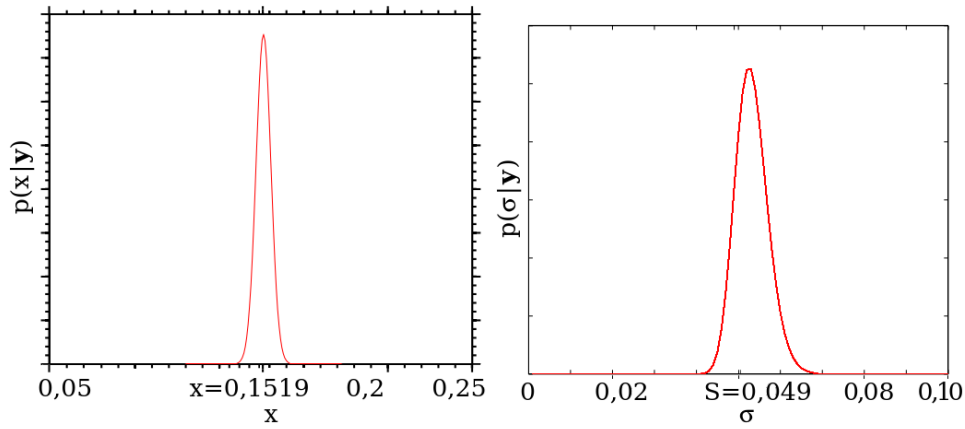


Figura 3.7: Probabilidades marginais $P(x|\mathbf{y})$ (esquerda) e $P(\sigma|\mathbf{y})$ (à direita), em função de σ e X , no casos em que $N = 100$.

$$\begin{aligned} E[(\sigma - E[\sigma|\mathbf{y}])^2|\mathbf{y}] &= E[\sigma^2|\mathbf{y}] - (E[\sigma|\mathbf{y}])^2 \\ &= \frac{NS^2}{N-4} - \frac{1}{2} \left(\frac{\Gamma(\frac{N-3}{2})}{\Gamma(\frac{N-2}{2})} \right)^2 NS^2, \end{aligned} \quad (3.17)$$

$$\begin{aligned} E[(x - E[x|\mathbf{y}])(\sigma - E[\sigma|\mathbf{y}]|\mathbf{y})] &= E[x\sigma|\mathbf{y}] - E[x|\mathbf{y}]E[\sigma|\mathbf{y}] \\ &= 0. \end{aligned} \quad (3.18)$$

Observa-se que a variância é finita a partir de $N > 4$. Um exemplo similar pode ser encontrado na referência [15].

3.5

Estimação por mínimos quadrados

Considere um processo cujo ruído apresenta distribuição gaussiana. Então os dados atuais são representados por

$$\mathbf{y} = \hat{\mathbf{y}}(\mathbf{x}) + \mathbf{n}, \quad (3.19)$$

onde \mathbf{n} é o vetor ruído, $\hat{\mathbf{Y}}$ é o vetor dos dados livres de ruído, e \mathbf{X} é o vetor de parâmetros verdadeiros.

Supondo que $P(\mathbf{x})$ também é considerada uma constante na Eq.(3.1) tem-se que

$$p(\mathbf{x}|\mathbf{y}) \approx \mathbb{N}p(\mathbf{y}|\mathbf{x}). \quad (3.20)$$

Neste caso a função verosimilhança é dada por

$$\begin{aligned} p(\mathbf{y}|\mathbf{x}) &= p(\mathbf{n} = \mathbf{y} - \hat{\mathbf{y}}(\mathbf{x})) \\ &= \frac{1}{(2\pi)^{N/2} \sqrt{\det[\Gamma]}} e^{-\frac{1}{2} (\mathbf{y} - \hat{\mathbf{y}}(\mathbf{x}))^T [\Gamma]^{-1} (\mathbf{y} - \hat{\mathbf{y}}(\mathbf{x}))}, \end{aligned} \quad (3.21)$$

onde a matriz de covariância é definida por

$$[\Gamma] = E \left[(\mathbf{y} - \hat{\mathbf{y}}(\mathbf{x}))^T (\mathbf{y} - \hat{\mathbf{y}}(\mathbf{x})) \right]. \quad (3.22)$$

Tomando o logaritmo de $p(\mathbf{y}|\mathbf{x})$ tem-se

$$\log(p(\mathbf{y}|\mathbf{x})) = \text{cte} - \frac{1}{2} \varepsilon(\mathbf{x}; \mathbf{y}), \quad (3.23)$$

onde

$$\text{cte} = -\log \left((2\pi)^{N/2} \sqrt{\det[\Gamma]} \right), \quad (3.24)$$

e a função de desfasagem é definida como

$$\varepsilon(\mathbf{x}; \mathbf{y}) = (\mathbf{y} - \hat{\mathbf{y}}(\mathbf{x}))^T [\Gamma]^{-1} (\mathbf{y} - \hat{\mathbf{y}}(\mathbf{x})). \quad (3.25)$$

Pode-se observar que minimizar $\varepsilon(\mathbf{x}; \mathbf{y})$ maximiza a função de verosimilhança. Este método é conhecido como mínimos quadrados. É dizer que o método de mínimos quadrados é equivalente ao método de máximo de verosimilhança para sistemas lineares com ruído gaussiano.

3.6

Estimação num sistema com 1 grau de liberdade

Como exemplo de aplicação do método dos mínimos quadrados, considere o sistema massa-mola-amortecedor de um grau de liberdade apresentado no capítulo 2, cuja dinâmica é descrita por

$$\ddot{z}(t) + 2\zeta\omega_0\dot{z}(t) + \omega_0^2 z(t) = \frac{f(t)}{m} - \ddot{x}_b(t) = (m^{-1})q_z(t), \quad (3.26)$$

onde $z(t) = x(t) - x_b(t)$ é a coordenada da massa relativa à base, $\omega_0 = \sqrt{\frac{k}{m}}$ a frequência natural do sistema e $\zeta = \frac{c}{2\sqrt{mk}}$ o fator de amortecimento.

Supondo o forçamento na mola como sendo $f(t) = 0$, só se tem o movimento de base dado pela seguinte equação

$$q_z(s) = -m\ddot{x}_b(t) = -mA_b \sin(\Omega_b t), \quad (3.27)$$

onde A_b e Ω_b são parâmetros que correspondem com a amplitude e a frequência da aceleração da base do sistema.

Assim, segue que

$$\begin{aligned} \hat{z}(t) &= \int_0^t \frac{e^{-\zeta\omega_0(t-s)}}{m\omega_a} \sin(\omega_a(t-s))(-mA_b \sin(\Omega_b s)) ds \\ &= \left(\frac{-1}{\omega_a} \int_0^t \sin(\omega_a(t-s)) \sin(\Omega_b s) e^{-\zeta\omega_0(t-s)} ds \right) A_b \\ &= C(\Omega_b, t) A_b, \end{aligned} \quad (3.28)$$

onde A_b e Ω_b serão tratados como parâmetros linear (X_l) e não linear (X_{nl}) respectivamente.

Supondo que os parâmetros $A_b = X_l$ e $\Omega_b = X_{nl}$ não são conhecidos, procura-se estimá-los pelo método de mínimos quadrados a partir de observação do deslocamento $z(t)$.

Fazendo uso MATLAB, são geradas N amostras do deslocamento, representado pela variável aleatória Z . Esses valores são armazenados no vetor $\mathbf{Z} = (Z_1, Z_2, \dots, Z_i, \dots, Z_N)^T$, onde $Z_i = Z(t_i)$ é a observação do deslocamento Z ao tempo t_i . Assim, pode-se escrever cada valor medido como o valor verdadeiro, dado pela Eq.(3.29), mais um ruído n_i estocástico independente de tipo gaussiano com média zero e desvio padrão σ_n . Logo,

$$\begin{aligned} Z_i &= \hat{z}_i + n_i \\ &= C(\Omega_b, t_i) A_b + n_i, \end{aligned} \quad (3.29)$$

ou em forma vetorial

$$\begin{aligned} \mathbf{Z} &= \hat{\mathbf{z}} + \mathbf{n} \\ &= \mathbf{C}(\Omega_b) A_b + \mathbf{n}. \end{aligned} \quad (3.30)$$

Segundo a Eq.(3.21), sendo $\mathbf{X} = [X_{nl}, X_l]^T$ o vetor de parâmetros, tem-se

$$\begin{aligned} p(\mathbf{z}|\mathbf{x}) &= \frac{1}{(2\pi\sigma_n^2)^{N/2}} e^{-\frac{1}{2\sigma_n^2}(\mathbf{z}-\hat{\mathbf{z}})^T(\mathbf{z}-\hat{\mathbf{z}})}, \\ &= \frac{1}{(2\pi\sigma_n^2)^{N/2}} e^{-\frac{1}{2\sigma_n^2}(\mathbf{z}-\mathbf{C}(\Omega_b)A_b)^T(\mathbf{z}-\mathbf{C}(\Omega_b)A_b)}, \end{aligned} \quad (3.31)$$

onde a matriz de covariância é $[\Gamma] = \sigma_n^2[I]$.

Se a probabilidade priori é quase uniforme, de forma tal que $p(\mathbf{z})$ pode ser considerada constante tem-se que a probabilidade posterior é dada por

$$p(\mathbf{x}|\mathbf{z}) \propto \aleph e^{-\frac{1}{2\sigma_n^2}(\mathbf{Z}-\mathbf{C}(\Omega_b)A_b)^T(\mathbf{z}-\mathbf{C}(\Omega_b)A_b)}, \quad (3.32)$$

onde, como dito anteriormente, \aleph é uma constante de normalização.

E a função de desfasagem, segundo a Eq.(3.25), é da forma

$$\begin{aligned} \varepsilon(\mathbf{x}; \mathbf{z}) &= \frac{1}{\sigma_n^2} (\mathbf{z} - \hat{\mathbf{z}})^T (\mathbf{z} - \hat{\mathbf{z}}) \\ &= \frac{1}{\sigma_n^2} (\mathbf{z} - \mathbf{C}(\Omega_b)A_b)^T (\mathbf{z} - \mathbf{C}(\Omega_b)A_b) \\ &= \frac{1}{\sigma_n^2} (\mathbf{z} - \mathbf{C}(x_{nl})x_l)^T (\mathbf{z} - \mathbf{C}(x_{nl})x_l). \end{aligned} \quad (3.33)$$

Desenvolvendo-se a equação acima obtêm-se

$$\varepsilon(\mathbf{x}; \mathbf{z}) = \frac{1}{\sigma_n^2} [\mathbf{z}^T \mathbf{z} + x_l \mathbf{C}(x_{nl})^T \mathbf{C}(x_{nl})x_l - 2\mathbf{z}^T \mathbf{C}(x_{nl})x_l]. \quad (3.34)$$

Considerando o parâmetro não linear como conhecido, diferenciando a Eq.(3.34) com respeito ao parâmetro linear, e igualando ao zero tem-se

$$\frac{\partial \varepsilon}{\partial x_l} = \frac{1}{\sigma_n^2} [2\mathbf{C}(x_{nl})^T \mathbf{C}(x_{nl})x_l - 2\mathbf{z}^T \mathbf{C}(x_{nl})] = 0, \quad (3.35)$$

donde

$$x_l = \frac{\mathbf{z}^T \mathbf{C}(x_{nl})}{\mathbf{C}(x_{nl})^T \mathbf{C}(x_{nl})}. \quad (3.36)$$

Logo, a função de desfasagem pode ser expressa em termos do parâmetro não linear apenas

$$\varepsilon(\mathbf{x}; \mathbf{z}) = \varepsilon(x_{nl}, x_l(X_{nl}); \mathbf{z}) = \bar{\varepsilon}(x_{nl}; \mathbf{z}). \quad (3.37)$$

Assim, procura-se estimar o valor de X_{nl} que minimize $\bar{\varepsilon}(x_{nl}; \mathbf{z})$, e logo com a Eq.(3.36) encontra-se a a estimação do valor para X_l .

Existem muitos métodos de minimizar funções, mas estes não são objetivo deste trabalho. Aqui foi usada a função `fminsearch` do MATLAB [9]. Esta, acha o mínimo de uma função escalar de várias variáveis, a partir de uma estimativa inicial. Geralmente é referida à otimizações irrestritas não lineares. O algoritmo implementado nesta é conhecido como método Simplex [10]. Este método é de procura direta então não utiliza gradientes. O que pode ser um problema na hora de procurar mínimos globais.

A figura 3.8 apresenta a solução da Eq.(3.26), junto com a resposta impulso unitário, utilizando as seguintes valores $m = 100 [km]$, $k = 10000 [N/m]$, $c = 40 [N \cdot s/m]$, $A_b = 1 [m]$ e $\Omega_b = 0,5 [1/s]$.

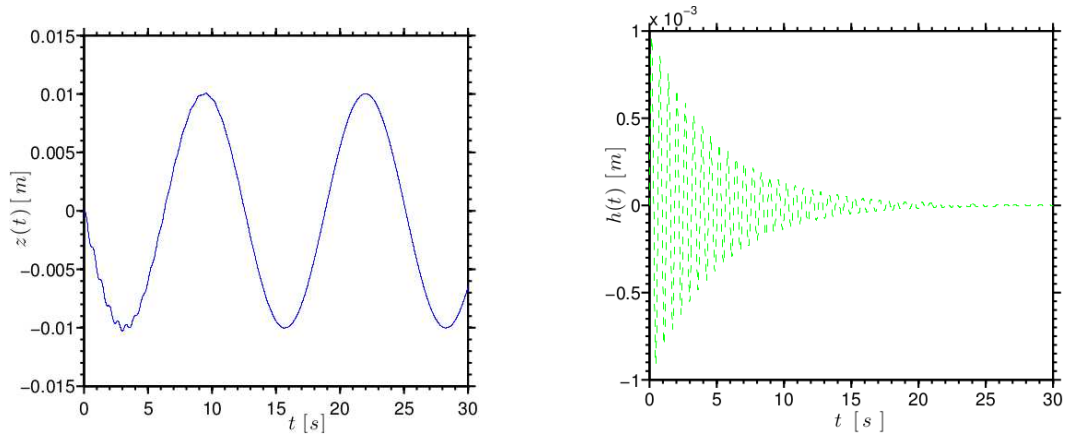


Figura 3.8: Deslocamento e função resposta ao impulso unitário do sistema com 1 grau de liberdade cujas propriedades estão sendo estimadas.

A Figura 3.9 mostra o comportamento da solução (primeira linha), da função de forçamento (segunda linha) e a convergência (terceira linha) para diferentes número de amostras, para um valor de desvio padrão de $\sigma_n = 0,002$, e chute inicial relativamente próximo ao valor real ($x_{nl_0} = 0,4 [1/s]$). São realizados três eventos, com diferentes números de amostras, sendo para cada coluna, da esquerda para direita, respectivamente, $N = 16$, $N = 31$ e $N = 301$. Na mesma figura, as curvas de cor verde são dadas pelos valores originais, as de cor vermelho são as curvas identificadas pelo método de mínimos quadrados e as observações são representadas pelas cruzes azuis.

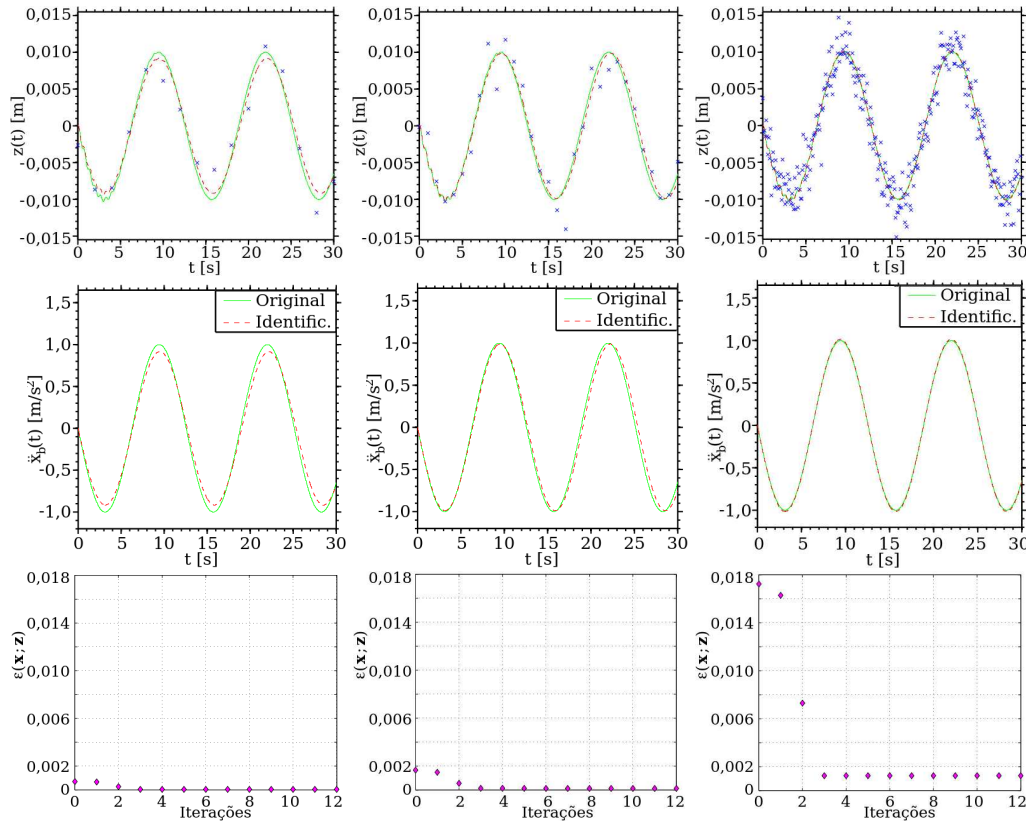


Figura 3.9: Deslocamentos, forçamento e convergências, para $\sigma_n = 0,002$.

Observa-se que para o valor de $\sigma_n = 0,002$, ou seja, 20% do valor da amplitude da solução original, o método consegue reproduzir bem a curva da solução, mesmo para uma quantidade pequena de observações. Os valores estimados para as configurações dadas na Figura 3.9 são apresentados na Tabela 3.1.

$\sigma_n = 0,002$	$N = 16$	$N = 31$	$N = 301$	Valores originais
X_{nl}	0,495781	0,494609	0,499844	0,5
X_l	0,915203	0,986691	1.013463	1,0

Tabela 3.1: Estimação de parâmetros referente à Figura 3.9.

Novamente são realizadas as simulações, mas agora com o valor de desvio padrão de $\sigma_n = 0,006$ (i.e., 60% do valor da amplitude da solução original), como é apresentado na Figura 3.10. O método não conseguiu estimar os valores dos parâmetros para $N = 16$ amostras. Então, são apresentados os casos em que $N = 31$, $N = 61$ e $N = 301$.

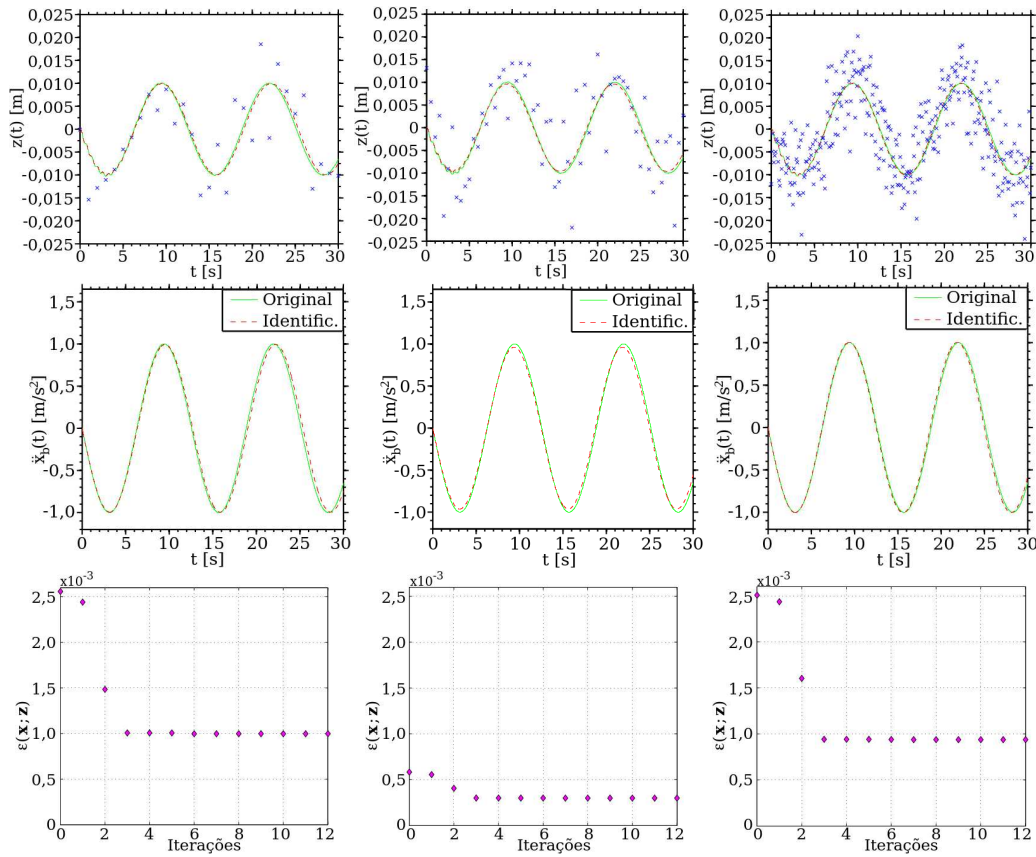


Figura 3.10: Deslocamentos, forçamento e convergências, para $\sigma_n = 0,006$.

A Tabela 3.2 apresenta os valores estimados para as configurações dadas na Figura 3.10. Como é de se esperar, o método começa a se comportar de forma instável para grandes dispersões (no caso, 60%).

$\sigma_n = 0,006$	$N = 31$	$N = 61$	$N = 301$	Valores originais
$X_{nl} [1/s]$	0,495313	0,501719	0,503045	0,5
$X_l [m]$	0,993527	0,994837	1.003204	1,0

Tabela 3.2: Estimação de parâmetros referente à Figura 3.10

A fim de ilustrar a instabilidade do método para grandes dispersões, a Figura 3.11 mostra os resultados obtidos ao tentar reproduzir as mesmas eventos da Figura 3.10 nos casos de $N = 31$ e $N = 61$ amostras. Conclui-se que no caso de grandes dispersão, o método de mínimos quadrados é instável para pequenas quantidades de amostras. No entanto, se o número de amostras for grande, o método converge e assim os parâmetros são estimados (por exemplo, para $N = 301$).

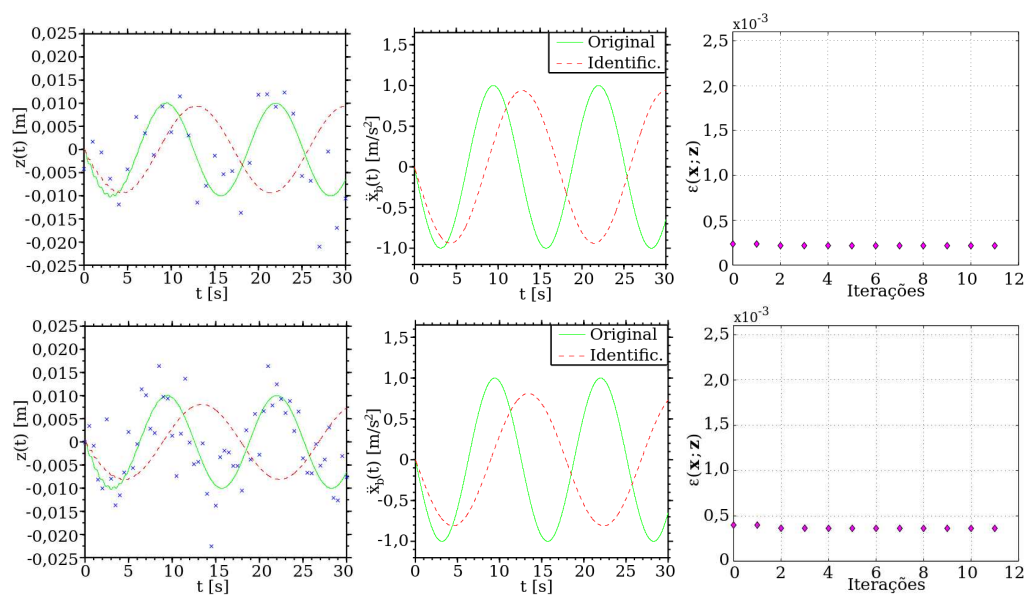


Figura 3.11: Deslocamentos, forçamento e convergências, para $\sigma_n = 0,006$, $N = 31$ e $N = 61$.

A Tabela 3.3 apresenta os valores estimados para as configurações dadas na Figura 3.11.

$\sigma_n = 0,006$	$N = 31$	$N = 61$	Valores originais
$X_{nl} [1/s]$	0,358594	0,486406	0,5
$X_l [m]$	0,978429	0,955785	1,0

Tabela 3.3: Estimação de parâmetros referente à Figura 3.11.

4

Estimação Paramétrica Recursiva

Neste capítulo apresenta-se o filtro de Kalman como uma ferramenta para se abordar um problema linear recursivo de estimação paramétrica e o filtro de Kalman estendido para problemas recursivos de natureza não linear.

4.1

Filtro de Kalman sem ruído no modelo

No capítulo anterior, assumiu-se que os dados foram coletados antes de serem processados para fazer a estimação de parâmetros. Neste processo deve-se achar a solução de um sistema de N -equações lineares de forma simultânea (onde N é o número de parâmetros não conhecidos). Entretanto, na prática, existem muitas situações nas que não é possível esperar para obter todos os dados e logo depois fazer a estimação dos parâmetros. Assim sendo, os algoritmos recursivos ganham terreno na prática, pois fornecem uma forma de se obter uma estimação que pode ser atualizada a cada novo dado coletado.

As ideias chaves do algoritmo recursivo do filtro de Kalman, sem erro no modelo, podem ser enunciadas como a seguir:

- O teorema de Bayes pode ser usado sequencialmente de forma a que considere várias observações.
- A PDF priori, de tipo gaussiana, é convertida na PDF posterior, também de tipo gaussiana, quando a função de verossimilhança é gaussiana.
- O uso de um conjunto pequeno de estatísticas suficientes com uma caracterização completa do estado de conhecimento dos parâmetros dados a PDF priori e todos os dados recoletados até o momento.

Suponha que a estimativa procurada do vetor de parâmetros \mathbf{X} para os k dados coletados até o momento é da forma

$$\mathbf{y}_k = [C_k] \mathbf{x}_k + \mathbf{n}_k, \quad (4.1)$$

onde \mathbf{n}_k sendo um processo gaussiano (ruído gaussiano) com média nula e matriz de covariância $[\Gamma_k]$.

Assuma também que, logo que a observação do novo dado \mathbf{Y}_k seja realizada, o estado de conhecimento de \mathbf{X} é representado pela PDF priori, a qual foi modelada como sendo gaussiana, com média $\bar{\mathbf{X}}_k$ e matriz covariância $[P_k]$.

Assim, é possível adotar o seguinte procedimento

1. Na medição dos dados \mathbf{Y}_k , para se calcular a média $\tilde{\mathbf{X}}_k$ e a matriz covariância $[\tilde{P}_k]$, as seguintes equações são utilizadas

$$[\tilde{P}_k] = \left([P_k]^{-1} + [C_k]^T [\Gamma_k]^{-1} [C_k] \right)^{-1}, \quad (4.2)$$

$$\tilde{\mathbf{x}}_k = [\tilde{P}_k] \left([C_k]^T [\Gamma_k]^{-1} \mathbf{y}_k + [P_k]^{-1} \right). \quad (4.3)$$

2. Usa-se a PDF a priori para o vetor de dados como a PDF posterior para o dado seguinte, de forma que

$$\bar{\mathbf{x}}_{k+1} = [\Phi_k] \tilde{\mathbf{x}}_k + \mathbf{b}_k, \quad (4.4)$$

e

$$[P_{k+1}] = [\Phi_k] [\tilde{P}_k] [\Phi_k]^T. \quad (4.5)$$

3. O passo de acima é repetido com os sucessivos vetores de dados atuais para cada observação, obtendo assim a sequencia de PDFs posteriores.

Uma implementação eficiente do algoritmo acima é dada pela fórmula de Woodbury [15]:

$$[\tilde{P}_k] = [P_k] - [P_k] [C_k]^T \left([\Gamma_k] + [C_k] [P_k] [C_k]^T \right)^{-1} [C_k] [P_k], \quad (4.6)$$

que não requer o cálculo de $[P]^{-1}$.

A inversa de $[\Gamma_k] + [C_k] [P_k] [C_k]^T$ envolve matrizes de tamanho determinado pelo vetor de dados em vez do vetor de parâmetros. Em particular, se os dados em cada passo são representados por um escalar, a inversão é uma simples divisão.

Devido à forma positiva definida das matrizes $[P_k]$, $[\tilde{P}_k]$ e $[P_k] [C_k]^T \left([\Gamma_k] + [C_k] [P_k] [C_k]^T \right)^{-1} [C_k] [P_k]$ ao se incluir o dado \mathbf{Y}_k ocorre uma redução na variância dos parâmetros estimados.

A forma atualizada para a média pode ser reescrita como a seguir

$$\tilde{\mathbf{x}}_k = \bar{\mathbf{x}}_k + [K_k] (\mathbf{y}_k - [C_k] \bar{\mathbf{x}}_k), \quad (4.7)$$

onde a matriz de ganho $[K_k]$ é dada por

$$[K_k] = [P_k] [C_k]^T \left([\Gamma_k] + [C_k] [P_k] [C_k]^T \right)^{-1}, \quad (4.8)$$

e a matriz de covariância atualizada é

$$[\tilde{P}_k] = ([I] - [K_k][C_k])[P_k]. \quad (4.9)$$

Assim, pode-se pensar o algoritmo recursivo discreto do filtro de Kalman, sem ruído no modelo, como em um sistema de retroalimentação o qual toma o erro da sinal como $\mathbf{n}_k = \mathbf{Y}_k - [C_k]\mathbf{X}$ baseado na estimação atual e é processado usando $[K_k]$ para produzir uma atualização a qual se soma com \mathbf{X}_k para obter $\tilde{\mathbf{X}}_k$.

4.2

Estimação linear sem ruído no modelo

Considere novamente o exemplo do sistema massa-mola-amortecedor com 1 grau de liberdade apresentado no capítulo 3, supondo agora que o parâmetro não linear Ω_b seja conhecido, e o parâmetro linear A_b é o que se deseja estimar.

Nesse caso tem-se o seguinte sistema de equações

$$\begin{cases} A_{b_{k+1}} = [\Phi_k] A_{b_k}, \\ Z_k = \hat{z}_k + n_k = C(t_k)A_b + n_k = C_k A_b + n_k \quad (\text{com } k = 1, 2, 3, \dots), \end{cases} \quad (4.10)$$

onde a matriz de evolução é $[\Phi_k] = 1$, devido ao fato de estar se estimando um número, que não evolui no tempo.

Suponha-se que a covariância das amostras é $[\Gamma] = \sigma^2 [I]$, e que o primeiro parâmetro A_{b_1} é conhecido (chute inicial). Segue que $\bar{A}_{b_1} = A_{b_1}$ e $P_1 = \infty$, donde conclui-se que $p(\bar{A}_{b_1}) = 1$, o que concorda com o fato do chute inicial ser conhecido.

Em seguida acha-se K_1 e \tilde{P}_1 através das Eqs.(4.8) e (4.9) respectivamente, sendo que $K_1 = C_1^{-1}$, onde $C_1 = C(t_1)$ e $P_2 = \tilde{P}_1 = \frac{\sigma^2}{C_1^2}$. Pela Eq.(4.7) para a primeira observação (Z_1) verifica-se então que

$$\tilde{A}_{b_1} = \bar{A}_{b_1} + C_1^{-1} (Z_1 - C_1 \bar{A}_{b_1}). \quad (4.11)$$

Depois, faz-se $\bar{A}_{b_2} = \tilde{A}_{b_1}$, $P_2 = \tilde{P}_1$, e com as Eqs.(4.8) e (4.9) acha-se K_2 e \tilde{P}_2 respectivamente, e substituindo na Eq.(4.7), obtem-se o valor atualizado de \tilde{A}_{b_2} e assim por diante.

Nas Figuras 4.1, 4.2 e 4.3, o leitor pode observar o comportamento da estimação recursiva de parâmetros, dado pelo filtro de Kalman sem ruído no modelo, para diferentes variabilidades nos dados e número de amostras usadas na estimação. Na primeira linha para um erro nos dados de 4%, com respeito ao valor da amplitude da resposta, dando um chute inicial de $A_{b_1} = 0,9$, próximo ao valor verdadeiro que se deseja estimar ($A_b = 1$). Já na segunda

linha são apresentados os resultados com erro nos dados de 40% e o chute inicial é longe do valor verdadeiro ($A_{b1} = 0,4$). Os dois casos são estudados para diferentes quantidades de amostras, sendo que a primeira, a segunda e a terceira coluna correspondem respectivamente a uma quantidade de amostras dadas por $N = 76$, $N = 151$ e $N = 301$.

A Figura 4.1 mostra o deslocamento do sistema, onde as curvas de cor verde correspondem às soluções verdadeiras, as de cor azul às reconstruídas e as cruzes representam as observações.

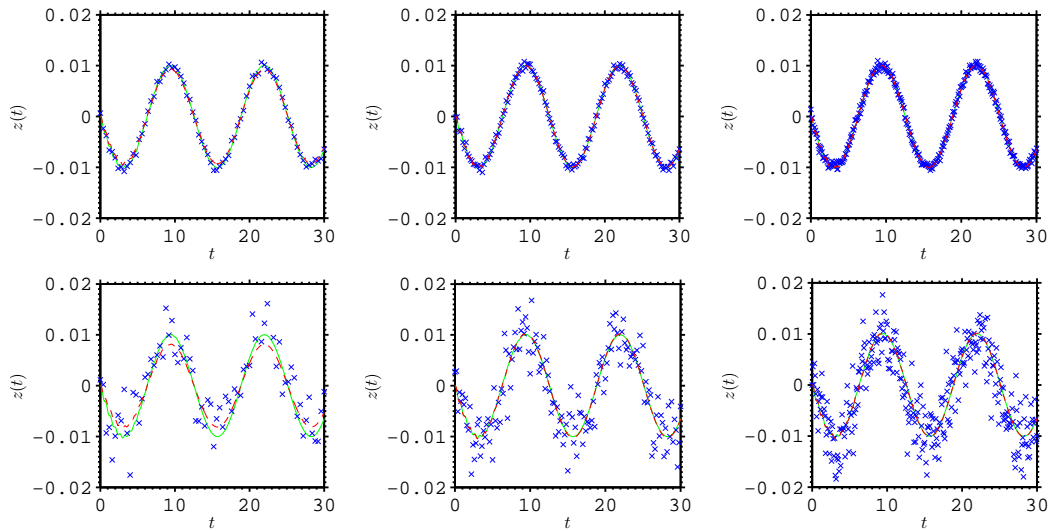


Figura 4.1: Curvas das soluções.

Já a Figura 4.2 apresenta os forçamentos ao qual o sistema está sujeito, junto com a curva reconstruída pelo método recursivo de filtro de Kalman sem ruído no modelo.

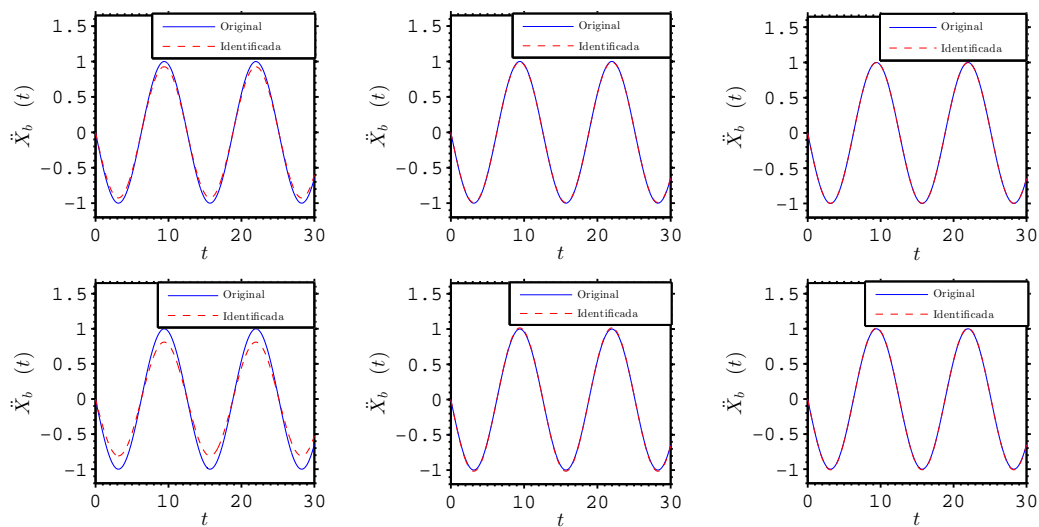


Figura 4.2: Curvas dos forçamentos original e reconstruído

A Figura 4.3 apresenta as curvas de convergência do modelo. As mesmas são representadas pelo valor melhorado da amplitude A para cada nova observação. O primeiro destaque da observação do comportamento do método é que aumentando o número de medições, a convergência melhora mas o valor final da mesma depende da acurácia das observações. Este comportamento pode ser compreendido ao se pensar que o valor do parâmetro A_{b_k} no tempo t_k é melhorado com a observação Z_k , que pode estar longe do valor verdadeiro, podendo assim piorar o valor melhorado \bar{A}_k .

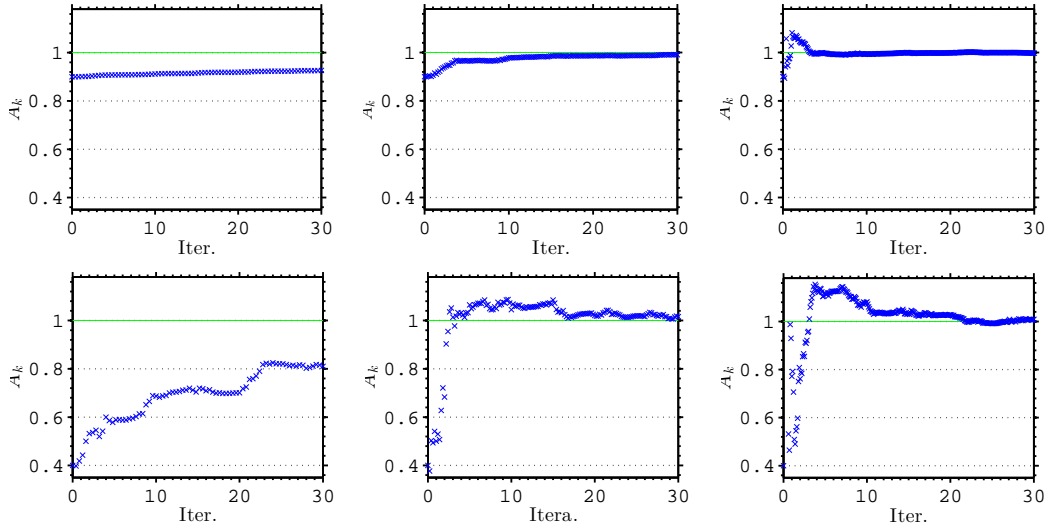


Figura 4.3: Curvas de convergência do modelo, representada pelo valor melhorado do parâmetro a estimar (A_b) para cada nova medição Z_k .

4.3

Filtro de Kalman com ruído no modelo

Considera-se agora a existência de ruído no modelo, de forma que a equação de evolução temporal é modificada adicionando um vetor erro \mathbf{e}_k e uma matriz de covariância do erro $[S_k]$, como a seguir

$$\mathbf{x}_{k+1} = [\Phi_k] \mathbf{x}_k + \mathbf{b}_k + \mathbf{e}_k, \quad (4.12)$$

e

$$[P_{k+1}] = [\Phi_k] [P_k] [\Phi_k]^T + [S_k]. \quad (4.13)$$

Assim, tendo em conta que logo antes da medição de \mathbf{Z}_k o estado de conhecimento é descrito por $\bar{\mathbf{X}}_k$ e $[P_k]$, e tomando a média da Eq.(4.12) obtêm-se

$$\bar{\mathbf{x}}_{k+1} = [\Phi_k] \bar{\mathbf{x}}_k + \mathbf{b}_k. \quad (4.14)$$

Sendo o ruído \mathbf{e}_k um processo gaussiano com média nula e matriz de covariância $[S_k] = R^2 [I]$, tem-se, como o resultado de medir \mathbf{Z}_k e atualizar o valor dos parâmetro, a seguinte sequência:

1. A covariância da PDF priori de \mathbf{X}_k

$$[\tilde{P}_k] = [P_k] - [P_k] [C_k]^T ([\Gamma_k] + [C_k] [P_k] [C_k]^T)^{-1} [C_k] [P_k]. \quad (4.15)$$

2. A matriz $[K_k] = [\tilde{P}_k] [C_k] [\Gamma_k]^{-1}$ (matriz de ganho de Kalman).
3. A média da PDF posterior de \mathbf{X}_k , é dizer, o estado do filtro de Kalman

$$\tilde{\mathbf{x}}_k = \bar{\mathbf{x}}_k + [K_k] (\mathbf{Z}_k - [C_k] \bar{\mathbf{x}}_k). \quad (4.16)$$

4. A covariância da PDF priori de \mathbf{X}_{k+1}

$$[P_{k+1}] = [\Phi_k] [\tilde{P}_k] [\Phi_k]^T + [S_k]. \quad (4.17)$$

5. A média da PDF priori de \mathbf{X}_{k+1} , é dizer, a predição de Kalman do novo estado

$$\bar{\mathbf{x}}_{k+1} = [\Phi_k] \tilde{\mathbf{x}}_k + \mathbf{b}_k. \quad (4.18)$$

4.4

Estimação linear de um parâmetro constante com ruído no modelo

Novamente pense no exemplo do sistema massa-mola-amortecedor com 1 grau de liberdade onde somente o parâmetro linear do forçamento (A_b) é não conhecido. Mas agora é considerado que o modelo matemático não considera todos os fenômenos envolvidos na dinâmica, de modo que quer-se corrigir isso com o filtro de Kalman através da inclusão de um ruído.

Assuma que tanto o erro no modelo e_k , quanto o erro das medições n_k , para cada passo de tempo, são do tipo gaussiano com média nula e matrizes de covariância $[S_k] = R^2 [I]$ e $[\Gamma_k] = \sigma_n^2 [I]$ respectivamente.

Assim, o problema é da forma

$$\begin{cases} A_{b_{k+1}} = \Phi_K A_{b_k} + e_k, \\ Z_k = \hat{z}_k + n_k = C_k A_b + n_k \quad (\text{com } k = 1, 2, 3, \dots), \end{cases} \quad (4.19)$$

Ou seja, $[\Phi_k] = 1$ e $\mathbf{b}_k = 0$. Com isso, as equações dos passos 4 e 5 são: $P_{k+1} = \tilde{P}_k + R^2$ e $\bar{X}_{k+1} = \tilde{X}_k$.

Dando um (pequeno) valor de 4% para a variância para observações, com respeito à amplitude da solução original e um valor de chute inicial afastado do valor verdadeiro de $A_{b_0} = 0,4$, variam-se os valores de variância para o erro do modelo ($R = 0,10; 0,25; 0,50; 0,75$, sendo do 10%, 25%, 50% e 75%, com

respeito ao valor original do parâmetro linear), e se observam as curvas para diferentes quantidades de observações ($N = 76 ; 151 ; 301$). Assim, tem-se as curvas como descrito a seguir.

A Figura 4.4 apresenta as curvas de convergência. Sendo para as colunas da esquerda para a direita os valores $N = 76$, $N = 151$, e $N = 301$ respectivamente, e de cima para baixo representam-se aos valores $R = 0, 10$, $R = 0, 25$ e $R = 0, 50$. Os valores da covariância e do chute inicial são mantidos iguais a $\sigma_n = 0.0004$ e a $A_{b_0} = 0,4$. Um comportamento rápido é observado ao longo da primeira coluna, enquanto a variância do erro do modelo cresce, a convergência parece ser mais rápida. Agora, observando a última coluna pode-se notar que enquanto R aumenta, a acurácia da predição aumenta. Mais na frente neste texto, vai-se voltar nesta discussão. O deslocamento do sistema e o forçamento ao qual ele está sujeito podem ser vistos na Figura 4.5. Outros casos não foram apresentados já que não apresentam muita diferença.

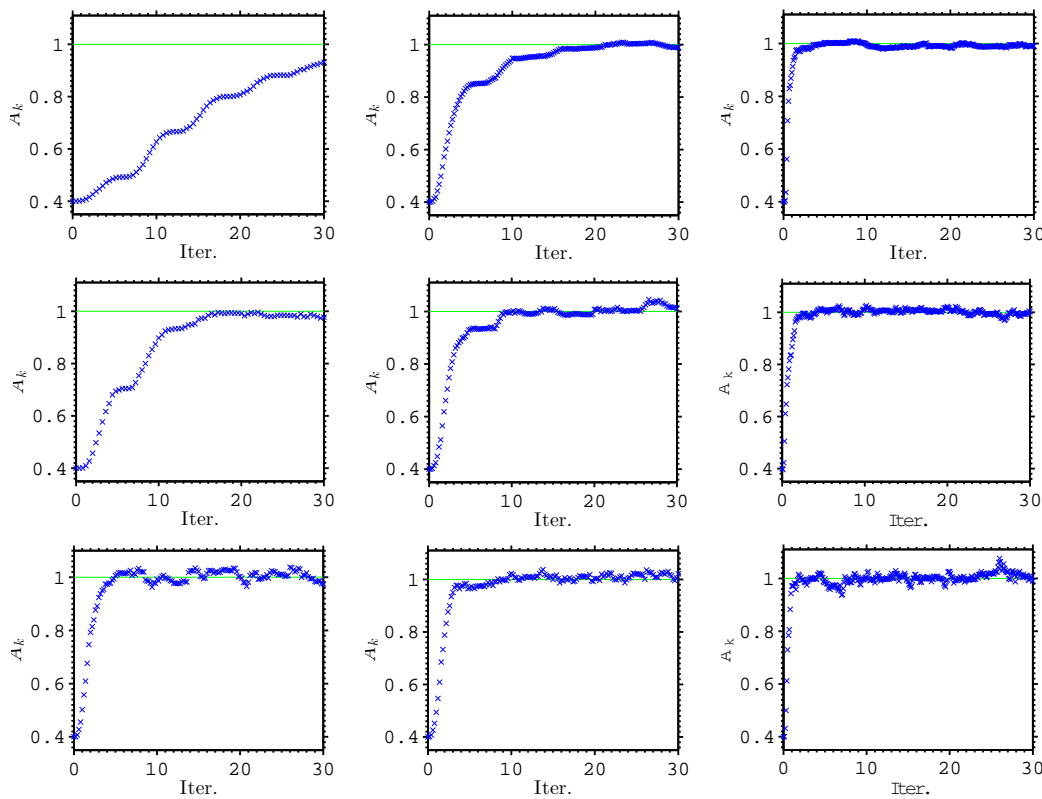


Figura 4.4: Curvas de convergência.

Também foram analisados os casos quando o valor do erro das amostras vai até um valor de 40% com respeito do valor da amplitude original do forçamento, variando o valor da variância do erro do modelo da mesma forma que antes, e para as mesmas quantidades de amostras. Mas, para estes casos, a convergência começa piorar. Assim, a Figura 4.6 apresenta curvas de

congruência só para os valores de $R = 0,1$ (acima) e $R = 0,25$ (abaixo). Nessa figura podem ser observados comportamentos similares aos achados na Figura 4.4 mas com acurácia pior.

Estes resultados despertam curiosidade de forma a tentar aprofundar a análise de convergência. Para isso, são apresentadas curvas de convergência ao longo de tempos maiores e de um maior número de medições na Figura 4.7. Primeiramente, tanto o erro do modelo como o das amostras são de 10% e de 4% respectivamente. Depois, aumenta-se somente o erro do modelo até 50% e mantém-se o erro das amostras em 4%. Logo faz-se ao contrário, o ruído das medições é aumentado até 40% e o erro no modelo é mantido em 10%. Por último, aumentam-se ambos até 50% para o modelo e 40% para as medições.

Pode-se ver que o erro nas medições apresenta uma maior influência que o erro no modelo. Uma outra coisa que pode-se observar é que ao longo das medições, o modelo sempre converge em direção ao parâmetro desejado. Isso é dizer que se é possível dispor de tempo para realizar observações o modelo sempre vai convergir para a solução, mas sua acurácia depende da qualidade das medições e do erro no modelo.

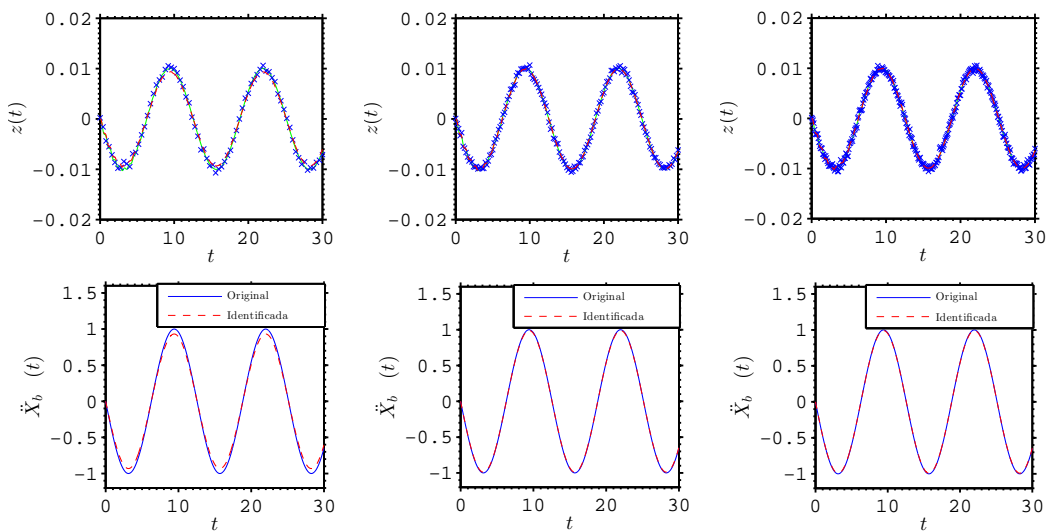


Figura 4.5: Curvas de soluções.

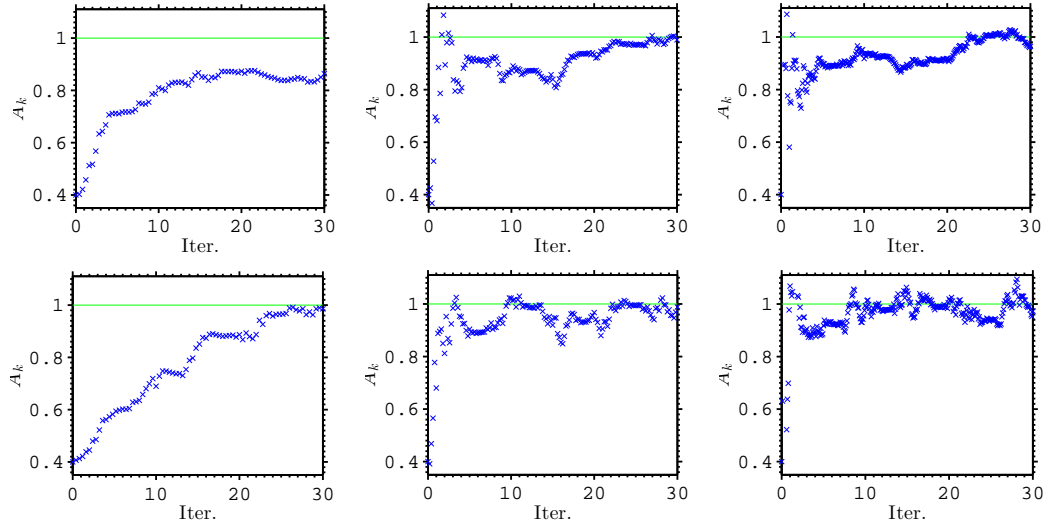


Figura 4.6: Curvas de iterações com $\sigma_n = 0,004$.

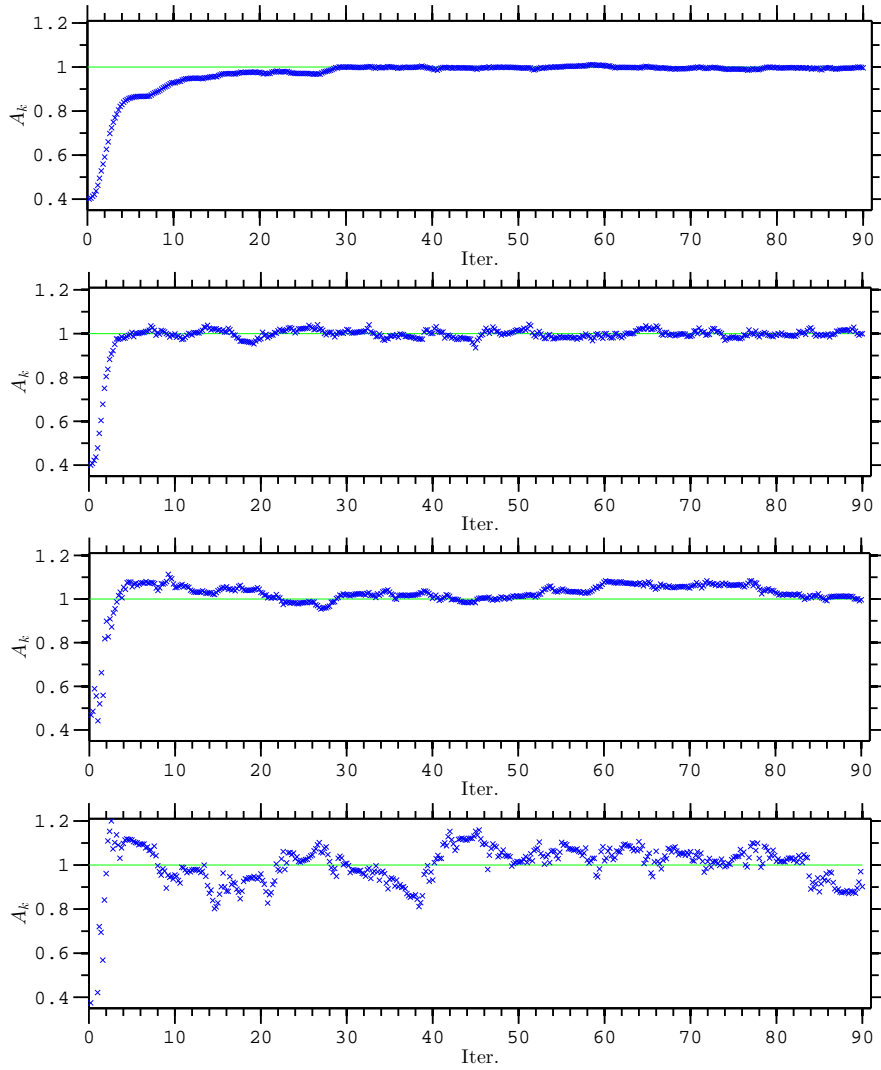


Figura 4.7: Curvas de convergência para um tempo de 90 s e para $N = 451$ medições.

4.5

Estimação linear de um parâmetro que evolui no tempo com ruído no modelo

Se o parâmetro a ser estimado evolui no tempo, o filtro de Kalman é capaz de acompanhar essa evolução. Pode-se imaginar como no mesmo exemplo da seção 4.4, e pensar como se ele estivesse embarcado num recinto o qual oscila com a mesma forma funcional dada pela Eq. 3.27. Desta vez, entretanto o sistema foi excitado em uma frequência de ressonância, de modo que a sua amplitude aumenta conforme

$$\ddot{x}_b = A_b t \sin(\Omega_b t). \quad (4.20)$$

Desta forma o problema a resolver é da forma

$$\begin{cases} A_{b_{k+1}} = \Phi_k A_{b_k} + e_k, \\ Z_k = \hat{z}_k + n_k, \end{cases} \quad (4.21)$$

onde \hat{z}_k é dado pela Eq.(2.20).

Se a matriz de evolução for $\Phi_k = 1$, da mesma forma que no caso estático, a predição do filtro de Kalman consegue acompanhar o crescimento do parâmetro.

Suponha que o valor inicial A_{b_0} é dado pela média inicial $\bar{A}_{b_0} = 0,7m/s$ e a covariância da priori $P_0 = 10$. Assim, a Figura 4.8 apresenta as curvas das soluções para quando $\sigma_n = 0.0004$, $R = 0.2$ e $N = 181$, da PDF priori $P(A_{b_0})$.

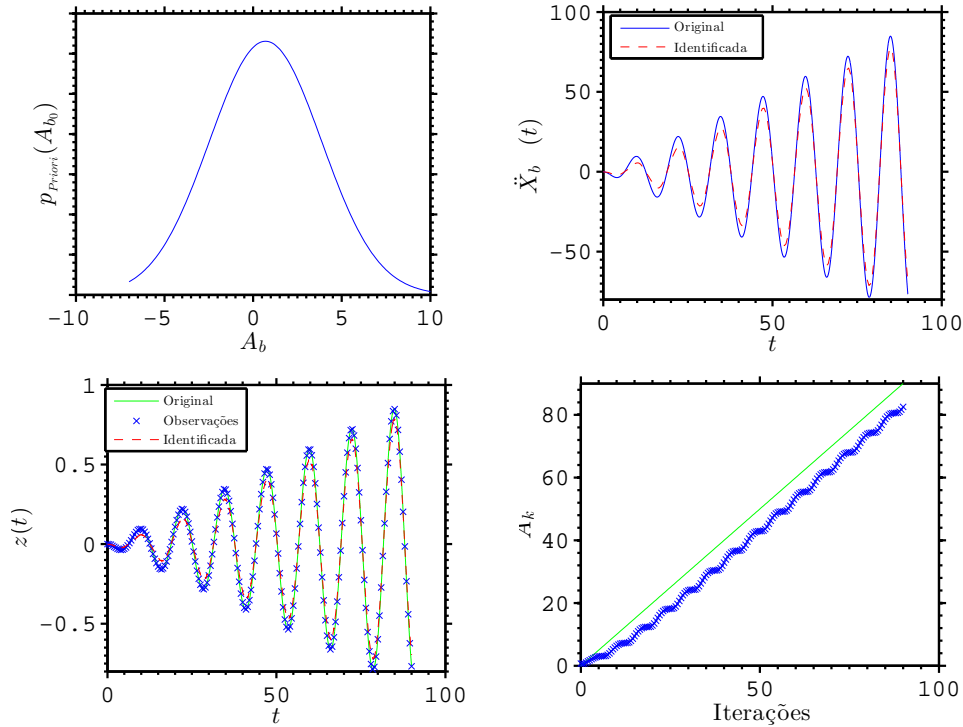


Figura 4.8: Função densidade de probabilidade priori de A_{b_0} .

Observa-se que se $\Phi_k = 1$, então não é possível a boa estimação do parâmetro desejado. Para contornar isso, observe que a evolução, sem ruído, de $X(t=0)$ até $X(t)$ é da forma

$$x(t) = \Phi(t)A_{b_0} = t A_{b_0}, \quad (4.22)$$

e que tomando a derivada temporal da Eq.(4.22) tem-se que $\dot{x}(t) = A_{b_0}$ e $\dot{\Phi}(t) = 1$. Assim, ao tempo t_k tem-se $\Phi(t_k) = t_k$ e discretizando [7] chega-se a

$$\Phi_k = \Phi(t_{k+1})(\Phi(t_k))^{-1} = t_{k+1}/t_k, \quad (4.23)$$

onde tem-se o seguinte problema

$$\begin{cases} A_{b_{k+1}} = \frac{t_{k+1}}{t_k} A_{b_k} + e_k, \\ Z_k = \hat{z}_k + n_k, \end{cases} \quad (4.24)$$

Resolvendo o sistema acima para as mesmas condições do exemplo anterior, têm-se os resultados apresentados na Figura 4.9, onde pode-se observar que o método aproxima o parâmetro sem problema algum.

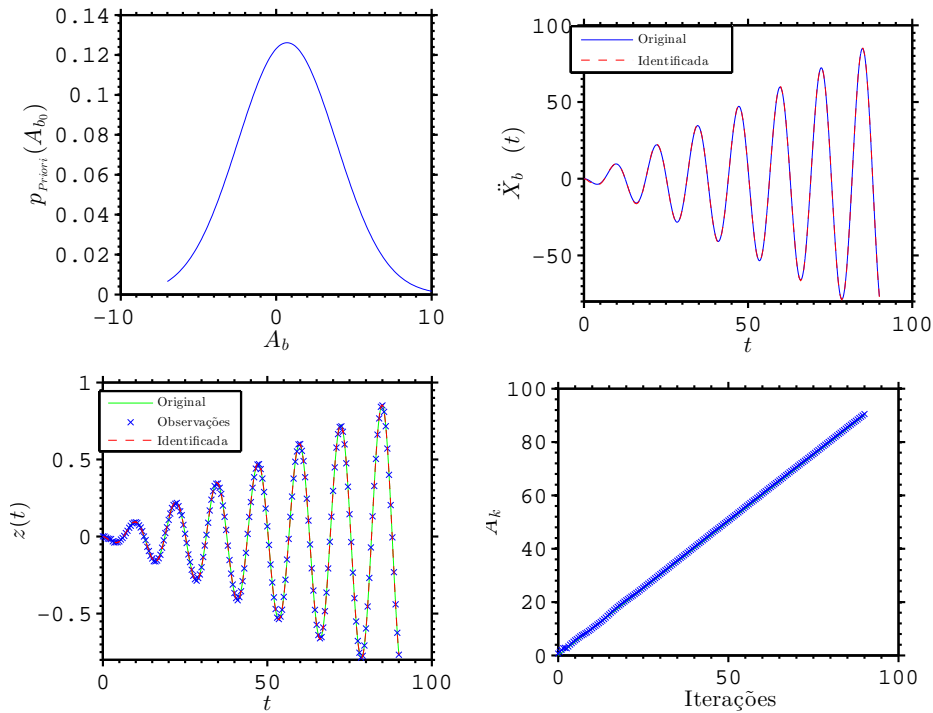


Figura 4.9: Função densidade de probabilidade priori de A_{b_0} .

Note-se que o filtro de Kalman consegue aproximar o comportamento do parâmetro mesmo se a evolução no tempo não for especificada. Mas se a evolução no tempo for corretamente especificada, o filtro apresenta um bom comportamento na estimação.

4.6

Filtro de Kalman estendido

Na prática a não linearidade dos parâmetros é coisa comum. Por isso o interesse em estender o filtro de Kalman para casos que apresentem não linearidades. Assim, o filtro de Kalman estendido pode, também, ser chamado de filtro de Kalman não linear. É claro que a não linearidade diz que a função de verosimilhança falha em ser gaussiana, mesmo quando o ruído é aditivo e gaussiano. Mas, se esta não linearidade não é forte e apresenta flutuações pequenas, é possível aproximar a PDF posterior por funções gaussianas e logo usar métodos para linearizá-las.

Relembrando, a evolução do filtro de Kalman linear foi dada pela matriz $[\Phi]$ e pelo vetor \mathbf{b} , atualizando a matriz covariância e o vetor de estado da seguinte forma

$$[P_{k+1}] = S_k + [\Phi_k] [\tilde{P}_k] [\Phi_k]^T, \quad (4.25)$$

e

$$\bar{\mathbf{x}}_{k+1} = [\Phi_k] \tilde{\mathbf{x}}_k + \mathbf{b}_k. \quad (4.26)$$

Mas, agora, a não linearidade pode apresentar-se na matriz evolução. Com isso a Eq.(4.26) toma a forma

$$\mathbf{x}_{k+1} = \Psi_k(\mathbf{x}_k) + \mathbf{e}_k, \quad (4.27)$$

ou nas variáveis de estado, com o que se tem

$$\mathbf{Z}_k = \mathbf{G}(\mathbf{x}_k) + \mathbf{n}_k, \quad (4.28)$$

ou para ambas juntas, onde Ψ e \mathbf{G} são funções do vetor de parâmetros \mathbf{X}_k e do tempo t_k representado por k .

Suponha que depois da medição \mathbf{Z}_k , é possível representar o estado de conhecimento de \mathbf{X}_k por uma distribuição gaussiana com média $\tilde{\mathbf{X}}_k$ e matriz covariância $[\tilde{P}_k]$. Assim, Ψ e \mathbf{G} são aproximados pelos seus polinômios de Taylor ao redor do valor $\tilde{\mathbf{X}}_k$ como a seguir

$$\Psi_k(\mathbf{x}_k) \approx \Psi_k(\tilde{\mathbf{x}}_k) + J\Psi_k(\tilde{\mathbf{x}}_k)(\mathbf{x}_k - \tilde{\mathbf{x}}_k), \quad (4.29)$$

e

$$\mathbf{G}_k(\mathbf{x}_k) \approx \mathbf{G}_k(\tilde{\mathbf{x}}_k) + J\mathbf{G}_k(\tilde{\mathbf{x}}_k)(\mathbf{x}_k - \tilde{\mathbf{x}}_k), \quad (4.30)$$

onde

$$J\Psi_k(\tilde{\mathbf{x}}_k) = \left[\frac{\partial \Psi_i}{\partial x_{kj}} \right], \quad (4.31)$$

e

$$J\mathbf{G}_k(\tilde{\mathbf{x}}_k) = \left[\frac{\partial G_i}{\partial x_{kj}} \right], \quad (4.32)$$

são os Jacobianos de Ψ e \mathbf{G} respectivamente.

Tomando o valor esperado da Eq.(4.27) tem-se o valor médio de \mathbf{X}_{k+1} como a seguir

$$\bar{\mathbf{X}}_{k+1} = \Psi_k(\bar{\mathbf{x}}_k), \quad (4.33)$$

já que o ruído \mathbf{e}_k é considerado como sendo um processo gaussiano com média nula e matriz de covariância $[S_k]$.

Calculando a matriz de covariância $[P_{k+1}]$ como sendo o valor esperado de $(\mathbf{X}_{k+1} - \bar{\mathbf{X}}_{k+1})(\mathbf{X}_{k+1} - \bar{\mathbf{X}}_{k+1})^T$, obtem-se

$$[P_{k+1}] = J\Psi_k(\tilde{\mathbf{x}}_k)[\tilde{P}_k]J\Psi_k(\tilde{\mathbf{x}}_k)^T + [S_k]. \quad (4.34)$$

Assim, se os dados medidos são dados pela Eq.(4.28), os passos para o filtro de Kalman extendido podem ser reescritos como a seguir:

1. Calcula-se a matriz de ganho de Kalman a partir de $[P_k]$

$$[K_k] = [P_k] J\mathbf{G}_k(\tilde{\mathbf{x}}_k)^T (J\mathbf{G}_k(\tilde{\mathbf{x}}_k) [P_k] J\mathbf{G}_k(\tilde{\mathbf{x}}_k)^T + [\Gamma_k])^{-1}.$$

2. A covariância da PDF priori de \mathbf{X}_k

$$[\tilde{P}_k] = ([I] - J\mathbf{G}_k(\tilde{\mathbf{x}}_k) [K_k]) [P_k].$$

3. A média da PDF posterior de \mathbf{X}_k , é dizer, o estado do filtro de Kalman

$$\tilde{\mathbf{X}}_k = \bar{\mathbf{x}}_k + [K_k] (\mathbf{Z}_k - \mathbf{G}_k(\tilde{\mathbf{x}}_k)).$$

4. A covariância da PDF priori de \mathbf{X}_{k+1}

$$[P_{k+1}] = J\Psi_k(\tilde{\mathbf{x}}_k)[\tilde{P}_k]J\Psi_k(\tilde{\mathbf{x}}_k)^T + [S_k],$$

5. A média da PDF priori de \mathbf{X}_{k+1} , é dizer, a predição de Kalman do novo estado

$$\bar{\mathbf{X}}_{k+1} = \Psi_k(\tilde{\mathbf{x}}_k).$$

A seguir, é resolvido um exemplo com o fim de dar um maior entendimento do filtro de Kalman extendido.

4.7

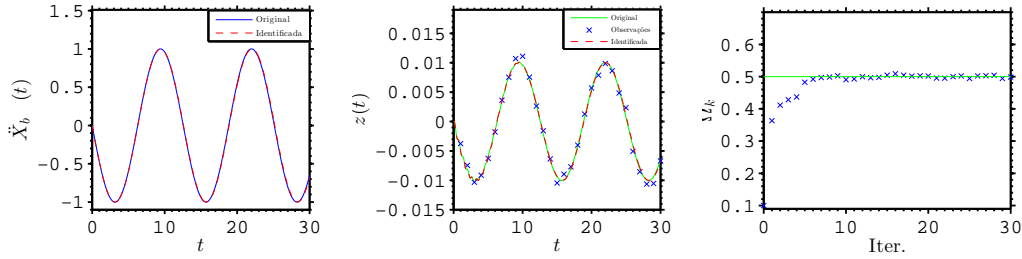
Estimação não linear com ruído no modelo

Agora, considera-se o exemplo do sistema com 1 grau de liberdade em que se deseja estimar o parâmetro não linear. Com isso assume-se que A_b é conhecido e Ω_b será estimado.

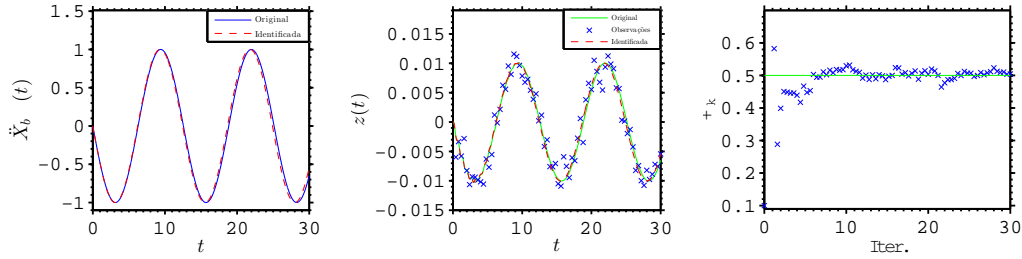
Supondo que os erros no modelo e_k e nas medições n_k , para cada passo de tempo t_k , são de tipo gaussiano com médias nulas e matrizes de covariância $[S_k] = R^2 [I]_{1 \times 1} = R^2$ e $[\Gamma_k] = \sigma_n^2 [I]_{1 \times 1} = \sigma_n^2$ respectivamente, tem-se o seguinte problema a ser resolvido

$$\begin{cases} \Omega_{b_{k+1}} = \Psi_k(\Omega_{b_k}) + e_k, \\ Z_k = \hat{z}_k + n_k = G_k(\Omega_{b_k}) + n_k. \end{cases} \quad (4.35)$$

$$R = 0,01, \sigma_n = 0,001 \text{ e } N = 31$$



$$R = 0,02, \sigma_n = 0,002 \text{ e } N = 76$$



$$R = 0,025, \sigma_n = 0,003 \text{ e } N = 151$$

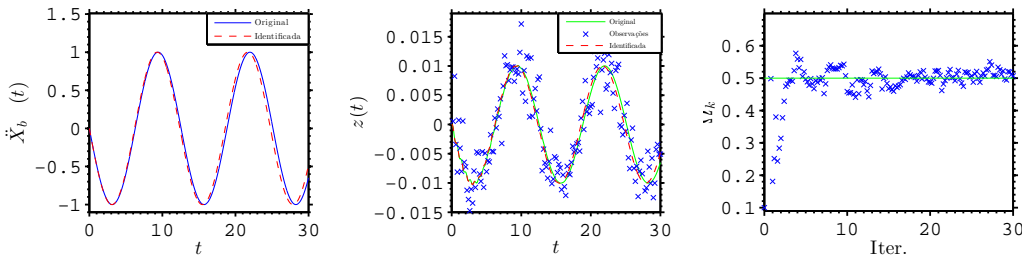


Figura 4.10: Curvas do forçamento, de solução e de convergência.

A Figura 4.10 apresenta as curvas de convergência para diferentes situações. Assim, sendo o chute inicial fixado em $\Omega_{b_0} = 0,1$, apresentam-se os casos limite, para quando o método é estável na convergência, para diferentes configurações de R , σ_n e N .

Essas curvas mostram um comportamento bom na hora de estimar o parâmetro quando este é um valor constante no tempo. Mas o método começa ter instabilidade e pode ter perda de convergência ao se aumentar o número de parâmetros que se quer estimar.

Um forte causativo deste comportamento é o fato de que no processo de linearização é truncada a serie de Taylor na derivada primeira. Pelo que, para ter uma acurácia melhor, tem-se que incluir mais termos da serie [7].

4.8

Filtro de Kalman estendido com correção de segunda ordem

O filtro de Kalman linearizado estendido surge da serie de Taylor truncada de primeira ordem nas expansões de $G(\mathbf{x})$ e $\Psi(\mathbf{x})$. Uma forma de aumentar a acurácia do modelo é mantendo um maior número de termos das expansões [7].

No caso de ter a serie de Taylor truncada na segunda ordem para aproximação das medições tem-se para cada componente

$$Z_i \approx J_i(\mathbf{G}(\hat{\mathbf{x}}))(\mathbf{x} - \hat{\mathbf{x}}) + \frac{1}{2}(\mathbf{x} - \hat{\mathbf{x}})^T H_i \mathbf{G}(\hat{\mathbf{x}})(\mathbf{x} - \hat{\mathbf{x}}) + \mathbf{e} \quad (4.36)$$

onde $J_i \mathbf{G}(\hat{\mathbf{x}})$ e $H_i \mathbf{G}(\hat{\mathbf{x}})$ são linhas das matrizes jacobiana e hessiana respectivamente.

Tomando o valor esperado da Eq.(4.36), e definindo o error de estimação como $\tilde{\mathbf{X}} = \bar{\mathbf{X}} - \mathbf{X}$, pode-se mostrar que

$$\bar{Z}_i = G_i(\tilde{\mathbf{x}}) + \varrho_i, \quad (4.37)$$

onde a correção do erro quadrático é dada por

$$\varrho_i = \text{tr}(H_i \mathbf{G} P(\tilde{\mathbf{x}})), \quad (4.38)$$

e $P(\tilde{\mathbf{x}}) = E\langle \tilde{\mathbf{x}} \tilde{\mathbf{x}}^T \rangle$ é a matriz de covariância a priori da estimação. Assim, o error quadrático ϱ_i pode ser adicionado ao filtro de Kalman estendido (FKE) para melhorar a acurácia

$$\hat{Z}_i = \bar{Z}_i = G_i(\tilde{\mathbf{x}}) + \varrho_i. \quad (4.39)$$

4.9

Estimação não linear com correção de segunda ordem

Considere agora como exemplo o problema massa-mola-amortecedor com dois graus de liberdade apresentado na seção 2.3.

Assim, o problema a ser resolvido é

$$\begin{cases} \Omega_{b_{k+1}} = \Psi_k(\Omega_{b_k}) + e_k, \\ \mathbf{Z}_k = \hat{\mathbf{z}}_k + \mathbf{n}_k = \mathbf{G}(\tilde{\mathbf{x}}) + \boldsymbol{\varrho} + \mathbf{n}_k, \end{cases} \quad (4.40)$$

onde $\mathbf{G}(\tilde{\mathbf{x}}) = \begin{pmatrix} G_1(\tilde{\mathbf{x}}) & G_2(\tilde{\mathbf{x}}) \end{pmatrix}^T$ e $\boldsymbol{\varrho} = \begin{pmatrix} \varrho_1 & \varrho_2 \end{pmatrix}^T$.

Então, supondo que $\Omega_b = X_{nl}$ é o parâmetro não linear que não se conhece, procura-se estimá-lo a partir de observação dos deslocamentos. Para isso são geradas N amostras de cada posição de cada massa, através de uma rotina de MATLAB que dá os valores dos deslocamentos $\mathbf{Z} = (Z_1, Z_2)^T$ de forma analítica, e é adicionado um ruído de forma aleatória. O mesmo ruído é considerado gaussiano identicamente distribuído com média nula e matriz de covariância $[\Gamma] = \sigma_n^2 [I]_{2 \times 2}$. O cálculo dos erros quadráticos ϱ_i , as matrizes hessianas, também é feito via MATLAB.

Sendo $\mathbf{Z}_k = \mathbf{Z}(t_k)$ a observação do vetor de deslocamentos \mathbf{Z} ao tempo t_k , pode-se escrever cada medição como

$$\begin{aligned} \mathbf{Z}_k &= \hat{\mathbf{z}}_k + \mathbf{n}_k, \\ &= \mathbf{C}(\Omega_b, t_k) \mathbf{A}_b + \mathbf{n}_k. \end{aligned} \quad (4.41)$$

Nesse caso a matriz de evolução é $\Psi_k = \Psi(t_{k+1}, \Omega_b) / \Psi(t_k, \Omega_b) = 1$.

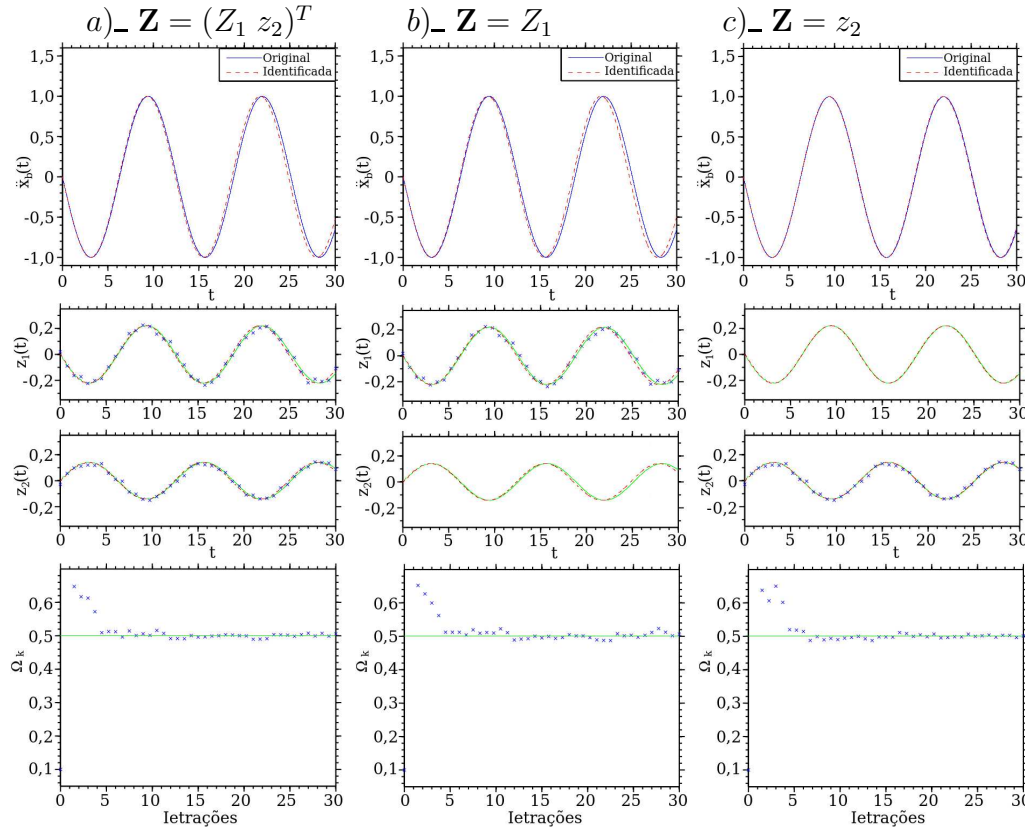


Figura 4.11: Curvas do forçamento, de solução e de convergência.

A seguir apresenta-se as curvas para o caso em que se quer estimar o

parâmetro Ω_b ao observar os deslocamentos em ambas massas, e em cada uma por separado, é dizer, quando são observados os deslocamentos Z_1 e Z_2 juntos e separados.

A Figura 4.11 apresenta as curvas de convergência, do forçamento e dos deslocamentos, para diferentes situações. Sendo $\Omega_{b_0} = 0, 1$ o chute inicial para o parâmetro, $R = 0,015$, $\sigma_n = 0,015$ e $N = 41$, apresentam-se as estimações para quando as observações são a) $\mathbf{Z} = (Z_1 \ Z_2)^T$, b) $\mathbf{Z} = Z_1$ e c) $\mathbf{Z} = z_2$.

Pode-se perceber que o filtro de Kalman estendido com erro quadrático funciona muito bem para qualquer uma das situações tratadas.

5

Estimação de parâmetros de uma bancada experimental

Neste capítulo aborda-se o estudo de caracterização de uma bancada experimental desenhada para um estudo de uma coluna de perfuração em escala reduzida. O estudo da mesma será abordado tanto pelo método de mínimos quadrados quanto pelos métodos filtro de Kalman e filtro de Kalman estendido, com o fim de estimar os parâmetros envolvidos na modelagem desta bancada.

5.1

Descrição da bancada experimental

As perfurações de poços são de grande interesse na indústria do petróleo. Nos últimos anos, muitas pesquisas sobre vibrações em colunas de perfuração têm sido publicadas, a exemplos [2, 12].

A realização de experimentos da dinâmica de colunas de perfuração em escala reduzida, possibilita um melhor entendimento dos fenômenos envolvidos. Para tanto, ao levar a cabo estudos experimentais, são requeridos coleções de dados a fim de estimar parâmetros característicos do comportamento dinâmico.

Neste capítulo, usa-se a modelagem de três graus de liberdade para uma coluna de perfuração em escala reduzida. Desta forma, na seção seguinte, apresenta-se as equações de movimento torcional da bancada experimental com intuito de validar os modelos de estimação apresentados neste trabalho.

Uma melhor descrição da bancada experimental encontra-se na referência [2].

5.2

Modelagem da bancada experimental

A modelagem consiste de um motor de corrente contínua acoplado a um eixo de baixa rigidez (que simula a coluna de perfuração) responsável pela transmissão da dinâmica torcional entre os graus de liberdade. A inercia do motor é concentrada no rotor r_2 , o qual é acoplado ao extremo do eixo-coluna, sendo este o segundo grau de liberdade. O rotor r_1 é posicionado no outro extremo da coluna, sendo considerado como o primeiro grau de liberdade. A representação de uma broca de perfuração, com atrito e inercia, é associada ao

rotor r_1 . A inércia do eixo-coluna é desprezível em comparação com as outras inércias envolvidas na dinâmica torcional do sistema. Ambos rotores giram em torno a seu centro geométrico e são desprezados os movimentos laterais. O amortecimento estrutural do material do eixo-coluna (amortecimento histerético) é desprezado. As vibrações torcionais são induzidas por atrito seco no primeiro grau de liberdade t_1 . Por fim, o eixo do motor é acoplado ao eixo-coluna por meio de um sistema de redução (de fator de $i_m = 8 : 1$). Na Figura 5.1 é ilustrado um esquema representativo da bancada experimental a ser modelada.

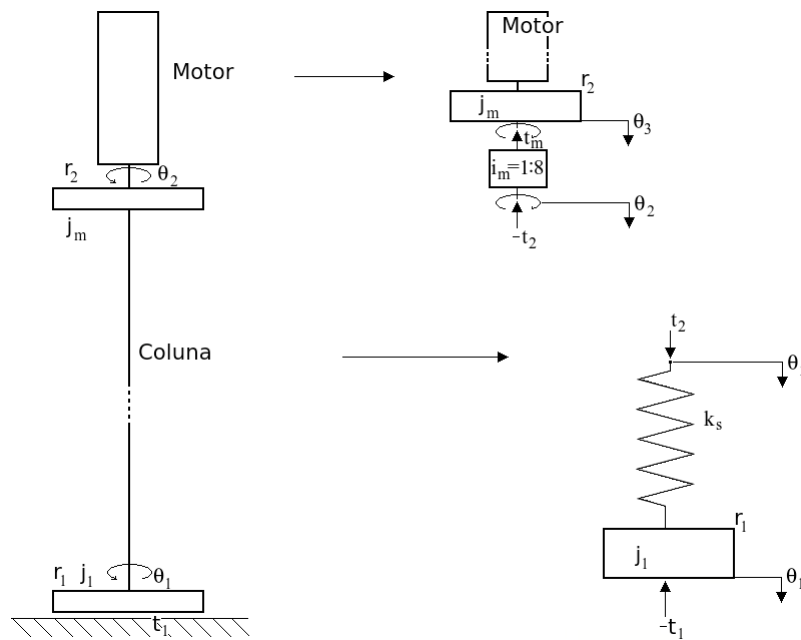


Figura 5.1: Esquema da bancada experimental, horizontal, como vista de cima.

Para o bom funcionamento do modelo preditivo, como em toda modelagem, é requerido o conhecimento dos parâmetros envolvidos. Porém, em algumas ocasiões, por diferentes razões, pode-se não ter o conhecimento ou ter informações parciais sobre alguns parâmetros. Neste casos, a Tabela 5.1 apresenta os parâmetros conhecidos e desconhecidos.

Constantes da coluna			
Parâmetros	Descrição	Valores	Unidades
j_1	Momento de inercia	$28,3 \times 10^{-3}$	$kg\ m^2$
l_s	Comprimento	2,4	m
D_s	Diâmetro	$3,0 \times 10^{-3}$	m
m_{r_1}	Massa do rotor 1	6,4	kg
k_s	Rigidez torcional	$2,63 \times 10^{-1}$	Nm/rad
r_{r_1}	Raio do rotor 1	$18,8 \times 10^{-2}$	m
t_1	Torque de atrito no rotor 1	—	m
Constantes do motor			
Parâmetros	Descrição	Valores	Unidades
j_m	Momento de inercia	$3,67 \times 10^{-4}$	$kg\ m^2$
l_m	Indutância interna	$1,1 \times 10^{-3}$	H
r_m	Resistência interna	0,33	Ω
k_T	Constante de torque	0,12	Nm/A
k_e	Constante de voltagem	$601,606 \times 10^{-4}$	$V/(rad/s)$
t_f	Torque por atrito interno	0,1	Nm
c_m	Amortecimento interno	$1,784 \times 10^{-4}$	$Nm/(rad/s)$
i_m	Fator redução engrenagem	1/8	8 : 1
v	Voltagem nominal	42	V

Tabela 5.1: Parâmetros da bancada experimental.

Começando pelo equacionamento do motor, o que é dado pela equação de queda de potenciais,

$$l_m \dot{e}_{(t)} + r_m e_{(t)} + k_e \dot{\theta}_3 = v, \quad (5.1)$$

e pela equação de torques, com acoplamento do eixo do motor ao eixo-coluna por meio do sistema de redução. Assim, tendo em conta nesta modelagem a inércia do motor j_m concentrada no rotor r_2 [2],

$$j_m \ddot{\theta}_3 + \dot{\theta}_3 c_m - k_T e_{(t)} + t_f = -t_m, \quad (5.2)$$

onde $e_{(t)}$ é a corrente do motor e θ_3 é o angulo do eixo do motor, e t_m é o torque torque no motor dado pelo acoplamento entre o motor e a redução.

O angulo do eixo do motor θ_3 é relacionado com o angulo do eixo-coluna θ_2 pelo fator de redução i_m como a seguir

$$i_m \theta_3 = \theta_2 . \quad (5.3)$$

O eixo-coluna transmite, pela sua dinâmica torcional, o torque da saída da caixa de redução (θ_2 - segundo grau de liberdade) até o rotor r_1 , de inércia j_1 , posicionado no extremo onde é definido o primeiro grau de liberdade θ_1 ,

$$j_1 \ddot{\theta}_1 + k_s(\theta_1 - \theta_2) = -t_1 , \quad (5.4)$$

e

$$-k_s(\theta_1 - \theta_2) = t_2 , \quad (5.5)$$

onde t_2 é o torque aplicado no eixo-coluna, no extremo do segundo grau de liberdade, dado pelo acoplamento com a redução do motor.

O torque no motor t_m e o torque no segundo grau de liberdade t_2 são relacionados pelo fator de redução,

$$t_m = i_m t_2 . \quad (5.6)$$

Das Eqs.(5.2), (5.3), (5.6) e (5.5) tem-se

$$j_m \ddot{\theta}_2 - i_m^2 k_s(\theta_1 - \theta_2) - i_m k_T e(t) + \dot{\theta}_2 c_m = -i_m t_f . \quad (5.7)$$

Pelo fato de que a corrente é a derivada temporal da carga elétrica, é conveniente usa-la como a seguir

$$\dot{q}(t) = \frac{dq(t)}{dt} = e(t) . \quad (5.8)$$

Desta forma, pelas Eqs.(5.1), (5.4), (5.7) e (5.8), pode-se escrever a dinâmica da bancada experimental como a seguir

$$\begin{aligned} \begin{pmatrix} j_1 & 0 & 0 \\ 0 & j_m & 0 \\ 0 & 0 & l_m \end{pmatrix} \begin{pmatrix} \ddot{\theta}_1 \\ \ddot{\theta}_2 \\ \ddot{q}(t) \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & c_m & -i_m k_T \\ 0 & k_e/i_m & r_m \end{pmatrix} \begin{pmatrix} \dot{\theta}_1 \\ \dot{\theta}_2 \\ \dot{q}(t) \end{pmatrix} \\ + \begin{pmatrix} k_s & -k_s & 0 \\ -i_m^2 k_s & i_m^2 k_s & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \theta_1 \\ \theta_2 \\ q(t) \end{pmatrix} = \begin{pmatrix} -t_1 \\ -i_m t_f \\ v \end{pmatrix} , \end{aligned} \quad (5.9)$$

onde, o sistema é da forma

$$[j] \ddot{\mathbf{z}} + [c] \dot{\mathbf{z}} + [k] \mathbf{z} = \mathbf{f} . \quad (5.10)$$

E logo, convenientemente, pode ser escrito no espaço de estados, como a seguir

$$\dot{\mathbf{y}} = [A] \mathbf{y} + \mathbf{F} , \quad (5.11)$$

onde

$$\mathbf{y} = \begin{pmatrix} \theta_1 \\ \theta_2 \\ q \\ \dot{\theta}_1 \\ \dot{\theta}_2 \\ \dot{q}(t) \end{pmatrix}, \quad \mathbf{F} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ -t_1/j_1 \\ -i_m t_f/j_m \\ v/l_m \end{pmatrix}, \quad (5.12)$$

$$[A] = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ -k_s/j_1 & k_s/j_1 & 0 & 0 & 0 & 0 \\ i_m^2 k_s/j_1 & -k_s/j_m & 0 & 0 & -c_m/j_m & k_T/(i_m j_m) \\ 0 & 0 & 0 & 0 & -i_n k_e/l_m & -r_m/l_m \end{pmatrix}. \quad (5.13)$$

O sistema dado pela Eq.(5.11) é resolvido numericamente usando o programa de cálculos MatLab.

Na seções seguintes são estimados os parâmetros de amortecimento interno do motor c_m e a constante de voltagem k_e , utilizando o método dos mínimos quadrados, e o torque de atrito no rotor r_1 usando filtro de Kalman extendido. Para este último caso, é proposto um modelo de atrito correspondente com um problema real, achado em uma bancada experimental.

Com o intuito de fornecer estimações dos parâmetros desconhecidos, é desenvolvida a análise numérica da bancada junto com os métodos de mínimos quadrados (não recursivo) e com o filtro de Kalman extendido (recursivo).

5.3

Estimação por mínimos quadrados

Primeiramente o método de mínimos quadrados é usado para estimar tanto o amortecimento interno do motor c_m como a constante de voltagem k_e . Logo, o vetor de parâmetros é

$$\mathbf{x} = \mathbf{x}_{nl} = \begin{pmatrix} c_m \\ k_e \end{pmatrix}. \quad (5.14)$$

Dado que no laboratório só é possível medir os valores das posições θ_1 e θ_2 e suas respectivas velocidades $\dot{\theta}_1$ e $\dot{\theta}_2$, se terá o vetor das amostras

$$\boldsymbol{\eta} = \begin{pmatrix} \theta_1 & \theta_2 & \dot{\theta}_1 & \dot{\theta}_2 \end{pmatrix}^T = \hat{\boldsymbol{\eta}} + \mathbf{n}, \quad (5.15)$$

onde $\hat{\boldsymbol{\eta}}$ é dado pela resolução numérica da Eq.(5.11), e \mathbf{n} é o ruído nas medições. Como sempre, esse ruído é suposto como sendo um processo estocástico independente identicamente distribuido com densidade de probabilidade de tipo gaussiana, com média nula e matriz de covariância $[\Gamma] = \sigma_n^2 [I]$.

Assim, a função de desfasagem, segundo a Eq.(3.25), é dada por

$$\varepsilon(\mathbf{x}; \boldsymbol{\eta}) = \left(\frac{1}{\sigma_n^2} \right) (\boldsymbol{\eta} - \hat{\boldsymbol{\eta}}(\mathbf{x}))^t (\boldsymbol{\eta} - \hat{\boldsymbol{\eta}}(\mathbf{x})) . \quad (5.16)$$

O mínimo da função dada na Eq. 5.16 é procurado com a função FMINSEARCH do MATLAB.

Para verificar a convergência do modelo são usados os valores dos parâmetros da Tabela 5.2, juntamente com a solução do modelo dada pela Eq.(5.11) à qual é somada um erro de caracter aleatório. Usa-se este resultado para estimar o valor original de tal forma que possa-se testar a convergência do processo de estimação. Assim, uma vez que seja assegurada a estimação por mínimos quadrados, serão usados valores experimentais a fim de melhorar o estado do conhecimento dos parâmetros da bancada.

O experimento é realizado sem considerar o atrito no rotor r_1 ($t_1 = 0$) e dando uma entrada de voltagem de $v = 8 \text{ Volts}$. O desvio padrão das amostras é variado segundo mostrado na Tabela 5.2.

Assim, a Figura 5.2 apresenta um dos casos expostos na Tabela 5.2, para as curvas dos valores de θ_1 , θ_2 , $q(t)$, $\dot{\theta}_1$, $\dot{\theta}_2$, e $\dot{q}(t) = e(t)$. Denotando as medições pelas cruzes vermelhas, as curvas reconstruídas com cor verde e as originais com a cor azul. Pode-se observar uma boa convergência do método de mínimos quadrados.

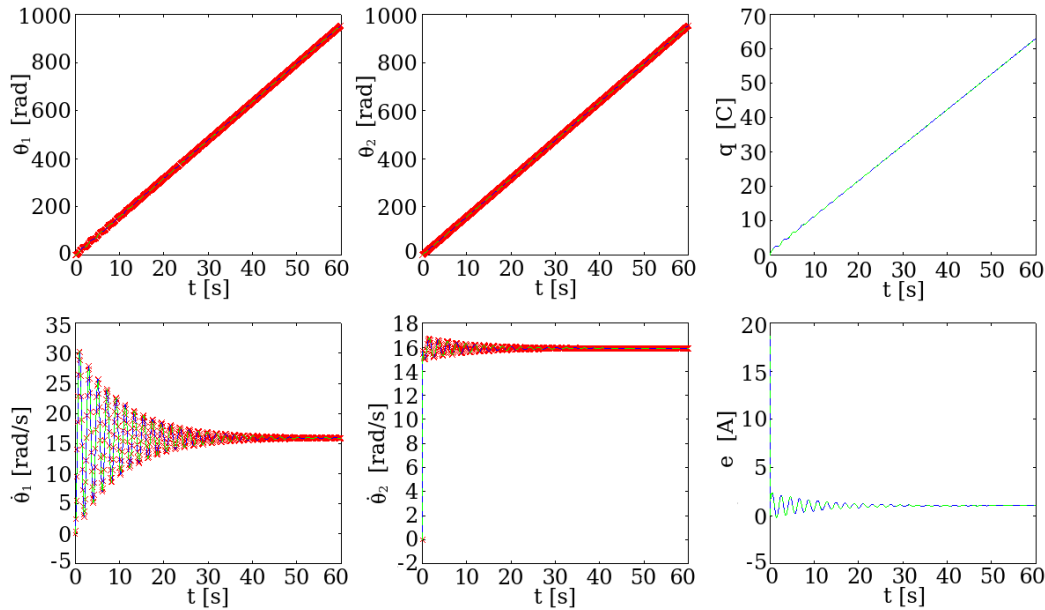


Figura 5.2: Curvas reconstruídas e medições para $\sigma_n = 0.01$

A curva de convergência do método é exibida na Figura 5.3.

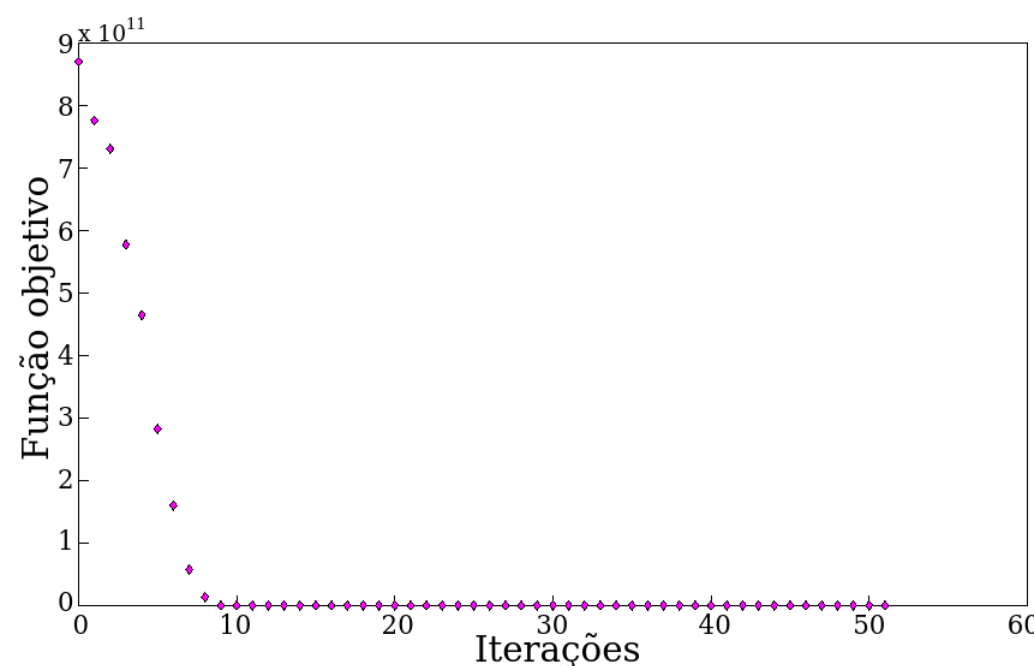


Figura 5.3: Convergência da função objetivo ($\varepsilon(\mathbf{x}; \boldsymbol{\eta})$) para $\sigma_n = 0.01$

Finalmente é construída a Tabela 5.2, na qual pode-se observar os erros de estimação para diferentes valores de desvio padrão. Assim tem-se uma ideia do bom comportamento para a estimação de parâmetros.

Valor	k_s	j_1
Analítico	19,099e5	601,606e4
Estimado com $\sigma_n = 0,001$	18,959e5	601,641e4
Error absoluto percentual	0,77 %	0,0064 %
Estimado para $\sigma_n = 0,01$	19,099e5	601,647e4
Error absoluto percentual	0,77 %	0,0068 %
Estimado para $\sigma_n = 0,1$	18,987e5	601,636e4
Error absoluto percentual	0,60 %	0,0051 %
Estimado para $\sigma_n = 1$	18987e5	601,642e4
Error absoluto percentual	0,71 %	0,0061 %
Estimado para $\sigma_n = 1$	18,943e5	601,883e4
Error absoluto percentual	0,82 %	0,046 %

Tabela 5.2: Parâmetros estimados com valores iniciais de $c_m = 0,001$ e $k_e = 0,01$

No seguinte é usado o método de mínimos quadrados para estimar o vetor de parâmetros dado na Eq.(5.17), como a seguir

$$\mathbf{x} = \begin{pmatrix} j_m & c_m & k_e & r_m & k_t & l_m & k_s & j_1 & t_f \end{pmatrix}^T. \quad (5.17)$$

O vetor de amostras é o mesmo que foi dado pela Eq.(5.15), só que agora é obtido de forma experimental. Porém, a função objetivo é a dada pela Eq.(5.16) com o vetor de parâmetros dado pela Eq.(5.17).

Realiza-se o experimento para um valor de voltagem de entrada de $v = 8$ [Volts] e com atrito $t_1 = 0$. Poderia-se fazer uma estimação por mínimos quadrados diretamente do vetor dado pela Eq.(5.17), mas existe uma limitação. Neste caso, o método simplex usado para minimizar a função objetivo converge rapidamente em mínimos locais ou diretamente não se consegue a convergência.

Para achar uma convergência com um número elevado de parâmetros foi criada, neste trabalho, uma heurística que minimiza a função objetivo para diferentes combinações de parâmetros, como apresentado no esquema da Figura 5.4. Desta forma, o método de mínimos quadrados é iterado primeramente para $\mathbf{x}_1 = (t_f \ j_m)^T$, de depois, sequencialmente, para $\mathbf{x}_2 = (j_m \ c_m)^T$, $\mathbf{x}_3 = (c_m \ k_e)^T$, $\mathbf{x}_4 = (k_e \ k_t)^T$, $\mathbf{x}_5 = (k_t \ r_m)^T$, $\mathbf{x}_6 = (r_m \ l_m)^T$, $\mathbf{x}_7 = (l_m \ k_s)^T$, $\mathbf{x}_8 = (k_s \ j_1)^T$, $\mathbf{x}_9 = (j_1 \ t_f)^T$. Para cada \mathbf{x}_i (com $i = 1, \dots, 9$) faz-se 10 iterações e repete-se todo o ciclo até o valor da função objetivo não mudar. Apesar da natureza heurística do algoritmo, os resultados a seguir mostram boa convergência.

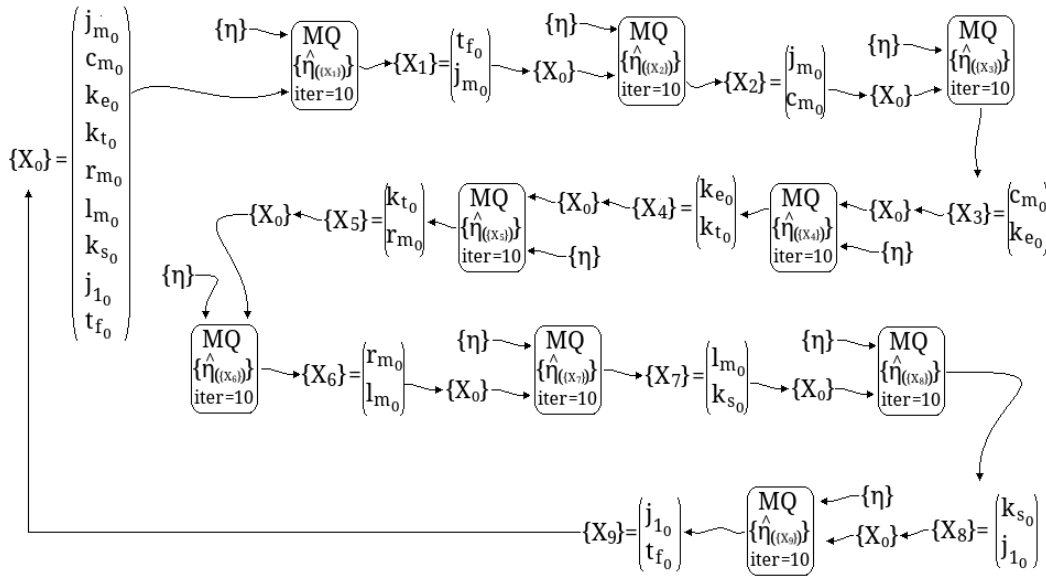


Figura 5.4: Heurística para minimizar a função objetivo.

A Figura 5.5 mostra as curvas para os valores medidos (azul) e as curvas reconstruídas (vermelho).

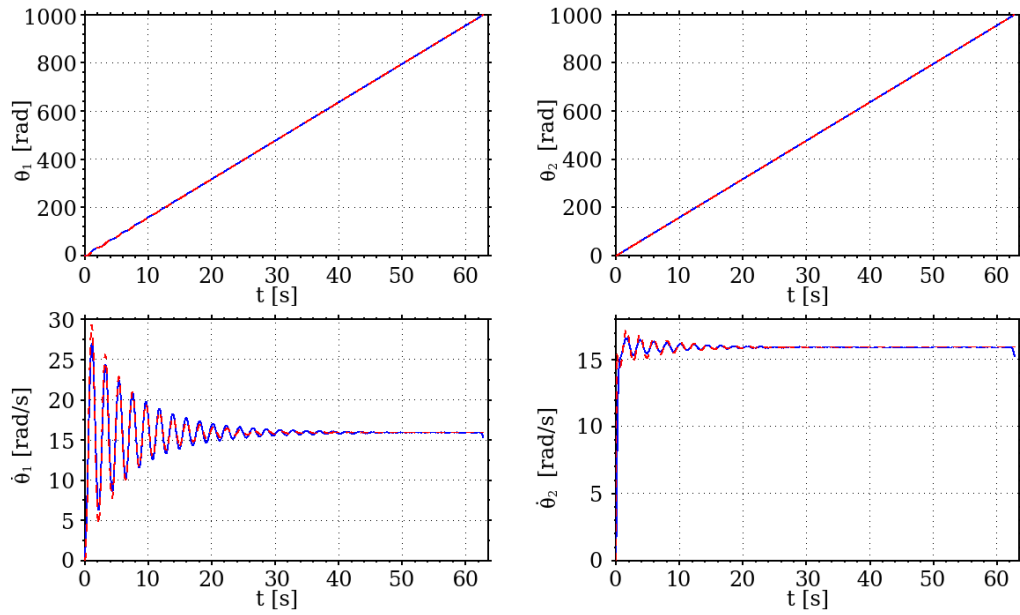


Figura 5.5: Valores medidos e curvas reconstruídas.

A curva de convergência é dada na Figura 5.6. A construção da mesma é dada pela continuação de cada curva de convergência das iterações do método descrito no esquema da Figura 5.4. Pode-se observar que o método converge para aproximadamente 400 iterações.

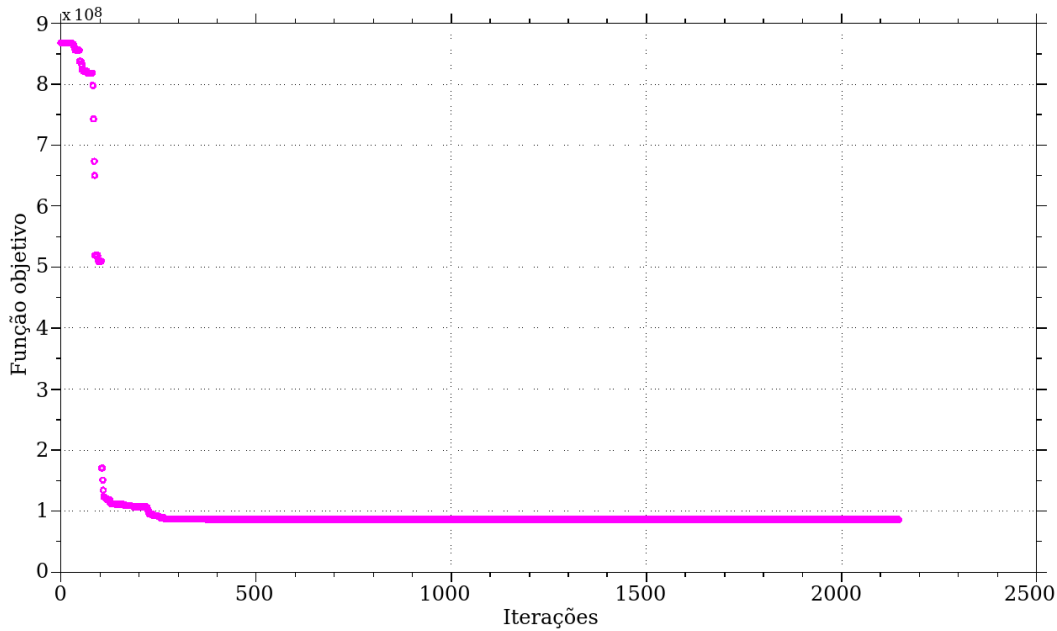


Figura 5.6: Convergência.

Os valores estimados são apresentados na Tabela 5.3 onde o erro relativo é com respeito aos valores apresentados na Tabela 5.1.

Parâmetros	Chute inicial	Estimação	error relativo (%)
\dot{j}_m	$3,67 \times 10^{-4}$	$4,698 \times 10^{-4}$	28 %
c_m	$1,91 \times 10^{-04}$	$1,78 \times 10^{-04}$	6,6 %
k_e	$601,606 \times 10^{-04}$	$587,723 \times 10^{-04}$	2,3 %
k_T	0,12	0,13	5 %
r_m	0,33	0,53	60 %
l_m	$1,10 \times 10^{-3}$	$0,84 \times 10^{-3}$	23,3 %
k_s	0,263233	0,2710	3 %
\dot{j}_1	$28,28 \times 10^{-3}$	$29,8 \times 10^{-3}$	5,4 %
t_f	0,1	0,1	0,0 %

Tabela 5.3: Parâmetros estimados

5.4

Estimação por filtro de Kalman estendido

Um valor do torque $\tau_1 = \tau_1(\dot{\theta}_1)$ simples é dado pelo atrito simulado no extremos r_1 e pelo modelo de Coulomb modificado, como a seguir

$$\tau_1(\dot{\theta}_1) = \begin{cases} \tau_{est} & \text{se } \dot{\theta}_1 < tol \\ \tau_{din} & \text{se } \dot{\theta}_1 \geq tol \end{cases} \quad (5.18)$$

onde tol é uma tolerância que descreve o limite entre os atritos estático e dinâmico.

Mas, a bancada experimental apresenta uma descentralização na direção transversal ao eixo de rotação, vide Figura 5.7. Assim, é importante caracterizá-la para melhorar o entendimento do problema geral.

A seguir, são testadas duas modelagens que descrevem o possível efeito produzido no torque τ_1 pela descentralização do sistema de atrito.

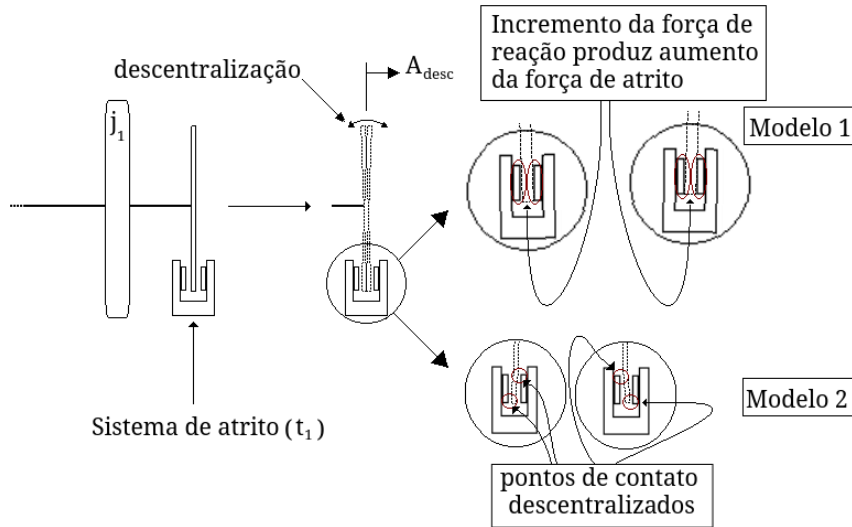
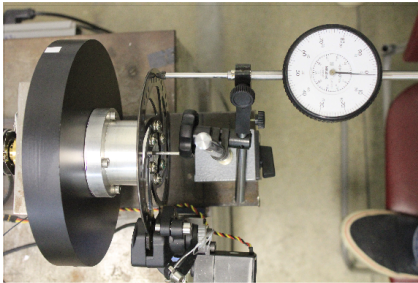


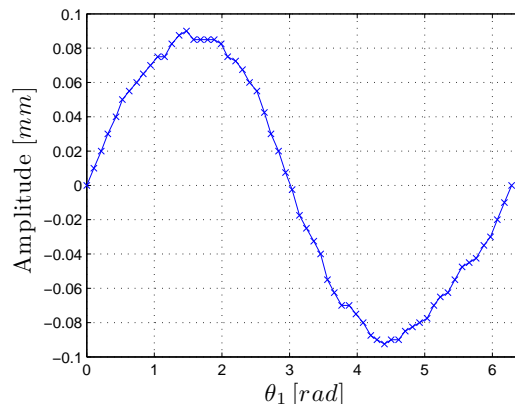
Figura 5.7: Descentralização do sistema de atrito posicionado em r_1 .

A Figura 5.7 apresenta um esquema de modelagens do atrito τ_1 . Estas ideias tem origem na observação da dinâmica da bancada. Ao ter uma descentralização como indicado na Figura 5.7, é de se suspeitar que ao variar o atrito dado pelo contato entre as superfícies, o torque fornecido por esta força de atrito também pode mudar. Assim, tanto o *modelo 1* como o *modelo 2* tem origem na variação da força de atrito devido à descentralização dita. Esta mudança depende em alguma proporção ($n_{\dot{\theta}_1}$) da rotação θ_1 do disco r_1 .

A Figura 5.8 apresenta a curva de amplitude da descentralização medida sobre a bancada, ao longo de uma volta (360°), mostrando que para um ponto fixo, a amplitude da descentralização cumpre um ciclo de a amplitude máxima $0.09 [mm]$ aproximadamente. Observa-se que a cada volta dada pelo disco r_1 , a amplitude A_{desc} do disco de atrito completa um ciclo.



5.8(a): Sistema de medição.



5.8(b): Amplitude para um ciclo do rotor.

Figura 5.8: Medição da descentralização do dispositivo de atrito.

Para a primeira modelagem (*modelo 1* na Figura 5.7), o aumento do deslocamento do disco entre as pastilhas de freio produz um aumento da força de contato entre estes, o que induz um aumento no torque de atrito. A seguir a descrição matemática proposta para o *modelo 1*,

$$\tau_{1_{mod1}}(\dot{\theta}_1 \geq tol) = \tau_{din} + A_{\tau_{din}mod1} \left(1 - \cos(n_{\dot{\theta}_1} \dot{\theta}_1 t) \right), \quad (5.19)$$

onde τ_{din} é o torque sem o feito da descentralização e $A_{\tau_{din}}$ é a amplitude de oscilação do torque médio induzida pela descentralização.

Para a segunda modelagem (*modelo 2* na Figura 5.7), o aumento do deslocamento do disco entre as pastilhas de freio produz uma diminuição da superfície de contato entre estes, o que induz uma diminuição da força normal de contato, tendo assim, uma diminuição do torque de atrito. A seguir a descrição matemática proposta para o *modelo 2*,

$$\tau_{1_{mod2}}(\dot{\theta}_1 \geq tol) = \tau_{din} - A_{\tau_{din}mod2} \left(1 - \cos(n_{\dot{\theta}_1} \dot{\theta}_1 t) \right). \quad (5.20)$$

Na Figura 5.9 pode ser visto um esquema representativo do comportamento das modelagens. Observa-se que o valor de proporção de frequência é $n_{\dot{\theta}_1} = 2$.

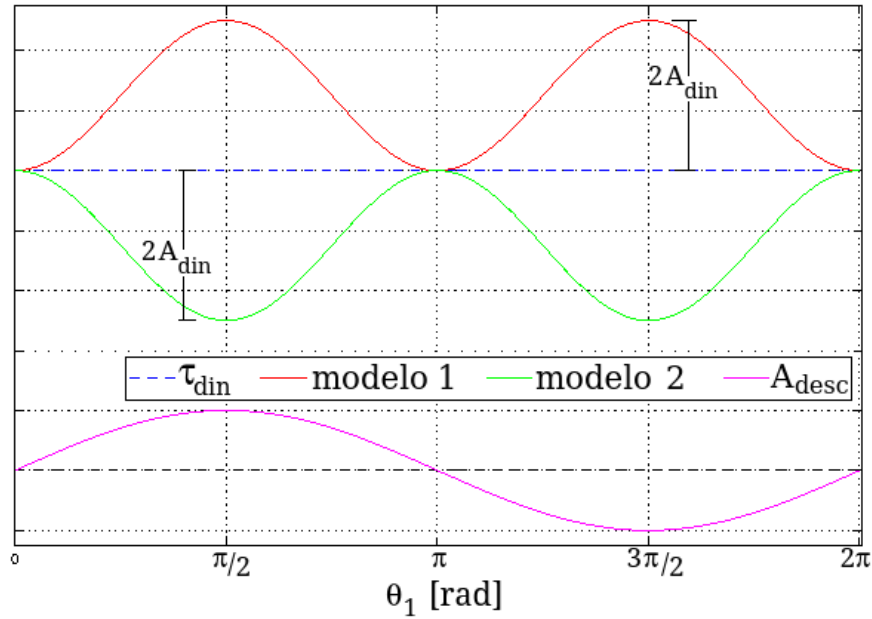


Figura 5.9: Esquemas das modelagens para o torque de atrito τ_1 .

No seguinte, usa-se o problema descrito pela Eq.(5.11) com o atrito dado pelas Eqs.(5.19) e (5.20). A voltagem de entrada é de $v = 8$ [Volts] como nos casos anteriores. Logo, quer-se estimar a amplitude de oscilação da descentralização $A_{\tau_{din}}$. O torque de atrito dinâmico τ_{din} é medido sobre a bancada e será considerados como conhecida.

Para comprovar o funcionamento da estimação é resolvido o sistema de forma numérica por uma rotina feita em código MATLAB usando ambas modelagens propostas. As amostras são geradas somando à solução numérica um erro do tipo gaussiano independente e identicamente distribuído com média nula e matriz covariância $[\Gamma] = \sigma_n^2 [I]_{1 \times 1}$ (onde σ_n é o desvio padrão).

Então, segundo o visto no capítulo 4, o problema a resolver para ambas modelagens, segundo o método de filtro de Kalman extendido, é

$$\begin{cases} A_{\tau_{din} k+1} = \Phi_K(A_{\tau_{din} k}) + e_k, \\ \boldsymbol{\eta}_k = \hat{\boldsymbol{\eta}}_k + \mathbf{n}_k = \mathbf{C}_k(A_{\tau_{din} k}) + \mathbf{n}_k, \\ \approx \mathbf{J}(\mathbf{C}_k(A_{\tau_{din} k})) A_{\tau_{din} k} + \mathbf{n}_k, \end{cases} \quad (5.21)$$

onde $\Phi_K = A_{\tau_{din} k}$ (dado que o parâmetro a estimar é constante), $\boldsymbol{\eta}$ é o dado pela Eq.(5.15) e $\mathbf{J}(\mathbf{C}_k(A_{\tau_{din} k}))$ é o jacobiano de $\mathbf{C}_k(A_{\tau_{din} k})$, que é calculado numericamente. As seguintes condições iniciais são usadas

$$\begin{aligned} A_{\tau_{din} k=0} &= 0,05, \\ [P_{k=0}] &= \frac{\sigma_n^2}{\|\mathbf{J}(\mathbf{C}_{k=0}(A_{\tau_{din} k=0}))\|^2}. \end{aligned} \quad (5.22)$$

A seguir serão apresentados gráficos que mostram o comportamento destas modelagens para os desvios padrões $\sigma_n = 0,05$ e $R = 0.02$ das amostras e do modelo respetivamente.

A Figura 5.10 se corresponde com o *metodo 1*. Esta apresenta as curvas da solução analítica (cor azul), as soluções aproximadas (cor vermelho) e os valores de medição gerados (cruzes de cor verde) para os deslocamentos angulares e as velocidades angulares nos dois graus de liberdade.

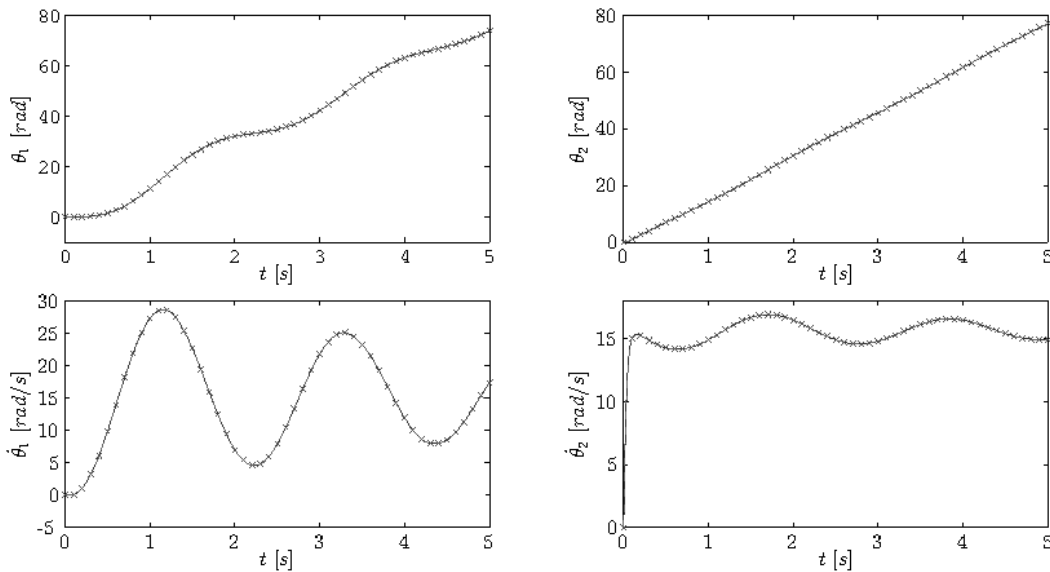


Figura 5.10: Observações, solução analítica e curvas reconstruídas para o *Modelo 1*.

A Figura 5.11 também se corresponde com o *metodo 1*. Nesta são apresentadas as curvas do torque analítica e reconstruída e a convergência para esta modelagem.

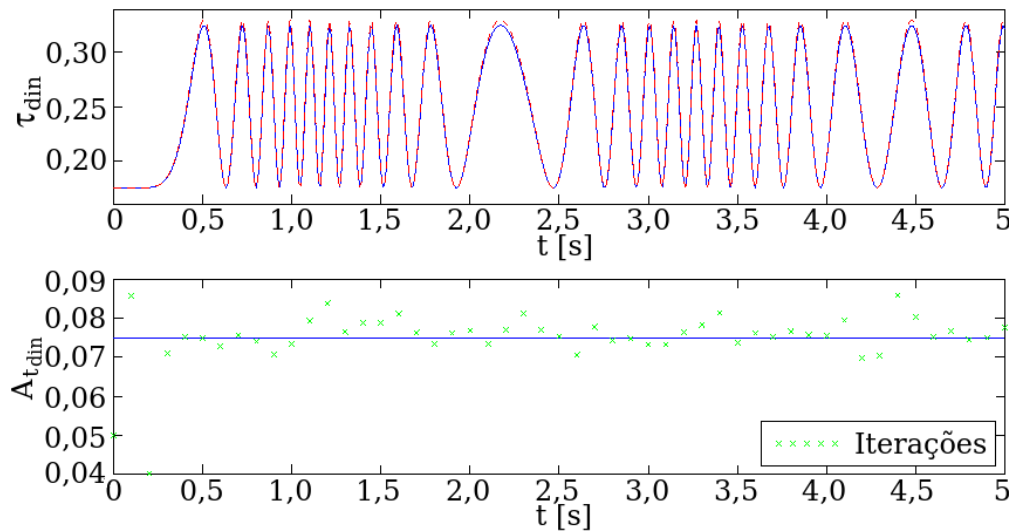


Figura 5.11: Torque e convergência para o *Modelo 1*.

Nas Figuras 5.10 e 5.11 pode-se observar o bom comportamento. Neste caso é possível perceber que a modelagem começa a perder amplitude na predição do torque.

A seguir o problema é resolvido para a segunda modelagem de τ_1 com as mesmas configurações. Assim, a Figura 5.12 (se corresponde com o *metodo 1*) apresenta as curvas da solução analítica (cor azul), as soluções aproximadas (cor vermelho) e os valores de medição gerados (cruzes de cor verde) para os deslocamentos angulares e as velocidades angulares nos dois graus de liberdade.

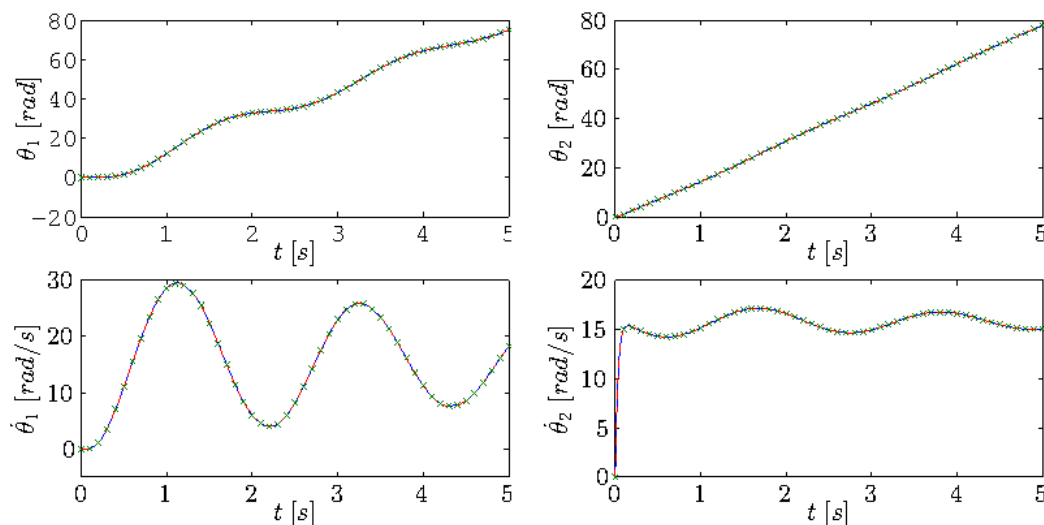


Figura 5.12: Observações, solução analítica e curvas reconstruídas para o *Modelo 2*.

E na Figura 5.13 são apresentadas as curvas do torque analítica e reconstruída e a convergência para o *metodo 2*.

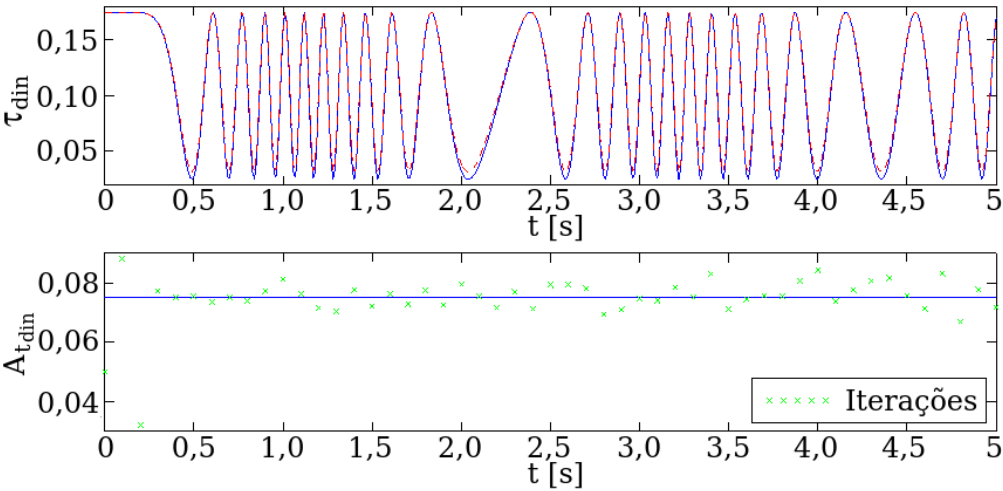


Figura 5.13: Torque e congengência para o *Modelo 2*.

Também, nas Figuras 5.12 e 5.13, observar-se o bom comportamento desta modelagem. E, igualmente nas Figuras 5.10 e 5.11, é possível perceber que a modelagem começa a perder amplitude na predição do torque.

Na Tabela 5.4 apresenta-se os valores das convergências para diferentes valores de desvío padrão. Assim, como feito para o método de mínimos quadrados, agora tem-se uma ideia do comportamento para a estimação do parâmetro e isto pode ser usado como medida de confiabilidade do método para caraterizações futuras da bancada experimental.

σ_n	<i>metodo 1</i>		<i>metodo 2</i>	
	$A_{t_{din}}$	<i>desvio relativo %</i>	$A_{t_{din}}$	<i>desvio relativo %</i>
0,0005	0,075	0,039 %	0,075	0,1 %
0,005	0,075	0,2 %	0,074	1,1 %
0,05	0,078	3,33 %	0,072	4,5 %
0,1	0,074	1,7 %	0,069	7,6 %
0,5	0,066	12,1 %	0,086	14,0 %
1	0,098	30,4 %	0,042	44,4 %

Tabela 5.4: Estimações por FKE para as modelagens *metodo 1* e *metodo 2*.

6

Conclusões

Neste capítulo são apresentadas observações gerais do trabalho, são dadas sugestões para possíveis trabalhos futuros e são apresentadas duas publicações.

6.1

Contribuições deste trabalho

A partir da inferência estatística bayesiana, foi definida a estimação paramétrica pelo método de mínimos quadrados. Foram resolvidos experimentos de um grau de liberdade para o caso em que os parâmetros foram tanto lineares quanto não lineares. A minimização da função objetivo foi feita pelo método de procura de mínimo conhecido como M-simplex, o qual foi dado pela função `fminsearch` no código de programação MATLAB. Demostrou-se que o método MQ é um bom estimador mas que a minimização da função objetivo além de ser lenta pode não convergir se as curvas a serem otimizadas não foram curvas suaves (ie., podendo cair em mínimos não globais).

Também usou-se o algoritmo recursivo de filtro de Kalman para estimação. Dado que o filtro de Kalman resolve só problemas lineares foi apresentada uma variação do mesmo, conhecida como filtro de Kalman estendido, a que consta da linearização das equações de evolução do FK pelo polinômio de Taylor e truncado este no primeiro grau. Devido a que esta aproximação apresenta uma rápida perda de acurácia na estimação de parâmetros, foi proposta uma forma de melhorar a precisão ao sumo o erro de truncamento da aproximação de Taylor à matriz jacobiana [7]. Vários exemplos foram resolvidos, para um e dois graus de liberdade e para parâmetros lineares com o FK e para não lineares com o FKE e com FKE com a correção ao segundo ordem. Também mostrou-se o comportamento do FK quando o parâmetro a ser estimado evolue no tempo, mostrando a importância das matrizes de evolução.

Em seguida foi abordada a modelagem de uma bancada experimental desenvolvida para estudos de uma coluna de perfuração em escala reduzida. A mesma constou da descrição da dinâmica do motor acoplado com a modelagem da coluna. As equações do modelo foram resolvidas de forma numérica usando os integradores numéricos `ode45` e `ode23t` do MATLAB. Estes resul-

tados foram utilizados para criar as amostras e poder testar os algoritmos de estimação. Estudando a dinâmica sem atrito no rotor, que simula a broca de perfuração, foi apresentada uma metodologia com a que comprovou-se a estimação por método de mínimos quadrados para os dados simulados. Logo foram utilizados dados experimental com o que foram obtidas estimações para 9 parâmetros envolvidos na dinâmica do sistema coluna-motor. Por último, foram construídas duas modelagen que para explicar situações observadas no dispositivo que fornece o torque de atrito na bancada experimental, de forma tal de aprofundá-los e contribuir na caracterização da dinâmica bancada experimental que tem-se desenvolvida no Laboratório de Vibrações Mecânicas (para maior informação desta bancada, o leitor pode-se dirigir à referência [2]). Para esta última caracterização foi usado o método de filtro de Kalman estendido e comprovou-se a sua estimação, mas por enquanto só pela simulação.

6.2

Observações deste trabalho

Uma rápida observação pode ser que ambos métodos conseguem a estimação com exatidão, a diferença é que mínimos quadrados demora mais para estimar. Uma das causas disto é que a minimização da função objetivo nem sempre é uma tarefa simples.

O método de filtro de Kalman é rápido em comparação com o método de mínimos quadrados, para sistemas lineares. Para este método as complicações são inversões de matrizes e a perda de precisão ao linearizar o caso não linear.

No método de mínimos quadrados, para minimizar a função objetivo, usa-se método M-Simplex, dado pela função de MATLAB `fminsearch` [9]. Este método encontra dificuldades ao calcular mínimos globais quando a função objetivo não é uma curva suave, caso contrário consegue estimações aceitáveis, mesmo quando a dispersão das amostras é grande.

O método de filtro de Kalman é uma boa ferramenta na hora de estimar parâmetros lineares, mas perde precisão quando a dispersão das amostras aumenta. Na hora de estimar parâmetros não lineares o filtro de Kalman estendido é usado. Este é baseado em uma linearização a que é obtida pelo truncamento no termo de segundo ordem da representação em polinômio de Taylor da função a linearizar. Este truncamento produz uma perda de precisão que afeta a predição do método.

6.3

Sugestões para trabalhos futuros

Uma primeira sugestão é de trabalhar com diferentes técnicas de minimização da função objetivo para o método de mínimos quadrados. Como exemplo podem ser algoritmos de procura pelo gradiente ou o método de Newton. Outra sugestão é a de testar algoritmos recursivos para parâmetros não lineares que utilizem truncamentos do polinômio de Taylor de ordens maiores.

Em todos os casos, o ruído das amostras é do tipo Gaussiano, mas poderia ser utilizado um outro tipo de ruído, dependendo o problema a ser resolvido, de forma tal que dar a estimação paramétrica a partir da inferência estatística bayesiana com ruídos de disferentes distribuições.

Para a caracterização da bancada experimental, na caraterização do sistema de atrito, espera-se poder fazer as medições para poder testar as modelagens aqui propostas. Também sugere-se testar com outros tipos de atrito de base (aqui proposto o atrito de Coulomb).

6.4

Publicações

- Fernando S. Buezas, Marta B Rosales, Rubens Sampaio and Mario Sandoval *UNCERTAINTY QUANTIFICATION IN THE SIMPLIFIED PROBLEM OF CONTACT BETWEEN TWO ELASTIC BODIES*. Proceedings of the 1st International Symposium on Uncertainty Quantification and Stochastic Modeling. February 26th to March 2nd, 2012, Maresias, São Sebastião, SP, Brazil.
- Mario G. Sandoval, Americo Cunha Jr y Rubens Sampaio. *ESTIMACIÓN DE PARÁMETROS DE UM PROBLEMA INVERSO ESTO-CÁSTICO: COMPARACIÓN DE DOS MÉTODOS*. XX ENIEF, XX Congreso sobre Métodos Numéricos y sus Aplicaciones. ENIEF 2013, 18 de Noviembre de 2013, Mendoza, Argentina.

Referências Bibliográficas

- [1] BOLSTAD, W. M.. **Introduction to Bayesian Statistics**. Wiley-Interscience, 2nd edition, 2007.
- [2] CAYRES, B. C.. **Numerical and Experimental Analysis of the Nonlinear Torsional Dynamics of a Drilling System**. M.Sc. Thesis, Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, Brazil, 2013.
- [3] DA SILVA NETO, A. J.; MOURA NETO, F. D.. **Problemas Inversos: Conceitos Fundamentais e Aplicações**. EdUERJ, Rio de Janeiro, 2005.
- [4] DAMIEN, P.; DELLAPORTAS, P.; POLSON, N. G. ; STEPHENS, D. A.. **Bayesian Theory and Applications**. Oxford, 2013.
- [5] ERKKI-SOMERSALO, J.-P.-K. .. **Computational and Statistical Methods for Inverse Problems (Applied Mathematical Sciences)**. Berlin Heidelberg NewYork, Hong Kong London, Milan Paris Tokyo; Springer, June 20, 2004.
- [6] GORDON, C.; WEBB, D. L. ; WOLPERT, S.. **One cannot hear the shape of a drum**. American Mathematical Society, 27:134–138, 1992.
- [7] GREWAL, M. S.; ANDREWS, A. P.. **Kalman Filtering: Theory and Practice Using MATLAB**. Wiley-IEEE Press, 2008.
- [8] LUTES, L. D.; SARKANI, S.. **Random Vibrations: Analysis of Structural and Mechanical Systems**. Elsevier, 2004.
- [9] www.mathworks.com/help/matlab/ref/fminsearch.html.
- [10] NELDER, J. A.; MEAD, R.. **A simplex method for function minimization**. The Computer Journal, 7:308–313, 1965.
- [11] OLVER, P. J.; SHAKIBAN, C.. **Applied Mathematics Lecture Notes**. 2012.

- [12] RITTO, T.. **Numerical Analysis of the Nonlinear Dynamics of a Drill-string with Uncertainty Modeling**. PhD thesis, Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2010.
- [13] SIVIA, D. S.; SKILLING, J.. **Data Analysis: A Bayesian Tutorial**. Oxford University Press, 2nd edition, 2006.
- [14] STARK, H.; WOODS, J. W.. **Probability, Random Processes, and Estimation Theory for Engineers**. Prentice-Hall, 1979.
- [15] TAN, S.; FOX, C.. **Inverse Problems**. Lecture Notes, University of Auckland, 2008.
- [16] TARANTOLA, A.. **Inverse Problem Theory and Methods for Model Parameter Estimation**. SIAM, 2005.
- [17] VAN DER HEIJDEN, F.; DUIN, R.; DE RIDDER, D. ; TAX, D.. **Classification, Parameter Estimation, and State Estimation: An engineering Approach using Matlab**. John Wiley & Sons, Chichester, 2004.