

Identificação relevante ao controle preditivo - Método EN-PH de otimização numérica baseado em técnicas de regularização

Raphael de Medeiros Souto Maior Baltar

Dissertação de Mestrado submetida à Coordenação do Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal de Campina Grande como parte dos requisitos necessários para a obtenção do grau de Mestre em Ciências no Domínio da Engenharia Elétrica.

Área de Concentração: Instrumentação e Controle

Péricles Rezende Barros, Ph.D
Orientador

Campina Grande, Paraíba, Brasil
©Raphael de Medeiros Souto Maior Baltar, Fevereiro de 2015

FICHA CATALOGRÁFICA ELABORADA PELA BIBLIOTECA CENTRAL DA UFCG

B197i Baltar, Raphael de Medeiros Souto Maior.
Identificação relevante ao controle preditivo – Método EM-PH de
otimização numérica baseado em técnicas de regularização / Raphael de
Medeiros Souto Maior Baltar. – Campina Grande, 2015.
88 f. : il. color.

Dissertação (Mestrado em Engenharia Elétrica) – Universidade Federal
de Campina Grande, Centro de Engenharia Elétrica e Informática, 2015.

"Orientação: Prof. Dr. Péricles Rezende Barros".
Referências.

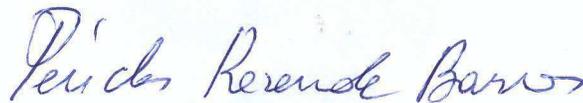
1. Controle. 2. Regularização. 3. Otimização Numérica. 4. Identificação
MRI. I. Barros, Péricles Rezende. II. Título.

CDU 621.337.1(043)

"IDENTIFICAÇÃO RELEVANTE AO CONTROLE PREDITIVO - MÉTODO EN-PH DE OTIMIZAÇÃO NUMÉRICA BASEADO EM TÉCNICAS DE REGULARIZAÇÃO"

RAPHAEL DE MEDEIROS SOUTO MAIOR BALTAR

DISSERTAÇÃO APROVADA EM 23/02/2015



PÉRICLES REZENDE BARROS, Ph.D., UFCG
Orientador(a)



JOSÉ SÉRGIO DA ROCHA NETO, D.Sc., UFCG
Examinador(a)



JAIDILSON JO DA SILVA, D.Sc., UFCG
Examinador(a)

CAMPINA GRANDE - PB

Agradecimentos

A Deus, pela dádiva da vida e todas as suas infinitas possibilidades.

Aos meus pais, pelo amor incondicional, os valores transmitidos e por estarem ao meu lado todos os dias, mesmo morando distante.

À minha namorada Diana, por enxugar o suor do meu rosto e por me inspirar todos os dias.

Aos meus familiares e amigos que, ao torcerem genuinamente por meu sucesso e felicidade, tornaram-se valiosas fontes de motivação.

Ao professor Péricles Rezende Barros por todas as lições dadas, a dedicação e comprometimento ao me orientar durante esta dissertação, e acima de tudo pela amizade construída.

Por fim, agradeço aos amigos verdadeiros do Laboratório de Instrumentação Eletrônica e Controle da UFCG, em especial a André Barbosa, Moisés Tavares, Rafael Lima, Stefânia Oliveira e Thiago Euzébio.

Resumo

Diferente dos controladores tradicionais, os controladores preditivos implementam as ações de controle a serem aplicadas no sistema a partir de predições das saídas do processo em um certo intervalo de tempo futuro. Isto é feito para que o controlador se antecipe a futuras variações da saída, e assim possa atuar com mais eficiência.

A abordagem mais direta para a obtenção de modelos apropriados para o controle preditivo é definir e minimizar a função de custo de múltiplos passos à frente. Isto pode ser alcançado utilizando um método de otimização não-convexa como o de Levenberg-Marquardt. Já no método PLS-PH, o problema de otimização não-linear é resolvido a partir de uma busca linear e da transformação dos preditores para o espaço de variáveis latentes (PLA, 2012).

Neste trabalho foi proposto um novo método de identificação multivariável MRI intitulado EN-PH (Elastic-Net Prediction Horizon) que consiste em um algoritmo de otimização numérica associado a técnica de regressão regularizada Elastic-Net. O novo método mostrou ser capaz de estimar os parâmetros do modelo do processo eficientemente, mesmo em presença de colinearidade entre os preditores.

Abstract

Unlike traditional controllers, predictive controllers implement the control actions to be applied to the system based on predictions of the process outputs at a certain future time interval. This is done so that the controller may anticipate future changes in the output, and thus can act more efficiently.

The most direct approach for obtaining appropriate models for predictive controllers is to define and minimize the multi-step ahead prediction error cost function. This can be achieved using a non-convex optimization method like the Levenberg-Marquardt algorithm. However, in the PLS-PH method, the nonlinear optimization problem is solved by a linear search method that performs a transformation of predictors for the latent variables space (PLA, 2012).

This work proposes a new MRI multivariable identification method entitled EN-PH (Elastic-Net Prediction Horizon) that consists of a numerical optimization algorithm associated with the regularized regression technique Elastic-Net. The new method has proved capable of estimate the parameters of the process model efficiently, even in the presence of collinearity among the predictors.

Sumário

| | | |
|----------|---|-----------|
| 1 | Introdução | 3 |
| 1.1 | Contexto | 3 |
| 1.2 | Revisão Bibliográfica | 5 |
| 1.2.1 | Controle preditivo | 5 |
| 1.2.2 | Identificação relevante ao controle preditivo | 6 |
| 1.2.3 | Regime de golfada em encanamentos-riser | 7 |
| 1.3 | Objetivos | 9 |
| 1.4 | Contribuições | 9 |
| 1.5 | Organização do Texto | 10 |
| 2 | Identificação relevante ao controle preditivo | 11 |
| 2.1 | Introdução | 11 |
| 2.2 | Controle Preditivo | 12 |
| 2.2.1 | Preditor | 13 |
| 2.2.2 | Obtendo um preditor linear | 14 |
| 2.2.3 | Função de custo de controle | 15 |
| 2.3 | MRI: abordagem paramétrica | 16 |
| 2.3.1 | Modelo | 16 |
| 2.3.2 | Preditor de um passo à frente | 17 |
| 2.3.3 | Preditor de múltiplos passos à frente | 17 |
| 2.3.4 | Minimizar a função de custo multi-passos à frente | 18 |
| 2.4 | Métodos de Variáveis Latentes (LVM) | 21 |
| 2.4.1 | Algoritmo Partial Least Squares (PLS) | 23 |
| 2.5 | Regularização | 24 |
| 2.5.1 | Regressão Ridge | 27 |
| 2.5.2 | Regressão Lasso | 29 |
| 2.5.3 | Regressão Elastic-Net | 31 |
| 2.5.4 | Obtenção de λ | 32 |
| 2.6 | Conclusão | 36 |

| | | |
|----------|--|-----------|
| 3 | Identificação de sistemas baseada em otimização numérica | 37 |
| 3.1 | Introdução | 37 |
| 3.2 | Algoritmo Least Squares - Prediction Horizon (LS-PH) | 38 |
| 3.3 | Algoritmo Partial Least Squares - Prediction Horizon (PLS-PH) | 41 |
| 3.4 | Algoritmo Elastic Net - Prediction Horizon (EN-PH) | 42 |
| 3.5 | Conclusão | 45 |
| 4 | Avaliação das técnicas de identificação - Simulação/Experimento | 46 |
| 4.1 | Introdução | 46 |
| 4.2 | Indicadores de validação dos modelos | 47 |
| 4.2.1 | J_{LRPIEV} | 47 |
| 4.2.2 | R^2 | 47 |
| 4.2.3 | Coefficientes de Theil | 48 |
| 4.3 | Coluna de destilação | 49 |
| 4.4 | Regime de fluxo de golfadas em prospecção offshore de petróleo | 53 |
| 4.4.1 | Prospecção de petróleo offshore | 53 |
| 4.4.2 | Identificação a partir do simulador | 54 |
| 4.5 | Sistema de temperatura baseado em módulos Peltier | 60 |
| 4.5.1 | Módulo peltier | 60 |
| 4.5.2 | Circuito eletrônico | 62 |
| 4.5.3 | Interface de comunicação | 63 |
| 4.5.4 | Identificação | 64 |
| 4.6 | Conclusão | 67 |
| 5 | Conclusões e Sugestões de Trabalhos Futuros | 68 |
| 5.1 | Conclusões | 68 |
| 5.2 | Sugestões de Trabalhos Futuros | 69 |
| | Referências Bibliográficas | 70 |
| A | Modelo MIMO ARX | 74 |
| B | Regime de fluxo multifásico de golfadas | 77 |
| B.1 | Escoamento multifásico | 77 |
| B.1.1 | Regime de golfadas | 79 |
| B.2 | Modelo de Jahanshahi | 81 |
| B.2.1 | Equações de conservação de massa para oleoduto e riser | 83 |
| B.2.2 | Condições de entrada | 83 |
| B.2.3 | Condições de saída | 83 |

| | | |
|-------|--|----|
| B.2.4 | Modelo do oleoduto | 84 |
| B.2.5 | Modelo do riser | 86 |
| B.2.6 | Modelo do fluxo de gás na base do riser | 87 |
| B.2.7 | Modelo do fluxo de líquido na base do riser | 88 |
| B.2.8 | Modelo de distribuição das fases na saída da válvula | 88 |

Lista de Símbolos e Abreviaturas

ARMAX - *Autoregressive Moving Average eXogenous*

ARX - *Autoregressive Exogenous*

CRI - *Control Relevant Identification*

EN-PH - *Elastic Net - Prediction Horizon*

LM - *Levenberg-Marquardt*

LVM - *Latent Variables Methods*

LP - *Linear Programming*

LS - *Least Squares*

MIMO - *Multiple Input Multiple Output*

MPC - *Model Predictive Control*

MRI - *Model predictive control Relevant Identification*

MSE - *Mean Squared Error*

PEM - *Prediction Error Methods*

PLS - *Partial Least Squares*

PLS-PH - *Partial Least Squares - Prediction Horizon*

QP - *Quadratic Programming*

SISO - *Single Input Single Output*

Lista de Tabelas

| | | |
|-----|--|----|
| 4.1 | Índices de Theil de validação - Coluna de destilação | 52 |
| 4.2 | Índices de Theil de validação - Golfada | 59 |
| 4.3 | Índices de Theil de validação - Planta térmica | 66 |
| B.1 | Parâmetros do sistema oleoduto-riser | 82 |

Lista de Figuras

| | | |
|------|---|----|
| 2.1 | Malha típica de controladores preditivos | 12 |
| 2.2 | Representação de um MPC | 14 |
| 2.3 | LVM: Projeção no espaço das variáveis latentes | 22 |
| 2.4 | caminhos das regressões para o exemplo da leucemia | 33 |
| 2.5 | Exemplo de um gráfico de validação cruzada para a regularização | 35 |
| | | |
| 4.1 | Diagrama esquemático da coluna de destilação | 49 |
| 4.2 | Conjunto de dados de identificação da coluna de destilação | 50 |
| 4.3 | Curvas de R^2 para modelo ordem 2 - Coluna de destilação | 51 |
| 4.4 | Curvas de R^2 para modelo ordem 3 - Coluna de destilação | 51 |
| 4.5 | Curvas de R^2 para modelo ordem 4 - Coluna de destilação | 51 |
| 4.6 | Representação de um sistema de oleodutos e riser | 53 |
| 4.7 | Variação da abertura da válvula de saída Z | 55 |
| 4.8 | Saídas da simulação do modelo da golfada para os dados de identificação | 55 |
| 4.9 | Curvas de R^2 para modelo ordem 2 - Golfada | 57 |
| 4.10 | Curvas de R^2 para modelo ordem 3 - Golfada | 57 |
| 4.11 | Curvas de R^2 para modelo ordem 4 - Golfada | 58 |
| 4.12 | Foto da montagem do sistema. | 60 |
| 4.13 | Diagrama da célula peltier | 61 |
| 4.14 | Diagrama estrutural do módulo peltier | 61 |
| 4.15 | Montagem dos módulos peltier | 61 |
| 4.16 | Diagrama de ligação do sistema. | 62 |
| 4.17 | Tela do Software supervisorio. | 63 |
| 4.18 | Dados de identificação da planta térmica | 64 |
| 4.19 | Curvas de R^2 para modelo ordem 2 - Planta térmica | 65 |
| 4.20 | Curvas de R^2 para modelo ordem 3 - Planta térmica | 65 |
| 4.21 | Curvas de R^2 para modelo ordem 4 - Planta térmica | 65 |
| | | |
| B.1 | Padrões de escoamento de fluxo multifásico em uma tubulação horizontal. | 78 |
| B.2 | Padrões de escoamento de fluxo multifásico em uma tubulação vertical. | 79 |

| | |
|--|----|
| B.3 Fases da golfada no riser. | 80 |
| B.4 Sistema oleoduto-riser | 81 |

Capítulo 1

Introdução

1.1 Contexto

Identificação de sistemas é uma área da modelagem matemática que lida com o problema de formular modelos matemáticos de sistemas dinâmicos baseando-se em dados experimentais gerados pelos próprios sistemas (LJUNG, 1999), (ISERMANN; MÜNCHHOF, 2011), (PAYNE, 1977), (STOICA, 1989).

Existem principalmente duas formas de construir modelos matemáticos de sistemas dinâmicos. Uma das abordagens consiste na construção analítica dos modelos através de leis físicas e químicas inerentes aos processos analisados. Entretanto, na engenharia muitas vezes as informações acerca da dinâmica dos sistemas é escassa e insuficiente, por isso as leis e os parâmetros exatos dos fenômenos são desconhecidos. Assim, faz-se necessário realizar uma aproximação que descreva satisfatoriamente o processo para uma dada aplicação. Na prática, existe um comprometimento entre a complexidade e a precisão dos modelos, incertezas são toleradas desde que não comprometam a robustez ou o desempenho do sistema (MOOR, 1996).

Todo modelo obtido por técnicas de identificação de sistemas será inevitavelmente uma aproximação do processo real, assim a validade de um modelo está diretamente relacionada com o seu propósito de uso. Um modelo preciso pode ser utilizado para simular o comportamento dinâmico do sistema. Já a classe de métodos de identificação relevantes ao controle (CRI) tem como motivação a necessidade de explorar os modelos lineares em projetos de controladores baseados em modelos. Sendo assim, os métodos CRI propõem-se a gerar um modelo que seja adequado ao problema de controle em questão, criando uma sinergia entre os algoritmos de identificação e os de controle.

Na indústria, é crescente a demanda por controladores cada vez mais sofisticados. Na área de controle avançado, destaca-se o Controle Preditivo por Modelo (MPC). Em essência, o objetivo do MPC é otimizar a previsão do comportamento futuro do processo

a partir das variáveis manipuláveis. O MPC consiste em uma família de técnicas de controle multivariável de processos que é utilizada em uma grande quantidade cenários práticos, tais quais controle de processos industriais, controle estocástico em estatística, aplicações automotivas e aeroespaciais. Este tipo de controle é habilitado para lidar com restrições nas entradas e saídas, e por este motivo permite que o sistema seja controlado próximo de seu limite físico. Por isso geralmente os controladores preditivos apresentam desempenho superior aos controladores lineares tradicionais, especialmente para sistemas multivariáveis (MACIEJOWSKI, 2002).

A cada instante de amostragem os controladores preditivos resolvem um problema de otimização para computar as ações de controle sobre um horizonte de tempo futuro finito. Em seguida, a primeira ação de controle calculada naquele horizonte de tempo é aplicada ao sistema. No próximo instante de amostragem, esta estratégia é repetida. No MPC, o comportamento das variáveis controladas é previsto baseando-se no modelo do processo. Sendo assim, o desempenho do controle depende fortemente da fidelidade do modelo utilizado. Por este motivo o desenvolvimento do modelo é o passo mais crítico e também o mais demorado na implementação de um controlador preditivo (MORARI; LEE., 1999).

Colinearidade é um fenômeno estatístico em que duas ou mais variáveis de predição de um modelo são altamente correlacionadas, o que significa que uma variável pode ser prevista linearmente a partir das outras. Nesta situação, as estimativas dos coeficientes da regressão podem mudar de forma irregular em resposta a pequenas mudanças do modelo ou dos dados. Este é um fenômeno comum em processos reais, especialmente em sistemas MIMO (Multiple input and Multiple output). A colinearidade pode causar problemas na estimação dos parâmetros porque ela infla a variância dos parâmetros de regressão e potencialmente leva à identificação errada dos preditores relevantes em um modelo. Colinearidade é um problema grave quando um modelo é treinado (identificado) em dados de uma região ou tempo, e utilizado para prever o comportamento das saídas em outra região ou tempo, com uma estrutura de colinearidade diferente ou desconhecida.

Os métodos de regressão regularizados quebraram o paradigma do problema de regressão linear de que era primordial buscar um modelo sem polarização, ou a menor possível. Foi visto que ao permitir que as estimativas sejam levemente polarizadas foi possível reduzir suas variâncias e, em consequência, o erro médio quadrático total.

Nesse contexto, este trabalho tem como objetivo geral desenvolver um novo método de identificação MRI, intitulado Elastic Net - Prediction Horizon (EN-PH), capaz de lidar com a presença de colinearidade nos dados de identificação. O EN-PH é um método de otimização numérica que realiza uma busca linear com base nos métodos de regularização para estimar os parâmetros do modelo. O desempenho e as propriedades do método

são estudados em dois exemplos de simulação e um exemplo experimental, em todos os exemplos o método mostrou-se uma alternativa competitiva para identificação de sistemas.

1.2 Revisão Bibliográfica

Na revisão bibliográfica desta dissertação será abordada inicialmente a evolução do controle preditivo. Em seguida serão apresentadas as principais contribuições na área de identificação relevante ao controle preditivo. Por fim, serão vistos os principais trabalhos acerca do tema regime de golfada em encanamentos-riser.

1.2.1 Controle preditivo

O Controle Preditivo foi proposto independentemente por diversos pesquisadores em uma mesma época. Apesar de data de publicação fornecer uma ideia de precedência, o Controle Preditivo foi originalmente desenvolvido por profissionais da indústria que implementaram os conceitos inerentes ao MPC alguns anos antes das primeiras publicações aparecerem (BOHLIN; ASTROM, 1965).

A patente mais antiga relacionada ao Controle Preditivo foi obtida por Martin-Sanchez em 1976, cujo método foi chamado de *Adaptive Predictive Control*. Assim como o nome indica, a ênfase estava no uso do modelo interno para obter um controle adaptativo, isto era feito pela adaptação do modelo e otimização dos sinais de controle.

Em (RICHALET et al., 1978), Richalet, engenheiro da companhia francesa Adersa, propôs um algoritmo de controle preditivo sob o nome de *Model Predictive Heuristic Control*. O *software* foi batizado como IDCOM, um anacronismo para Identificação e Comando. O objetivo nesse caso era de criar uma metodologia de controle que pudesse solucionar problemas difíceis de lidar com o convencional controlador PID, mas que também fosse baseado em conceitos intuitivos e de fácil sintonia. Manipulação de restrições e otimização não foram enfatizadas nesse método.

Segundo Morari e Lee (MORARI; LEE., 1999), as ideias de Controle Receding Horizon e Controle Preditivo por Modelo são anteriores à década de 60, como o preditor de Smith publicado em 1957. Apesar disso, o interesse nesse campo surgiu apenas nos anos 80 após publicação de artigos referentes ao Controle por Matriz Dinâmica (DMC) (CUTLER; RAMAKER, 1980) e Controle Preditivo Generalizado (GPC) (CLARKE; MOHTADI; S, 1987)

Embora os controladores preditivos tenham se originado na indústria petroquímica (CUTLER; RAMAKER, 1980), eles têm sido aplicados com sucesso em muitas outras áreas, e.g., controle de tráfego (BELLEMANS; SCHUTTER; MOOR, 2006), conversores de potência (KOURO et al., 2009) e indústria automotiva (HROVAT et al., 2012).

1.2.2 Identificação relevante ao controle preditivo

É reconhecido na literatura que a teoria de identificação de sistemas começou a ser estudada efetivamente nos anos 60 com os trabalhos apresentados em (ASTROM, 1965) e (KALMAN, 1966). O objetivo de Astrom e Bohlin foi elaborar modelos ARMAX SISO partindo de sequências de dados dos sinais de entrada e saída. Desde a apresentação desse artigo, inúmeras técnicas de identificação estatística surgiram denominadas *Métodos de Erro de Predição* (PEM), culminando em trabalhos reconhecidamente estabelecidos como (LJUNG, 1999) e (STOICA, 1989).

A necessidade da obtenção de modelos matemáticos para o desenvolvimento de algoritmos de controle não é exclusiva dos controladores preditivos. Um caso crítico ocorre nos algoritmos de controle robusto, nos quais, além do modelo da planta, é necessário informações acerca do erro de modelagem. Em (GEVERS, 2002) é revisado um dos grandes desafios existentes no início da década de 90, que era a divergência entre as estruturas de modelos obtidos por métodos de identificação tradicionais e as estruturas de modelos que o projeto dos controladores demandava.

A linha de pesquisa conhecida como MRI (MPC Relevant Identification) originou-se com o trabalho desenvolvido em (SHOOK; MOHTADI; SHAH, 1991). Aplicações de controle preditivo demandam modelos que forneçam previsões confiáveis sobre todo o horizonte de predição. Com essa premissa, em (SHOOK; MOHTADI; SHAH, 1991) apresentou-se um método cujo o objetivo era a obtenção de um modelo que fornecesse uma boa predição da resposta do sistema para todos os passos de tempo $i = 1, \dots, n_f$ à frente. Foi verificado que os parâmetros estimados das funções objetivo baseadas em preditores de múltiplos passos à frente, geralmente resultam em melhores modelos para aplicações MPC.

Em (GOPALUNI; PATWARDHAN; SHAH, 2004), a função de custo de múltiplos passos à frente do erro de predição é minimizada, e as propriedades da abordagem MRI são analisadas, demonstrando o potencial da utilização deste tipo de função de custo quando há discrepância modelo-planta, o que é comum na indústria de processo.

Em (PLA, 2012) foi realizada uma extensiva descrição dos desafios e abordagens existentes na área de MRI. E finalmente foi exposto um novo método, intitulado PLS-PH, que é uma abordagem de otimização numérica de busca linear que utiliza os métodos de variáveis latentes para estimar os parâmetros do modelo. Neste artigo foi demonstrado que o método PLS-PH pode superar consideravelmente abordagens paramétricas MRI convencionais, se há colinearidade no conjunto de dados de identificação.

Em (QUACHIO, 2012) foi apresentada uma avaliação da capacidade do método PLS-PH em gerar modelos lineares para realizar as previsões múltiplos passos à frente, para sistemas SISO e MIMO, com dados coletados em malha fechada. Também foi avaliada a

capacidade do algoritmo de identificar modelos não-lineares baseados na estrutura NARX polinomial.

Quando as suposições padrões dos mínimos quadrados são respeitadas, os coeficientes da regressão obtidos não apresentarão polarização. Além disso, dentre todas as técnicas de regressão lineares sem polarização, o modelo obtido pelo método dos mínimos quadrados também irá possuir o menor erro médio quadrático (MSE). Porém, em (TIBSHIRANI, 1996) demonstra-se que é possível produzir modelos com MSEs menores ao permitir que os parâmetros estimados tornem-se levemente polarizados. Também é reconhecido o fato de que um pequeno acréscimo na polarização pode produzir uma queda substancial na variância, resultando assim em um MSE menor do que o que seria obtido na regressão pelos mínimos quadrados ordinária. A estratégia dos métodos de regularização é combater a colinearidade na matriz de regressores produzindo modelos levemente polarizados que resultam em modelos de regressão onde o MSE geral é bastante competitivo (KUHN; JOHNSON, 2013).

1.2.3 Regime de golfada em encanamentos-riser

Um processo multivariável não-linear suscetível à colinearidade e de grande interesse nas indústrias de petróleo é o sistema de prospecção *offshore* com tubulação riser. Neste processo ocorre um fenômeno conhecido como golfada que é caracterizado por oscilações severas no fluxo multifásico da saída do sistema e na pressão do topo e da base do riser. A golfada é reconhecidamente um sério problema na prospecção *offshore* de petróleo por comprometer a segurança e a produtividade do processo, por isso muito esforço tem sido despendido a fim de prevenir a planta de operar em tal regime (BAI; BAI, 2005).

Uma solução convencional para este problema é reduzir manualmente a abertura da válvula de saída do riser para acelerar o aumento da pressão em sua base. Isto vai forçar a liberação da passagem de gás pelo ponto baixo e assim, estabilizar o fluxo. Porém esta solução reduz a produtividade da planta, especialmente em campos de extração com baixa energia (pressão e temperatura).

Todavia, a aplicação de sistemas de controle tem se mostrado uma abordagem eficiente na supressão da golfada, como mostram os testes experimentais em (GODHAVN; FARD; FUCHS, 2005). Para a elaboração desta e de outras estratégias de controle, é fundamental a criação de modelos matemáticos simples, mas que sejam capazes de descrever a dinâmica essencial do fenômeno da golfada. Com o objetivo de controlar o sistema, é mais importante captar as principais dinâmicas do início da golfada, não a golfada propriamente dita. Assim, características como comprimento e formato da golfada não são fundamentais na modelagem do sistema (JAHANSHAH, 2013).

Destacam-se três modelos dinâmicos na literatura que têm como escopo de concepção

o projeto de controladores. Storakaas foi pioneiro ao modelar o fenômeno da golfada em encanamentos risers, em (STORKAAS; SKOGESTAD, 2003) é apresentado seu modelo de espaço de estados que tem sido utilizado, por exemplo, na análise de controlabilidade (STORKAAS; SKOGESTAD, 2007). Em (MEGLIO; KAASA; PETIT, 2009) foi introduzido um modelo de três estados, representados por equações diferenciais ordinárias (EDO). Neste modelo, cinco parâmetros de sintonia são necessários para realizar o ajuste do modelo ao caso de interesse. Por fim, em (JAHANSHAH, 2013) foi desenvolvido um modelo simples de quatro estados e quatro parâmetros de sintonia. Em (MELAND, 2010) foram realizadas diversas comparações a partir de simulações dos dois últimos modelos e mostrou-se que o New Model e o DiMeglio têm respostas em frequência semelhantes e ambos resultam em boas representações do sistema riser real. Assim, o New Model destaca-se por apresentar menos parâmetros de sintonia e por estes serem totalmente independentes das propriedades físicas do sistema.

Como citado, na literatura existem diversos modelos matemáticos que descrevem este fenômeno, porém eles dependem de várias constantes de ajuste do modelo ao processo em questão. Assim, para realizar o projeto de um controlador dedicado a estabilizar o regime da golfada, é interessante utilizar um modelo obtido a partir de dados de identificação da própria planta.

1.3 Objetivos

Este trabalho teve como objetivos os seguintes itens:

- Estudar os métodos de identificação MRI;
- Analisar as propriedades dos métodos de regularização;
- Desenvolver um novo método de identificação MRI que agrega as características da regressão linear regularizada para criar o algoritmo de otimização numérica EN-PH.
- Avaliar o desempenho da técnica em aplicações simuladas e experimentais;

1.4 Contribuições

As contribuições apresentadas neste trabalho são:

- Desenvolvimento do método de identificação relevante ao controle preditivo EN-PH;
- A identificação de sistemas e análise comparativa de métodos de identificação MRI, incluindo o método proposto.

1.5 Organização do Texto

No capítulo 2 são apresentados os tópicos pertinentes à identificação de sistemas relevantes ao controle preditivo. Inicialmente é apresentada uma descrição geral do controle preditivo e suas principais características. Em seguida é apresentada uma revisão geral sobre os conceitos fundamentais dos métodos de identificação relevantes ao controle preditivo. Ao fim desta seção é apresentada uma técnica de identificação MRI tradicional que utiliza o método de otimização de Levenberg-Marquardt para minimizar a função de custo múltiplos passos à frente. Também é realizada uma revisão sobre os métodos de variáveis latentes e os métodos de regularização. No capítulo 3 são vistos os algoritmos de busca desenvolvidos para estimar os parâmetros dos modelos. Inicialmente, apresenta-se o algoritmo base Least Squares - Prediction Horizon (LS-PH), base para os outros dois métodos de busca linear apresentados. Em seguida o método Partial Least Squares - Prediction Horizon (PLS-PH) é apresentado como uma extensão do LS-PH baseado nos métodos de variáveis latentes. Por fim, define-se o algoritmo EN-PH como uma extensão do LS-PH baseando-se nos métodos de regularização. No capítulo 4 são introduzidos os índices de avaliação de desempenho dos métodos de identificação que serão utilizados nos exemplos. Em seguida, três aplicações foram submetidas aos métodos a fim de avaliar o desempenho preditivo, as características do modelo identificado e validar o novo método proposto. Finalmente, no capítulo 5 são expostas as conclusões finais do trabalho juntamente com as sugestões de trabalhos futuros.

Capítulo 2

Identificação relevante ao controle preditivo

2.1 Introdução

O desenvolvimento do modelo é o passo mais crítico e demorado na implementação de um controlador preditivo (MORARI; LEE., 1999). Existem diferentes estruturas de modelos capazes de representar aproximadamente o processo, a escolha deve ser realizada com base no propósito de sua utilização e na natureza de cada processo. Por exemplo, caso se deseje criar um simulador para estudar o comportamento dinâmico de um processo, um modelo sofisticado e preciso deve ser utilizado. Entretanto, se o objetivo é projetar um controlador para aquele processo, um modelo simplificado é geralmente preferível para facilitar a implementação da estratégia de controle.

Na identificação tradicional minimiza-se a função de custo do erro de predição a um passo à frente. Uma vez que no MPC o modelo realiza predições acerca do comportamento do processo por toda uma janela de predição, uma abordagem direta para MRI é a minimização da função de custo do erro de predição multi-passos à frente para ajustar o modelo ao conjunto de dados de identificação. Em (GOPALUNI; PATWARDHAN; SHAH, 2004) a função de custo do erro de predição multi-passos à frente foi minimizada, e analisou-se as propriedades do MRI demonstrando o potencial desta abordagem. Uma desvantagem deste método é que a função de custo de identificação é não linear em seus parâmetros, e em geral o problema de minimização é não convexo. A complexidade computacional da otimização não-convexa não vem a ser um empecilho, uma vez que o estágio de identificação do modelo é realizado *offline*.

É comum verificar algum grau de colinearidade entre as variáveis em conjuntos de dados colhidos de processos reais. Uma alta colinearidade entre os preditores significa que as variáveis do conjunto partilham quantidades de informações redundantes. Este

fenômeno pode causar problemas para a estimação dos parâmetros pois gera um aumento na variância dos parâmetros de regressão. Com isso, o resultado da identificação dos parâmetros do modelo relevantes ao processo é comprometido. Neste capítulo também serão apresentados os métodos de variáveis latentes e os métodos de regularização, que são duas diferentes abordagens para lidar com o problema de colinearidade nos dados de identificação.

2.2 Controle Preditivo

Controle preditivo, ou do inglês *Model Predictive Control* (MPC) é uma técnica avançada de controle multivariável de processos usada em diversas aplicações práticas, tais quais controle industrial de processos; controle estocástico, em economia; em aplicações automotivas e aeronáuticas. Tal abrangência deve-se ao fato de que o MPC é baseado em ideias bastante intuitivas e por considerar sistemas multivariáveis sujeitos a restrições nas entradas e saídas, desde a concepção do controlador (MACIEJOWSKI, 2002).

As premissas básicas, comuns aos diversos tipos de controladores preditivos, são descritas à seguir:

Na figura 2.1 é ilustrada uma malha de controle MPC típica:

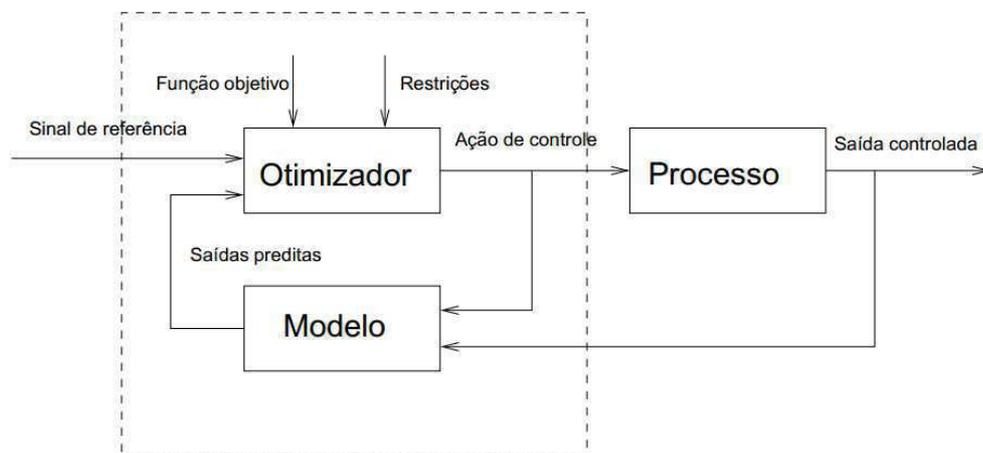


Figura 2.1: Malha típica de controladores preditivos

- Como passo inicial, o modelo do processo é usado para realizar as previsões acerca do comportamento do processo em uma dada janela de predição: $\hat{\mathbf{y}}_{k+i} = \forall i \in [1, 2, \dots, n_f]$. Onde $\hat{\mathbf{y}}_{k+i} \in \mathbb{R}^{1 \times n_0}$ são as previsões em $k+i$ a partir das informações disponíveis até o instante k ; n_0 , o número de saídas do processo; n_f , o número de passos futuros preditos, também chamado de horizonte de predição.

- A sequência de controle será obtida à partir da minimização de uma função de custo. Funções de custo típicas que são usadas no projeto de controladores devem ponderar: desvios entre saída e uma dada trajetória desejada e o esforço de controle. Se a função de custo for quadrática, o modelo linear e não houverem restrições de desigualdades impostas ao sistema, haverá uma expressão analítica para o controlador MPC. Caso contrário, otimização numérica pode ser usada para encontrar a ação de controle que minimiza a função de custo.
- A cada instante de amostragem uma sequência de controle é obtida, seguindo os passos anteriores, porém apenas a primeira ação de controle será, de fato, aplicada ao processo. Esta característica do MPC é conhecida como horizonte móvel. O motivo de se utilizar a política de horizonte móvel é que no instante $k + 1$ haverá informações acerca das saídas que não estavam disponíveis no instante k , assim, estes erros de predição podem ser contabilizados no cálculo da próxima ação de controle.

2.2.1 Preditor

O preditor é responsável por fornecer as predições das saídas do processo sobre uma janela de tempo futura. As predições são realizadas a partir do modelo do processo e são expressas em termos das informações conhecidas à priori das entradas, saídas e da sequência de controle futura. O preditor é o componente fundamental no projeto de um controlador preditivo, uma vez que as ações de controles são baseadas nas predições das variáveis controladas. Quão melhor for o preditor, melhor será a performance do controlador; assim, obter um preditor confiável é crucial em MPC.

Uma forma generalizada das predições futuras pode ser expressa como:

$$\hat{\mathbf{y}}_f(k) = f(\mathbf{u}_p(k), \mathbf{y}_p(k), \mathbf{u}_f(k), \mathbf{u}_{dof}(k)) \quad (2.1)$$

Na figura 2.2 é apresentado um exemplo dos sinais que compõem o MPC.

Os vetores linhas são definidos como:

$$\begin{aligned} \hat{\mathbf{y}}_f(k) &= [\hat{\mathbf{y}}(k+1) \dots \hat{\mathbf{y}}(k+n_f)] \rightarrow \text{Saídas preditas} \\ \left. \begin{aligned} \mathbf{u}_p(k) &= [\mathbf{u}(k-1) \dots \mathbf{u}(k-n_b+1)] \\ \mathbf{y}_p(k) &= [\mathbf{y}(k-1) \dots \mathbf{y}(k-n_a)] \end{aligned} \right\} \rightarrow \text{Dados passados conhecidos} \\ \left. \begin{aligned} \mathbf{u}_f(k) &= [\mathbf{u}(k+n_f-1) \dots \mathbf{u}(k+n_u)] \\ \hat{\mathbf{u}}_{dof}(k) &= [\mathbf{u}(k+n_u-1) \dots \mathbf{u}(k)] \end{aligned} \right\} \rightarrow \text{Sequência de controle futura} \end{aligned}$$

onde n_b é o horizonte do passado das entradas; n_a o horizonte passado das saídas; n_f

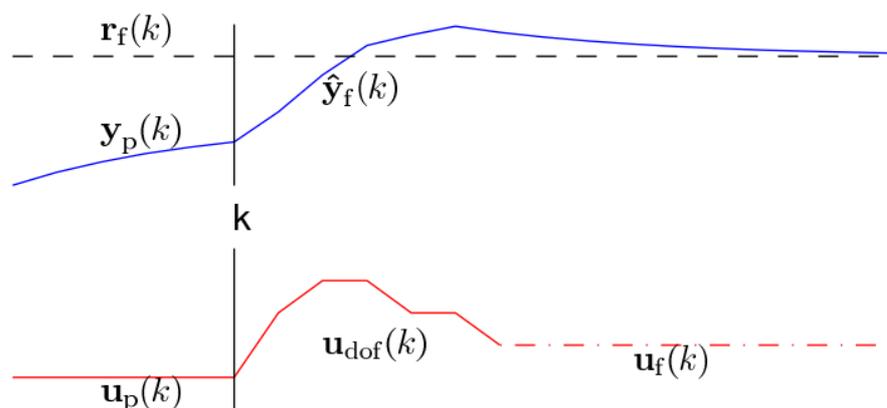


Figura 2.2: Representação de um MPC

é o horizonte de predição e n_u o horizonte de controle. Os detalhes de como se obter $f(\cdot)$ dependem da estratégia de identificação e/ou modelo utilizado. A subseção a seguir explica como obter um modelo linear para ser usado como preditor.

2.2.2 Obtendo um preditor linear

Assumindo uma estrutura linear, o preditor da equação (2.1) será:

$$\hat{y}_f(k) = \underbrace{[\mathbf{u}_p(k) \quad \mathbf{y}_p(k) \quad \mathbf{u}_f(k) \quad \mathbf{u}_{dof}(k)]}_{\mathbf{x}(k)} \boldsymbol{\theta} \quad (2.2)$$

Onde $\boldsymbol{\theta}$ é a matriz que contém os parâmetros do preditor linear e $\mathbf{x}(k)$ é o vetor de regressão no instante k .

Quando há um conjunto de dados de identificação, as seguintes matrizes podem ser formadas:

$$\mathbf{Y} = \begin{bmatrix} \mathbf{y}_f(1) \\ \vdots \\ \mathbf{y}_f(N) \end{bmatrix} \quad ; \quad \mathbf{X} = \begin{bmatrix} \mathbf{x}(1) \\ \vdots \\ \mathbf{x}(N) \end{bmatrix}$$

nas quais os vetores linha são obtidos como:

$$\begin{aligned} \mathbf{y}_f(k) &= [y(k+1) \dots y(k+n_f)] \\ \mathbf{x}(k) &= \underbrace{[u(k-1) \dots u(k-n_b+1)]}_{\mathbf{u}_p(k)}, \underbrace{[y(k+1) \dots y(k-n_a)]}_{\mathbf{y}_p(k)}, \\ &\quad \underbrace{[u(k+n_f-1) \dots u(k+n_u)]}_{\mathbf{u}_f(k)}, \underbrace{[u(k+n_u-1), \dots, u(k)]}_{\mathbf{u}_{dof}(k)} \end{aligned}$$

$\forall k \in [1 \dots N]$. Note que $\mathbf{y}_f(k)$ é um vetor linha com $n_y = n_0 n_f$ colunas, e $\mathbf{x}(k)$ é um vetor linha com $n_x = n_i(n_b - 1) + n_0 n_a + n_i n_f$ colunas. Assim, $\mathbf{Y} \in \mathbb{R}^{N \times n_y}$ e $\mathbf{X} \in \mathbb{R}^{N \times n_x}$.

Considerando o erro de modelagem tem-se:

$$\mathbf{Y} = \underbrace{\mathbf{X}\boldsymbol{\theta}}_{\hat{\mathbf{Y}}} + \mathbf{F}$$

onde F é o vetor do erro de modelagem e $\hat{\mathbf{Y}}$ é a matriz de predições. Dadas as matrizes \mathbf{X} e \mathbf{Y} , a abordagem de identificação mais simples é utilizar o LS (*least squares*) para ajustar o modelo:

$$\boldsymbol{\theta} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y}$$

2.2.3 Função de custo de controle

A expressão geral para a função de custo no MPC, em um dado instante, é dada por:

$$J_C = g(\mathbf{r}(k), \hat{\mathbf{y}}_f(k), \mathbf{u}_{dof}(k)) \quad (2.3)$$

onde $r(k)$ é o vetor de referência das saídas.

No projeto de controladores geralmente utiliza-se modelos lineares (QIN; BADGWELL, 2003), além disso a minimização de funções quadráticas com restrições é tratável matematicamente. Assim, a função de custo típica em MPC é dada por:

$$J_C = \|[\mathbf{r}_f(k) - \hat{\mathbf{y}}_f(k)] \mathbf{W}_y\|_F^2 + \lambda_u \|\Delta \mathbf{u}_{dof}(k) \mathbf{W}_u\|_F^2 \quad (2.4)$$

onde \mathbf{W}_y é a matriz de ponderação dos desvios das saídas; \mathbf{W}_u a matriz de ponderação do esforço de controle; λ_u , o peso geral das ações de controles. Também é comum utilizar incrementos das ações de controle $\Delta \mathbf{u}_{dof}$ ao invés da própria ação de controle a fim de obter um controle livre de *offset* (ROSSITER, 2003).

A minimização da função de custo na equação (2.3) sujeita a restrições de desigualdades nas MVs (Manipulated Variables) ou CVs (Controlled Variables) pode ser expressa como:

$$\min_{\Delta \mathbf{u}_{dof}(k)} J_C(k) \quad s.t. \text{ restrições} \quad (2.5)$$

No caso geral, deve-se utilizar um otimizador *online* para resolver o problema de minimização e obter a sequência de controle \mathbf{u}_{dof} . Se o modelo for linear e a função de custo quadrática, a minimização pode ser resolvida utilizando QP (Quadratic Programming). Caso a função de custo seja não linear em $\boldsymbol{\theta}$, geralmente utiliza-se métodos de otimização numérica não convexas.

2.3 MRI: abordagem paramétrica

Nesta seção, inicialmente será definida a estrutura do modelo utilizado, e a partir do modelo obtém-se o preditor do sistema. Em seguida, define-se a função de custo multi-passos à frente. Por fim, apresenta-se um método capaz de minimizar a função de custo de múltiplos passos à frente e a obter os parâmetros dos modelos.

2.3.1 Modelo

Num estudo inicial, o modelo SISO de Box-Jenkins será usado representando a expressão geral para um modelo LTI na forma de função de transferência:

$$y(k) = \frac{\mathbf{B}(z^{-1})}{\mathbf{A}(z^{-1})}u(k) + \frac{\mathbf{C}(z^{-1})}{\mathbf{D}(z^{-1})}\xi(k)$$

onde: $y(k)$ é a saída do processo; $u(k)$, a entrada do processo; e $\xi(k)$ um ruído branco. É comum não obter sucesso na estimação de $\mathbf{C}(z^{-1})$, além disso as características do distúrbio são frequentemente alteradas. Por esses motivos é uma prática comum em MPC usar o modelo CARIMA Rossiter:

$$\mathbf{C}(z^{-1}) = \mathbf{T}(z^{-1}) \quad ; \quad \mathbf{D}(z^{-1}) = \mathbf{A}(z^{-1})(1 - z^{-1}) \quad (2.6)$$

onde $\mathbf{T}(z^{-1})$ é considerado como uma escolha de projeto. O modelo SISO pode ser transformado na forma ARX:

$$y(k) = \frac{\mathbf{B}(z^{-1})}{\mathbf{A}(z^{-1})}u(k) + \frac{\mathbf{T}(z^{-1})}{\mathbf{A}(z^{-1})(1 - z^{-1})}\xi(k)$$

$$\frac{(1 - z^{-1})}{\mathbf{T}(z^{-1})}y(k) = \frac{\mathbf{B}(z^{-1})}{\mathbf{A}(z^{-1})} \frac{(1 - z^{-1})}{\mathbf{T}(z^{-1})}u(k) + \frac{1}{\mathbf{A}(z^{-1})}\xi(k)$$

$$y_f(k) = \frac{\mathbf{B}(z^{-1})}{\mathbf{A}(z^{-1})}u_f(k) + \frac{1}{\mathbf{A}(z^{-1})}\xi(k)$$

Assumindo que os dados de identificação foram previamente filtrados, uma estrutura ARX pode ser assumida.

$$y_f(k) = \frac{1 - z^{-1}}{\mathbf{T}(z^{-1})}y(k) \quad ; \quad u_f(k) = \frac{1 - z^{-1}}{\mathbf{T}(z^{-1})}u(k)$$

No apêndice A é apresentada a equivalência entre a estrutura ARX e a seguinte ex-

pressão para o caso MIMO:

$$\mathbf{y}(k) = \mathbf{x}(k-1)\boldsymbol{\theta} + \boldsymbol{\xi}(k) \quad (2.7)$$

com,

- $\mathbf{y}(k) = [y_1(k) \dots y_{n_0}(k)]$
- $\mathbf{x}(k-1) = [\mathbf{u}(k-1) \dots \mathbf{u}(k-n_b), \mathbf{y}(k-1) \dots \mathbf{y}(k-n_a)]$, é o vetor de regressão onde:
 - $\mathbf{u}(k) = [u_1(k) \dots u_{n_i}(k)]$
 - n_b : número de entradas armazenadas no modelo
 - n_a : número de saídas armazenadas no modelo
- $\boldsymbol{\theta}$ é a matriz de parâmetros que define o modelo. Sua dimensão é $n_x \times n_0$ (sendo n_x o número de colunas em \mathbf{x}_k , $n_x = n_i n_b + n_0 n_a$)
- $\boldsymbol{\xi}_k = [\xi_1(k) \dots \xi_{n_0}(k)]$

2.3.2 Preditor de um passo à frente

Se o modelo estimado é definido por θ , então as predições das saídas no instante $k+1$, a partir das informações disponíveis até o instante k são obtidas como:

$$\hat{\mathbf{y}}(k+1) = \mathbf{x}(k)\boldsymbol{\theta} \quad (2.8)$$

onde,

$$\mathbf{x}(k) = [\mathbf{u}(k) \dots \mathbf{u}(k+1-n_b), \mathbf{y}(k) \dots \mathbf{y}(k+1-n_a)]$$

Assumindo a inexatidão entre a planta e o modelo, $\xi(k)$ na equação (2.7) não será mais um ruído branco, e as saídas reais do processo são dadas por:

$$\mathbf{y}(k+1) = \hat{\mathbf{y}}(k+1) + \mathbf{e}(k+1)$$

sendo $\mathbf{e}(k+1)$ o erro de identificação no instante $k+1$ com dados das saídas disponíveis até o instante k .

2.3.3 Preditor de múltiplos passos à frente

A partir da equação (2.8), o modelo de predição de erro multi-passos à frente dispondo de informações das saídas até o instante k , e das entradas até o instante $k+n_f$ pode ser

formulado como:

$$\hat{\mathbf{y}}(k+j) = \mathbf{x}(k+j-1)\boldsymbol{\theta}, \quad \forall j \in [1, 2, \dots, n_f] \quad (2.9)$$

onde

$$\mathbf{x}(k+j-1) = [\mathbf{u}(k+j-1) \dots \mathbf{u}(k+j-n_b), \bar{\mathbf{y}}(k+j-1) \dots \bar{\mathbf{y}}(k+j-n_a)] \quad (2.10)$$

$$\bar{\mathbf{y}}(\alpha) = \begin{cases} \hat{\mathbf{y}}(\alpha) & \text{para } (\alpha) > k \\ \mathbf{y}(\alpha) & \text{para } (\alpha) \leq k \end{cases} \quad (2.11)$$

$\hat{\mathbf{y}}(k+n_f)$ pode ser obtido recursivamente, iniciando em $\hat{\mathbf{y}}(k+1)$ e utilizando a equação (2.9). As saídas reais podem ser expressas como:

$$\mathbf{y}(k+j) = \hat{\mathbf{y}}(k+j) + \mathbf{e}(k+j) \quad (2.12)$$

sendo $\mathbf{e}(k+j)$ o erro de identificação no instante $k+j$ com dados das saídas disponíveis até o instante k .

2.3.4 Minimizar a função de custo multi-passos à frente

A função de custo básica normalmente utilizada em MPC é dada por:

$$J_C = \|[\mathbf{r}_f(k) - \hat{\mathbf{y}}_f(k)]\mathbf{W}_y\|_F^2 + \lambda_u \|\Delta \mathbf{u}_{dof}(k)\mathbf{W}_u\|_F^2$$

onde \mathbf{W}_y é a matriz de ponderação dos desvios das saídas; \mathbf{W}_u a matriz de ponderação do esforço de controle; λ_u , o peso geral das ações de controles; $\mathbf{r}_f(k)$ a matriz de referência das saídas. Objetivando a simplificação da análise, será assumido que todas as saídas recebem o mesmo peso, então $\mathbf{W}_y = \mathbf{I}$; Além disso, $\lambda_u = 0$. A função de custo torna-se:

$$J_C = \|\mathbf{r}_f(k) - \hat{\mathbf{y}}_f(k)\|_F^2$$

que pode ser reescrito como

$$J_C = \sum_{j=1}^{n_f} \underbrace{\|\mathbf{r}(k+j) - \hat{\mathbf{y}}(k+j)\|_F^2}_{\triangleq \mathbf{c}(k+j)} \quad (2.13)$$

No controle preditivo a diferença entre as predições e suas referências é minimizada. Porém, em MRI o objetivo é tratar o descasamento entre o modelo e o processo real, então

minimiza-se os desvios entre os valores reais das saídas e suas referências:

$$\tilde{J}_C \triangleq \sum_{j=1}^{n_f} \|\mathbf{r}(k+j) - \mathbf{y}(k+j)\|_F^2$$

Utilizando a equação (2.12), a função de custo pode ser expressa como função dos erros de controle e de identificação:

$$\begin{aligned} \tilde{J}_C &= \sum_{j=1}^{n_f} \|\mathbf{r}(k+j) - (\hat{\mathbf{y}}(k+j) + \mathbf{e}(k+j))\|_F^2 \\ &= \sum_{j=1}^{n_f} \|\underbrace{\mathbf{r}(k+j) - \hat{\mathbf{y}}(k+j)}_{\mathbf{c}(k+j)} - \mathbf{e}(k+j)\|_F^2 \\ &= \sum_{j=1}^{n_f} \|\mathbf{c}(k+j) - \mathbf{e}(k+j)\|_F^2 \end{aligned}$$

que pode ser reescrita como:¹

$$\tilde{J}_C = \sum_{j=1}^{n_f} \text{tr}(\mathbf{c}^T(k+j)\mathbf{c}(k+j)) \quad (2.14a)$$

$$+ \sum_{j=1}^{n_f} \text{tr}(\mathbf{e}^T(k+j)\mathbf{e}(k+j)) \quad (2.14b)$$

$$- \sum_{j=1}^{n_f} \text{tr}(\mathbf{c}^T(k+j)\mathbf{e}(k+j) + \mathbf{e}^T(k+j)\mathbf{c}(k+j)) \quad (2.14c)$$

Note que o somatório em (2.14a) é igual à função de custo minimizada no MPC obtida na equação (2.13). Já o somatório em (2.14c) representa a correlação cruzada entre os erros de controle e identificação, e de acordo com (GEVERS, 2002) pode ser resolvido utilizando ajuste iterativo. Porém esta estratégia está fora do escopo deste trabalho.

Uma vez que o somatório em (2.14b) depende apenas do erro de identificação, ele é o componente mais relevante do \tilde{J}_C para esta seção, e será referido como J_{lrpi} (do inglês Long Range Prediction Identification cost index). O J_{osapi} (do inglês One Step Ahead Prediction Identification cost index) normalmente utilizado em métodos de predição de

¹A norma de Frobenius pode ser expressa usando o operador traço: $\|\mathbf{A}\|_F^2 = \text{tr}(\mathbf{A}^T \mathbf{A})$

erro (PEM) como os mínimos quadrados, é obtido fazendo $n_f = 1$.

$$J_{lrpi}(k) = \sum_{j=1}^{n_f} \|\mathbf{e}(k+j)\|_F^2 \quad (2.15)$$

$$J_{osapi}(k) = \|\mathbf{e}(k+1)\|_F^2 \quad (2.16)$$

Intuitivamente, é fácil perceber que no escopo de controle preditivo o índice J_{lrpi} é mais apropriado do que J_{osapi} . MRI foi estudado minunciosamente em (GOPALUNI; PATWARDHAN; SHAH, 2004) e verificou-se melhor desempenho do controlador preditivo quando estima-se os parâmetros do modelo minimizando J_{lrpi} ao invés de J_{osapi} . Também foi provado que qualquer experimento que é informativo o suficiente para a abordagem de um passo à frente, também será informativo o suficiente para a abordagem multi-passos à frente. Com isso, se o modelo destina-se a ser usado em MPC, o índice a ser minimizado no estágio de identificação é o J_{lrpi} . O índice pode ser avaliado para um conjunto de dados como:

$$J_{LRPI} = \sum_{k=1}^N J_{lrpi}(k) = \sum_{k=1}^N \sum_{j=1}^{n_f} \|\mathbf{e}(k+j)\|_F^2 \quad (2.17)$$

onde,

$$\mathbf{e}(k+j) = \mathbf{y}(k+j) - \mathbf{x}(k+j-1)\boldsymbol{\theta} \quad (2.18)$$

A partir da equação (2.10), $\mathbf{x}(k+j-1)$ contém os elementos $\hat{\mathbf{y}}(k+j-\beta), \forall \beta \in [1, 2, \dots, j-1]$, que por sua vez, depende de $\boldsymbol{\theta}$; Por isso, este é um problema não linear em $\boldsymbol{\theta}$ e não há uma solução fechada como no caso da abordagem um passo à frente.² Assim, faz-se necessário recorrer a algoritmos de otimização numérica iterativa para minimizar J_{LRPI} .

Métodos dos mínimos quadrados não lineares envolvem uma melhoria iterativa dos valores dos parâmetros, de modo a reduzir a soma do quadrado dos erros da função de custo. O método de Levenberg-Marquardt (LM) é uma das técnicas padrão para resolver problemas de mínimos quadrados não lineares (MORE, 1977). O método LM é uma combinação de duas técnicas de minimização: o método de gradiente descendente e o método de Gauss-Newton. No método de gradiente descendente, a soma dos quadrados dos erros é reduzida ao atualizar os parâmetros na direção da maior redução da função de custo. No método de Gauss-Newton, a soma dos erros quadráticos é reduzida assumindo que a função dos mínimos quadrados é localmente quadrática, e encontrando o mínimo desta aproximação. O método de Levenberg-Marquardt age mais como um método de

²A abordagem um passo à frente é linear em $\boldsymbol{\theta}$ pois requer apenas \mathbf{x}_{k+1} que é independente de $\boldsymbol{\theta}$.

gradiente descendente quando os parâmetros estão longe de seu valor ideal, e age mais como o método de Gauss-Newton quando os parâmetros são perto de seu valor ideal.

2.4 Métodos de Variáveis Latentes (LVM)

Uma das possíveis abordagens para lidar com dados de identificação correlacionados é utilizar os métodos de variáveis latentes (LVM). LVMs transformam os dados ruidosos e correlacionados em um conjunto de dados reduzido e informativo, no qual a identificação pode ser realizada com sucesso. O método dos mínimos quadrados parciais (PLS) é um LVM que determina as variáveis latentes minimizando a correlação entre os escores de entrada e saída.

Em problemas de regressão linear da forma

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\theta} + \mathbf{F}$$

é possível utilizar o método dos mínimos quadrados e obter a matriz de parâmetros $\boldsymbol{\theta}$ como:

$$\boldsymbol{\theta} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y}$$

Porém, na ocorrência de correlação nos dados de identificação, a matriz $(\mathbf{X}^T \mathbf{X})$ é mal condicionada. Por isso, a obtenção de $\boldsymbol{\theta}$ pelos mínimos quadrados não será alcançada com sucesso, causando grande erro de variância. A fim de solucionar este problema, realiza-se a regressão no espaço das variáveis latentes.

Dados \mathbf{Y} e \mathbf{X} , a redução das variáveis pode ser realizada como :

$$\mathbf{Y} = \mathbf{U}\mathbf{Q}^T + \mathbf{F} \quad (2.19)$$

$$\mathbf{X} = \mathbf{T}\mathbf{P}^T + \mathbf{E} \quad (2.20)$$

onde $\mathbf{Y} \in \mathbb{R}^{N \times n_y}$ é o espaço das saídas; $\mathbf{U} \in \mathbb{R}^{N \times n_{lv}}$, os escores de saída; $\mathbf{Q} \in \mathbb{R}^{n_y \times n_{lv}}$, as cargas de saída; $\mathbf{F} \in \mathbb{R}^{N \times n_y}$, os resíduos de saída; $\mathbf{X} \in \mathbb{R}^{N \times n_x}$ é o espaço das entradas; $\mathbf{T} \in \mathbb{R}^{N \times n_{lv}}$, os escores de entrada; $\mathbf{P} \in \mathbb{R}^{n_x \times n_{lv}}$, as cargas de entrada; $\mathbf{E} \in \mathbb{R}^{N \times n_x}$, os resíduos de entrada;

Conforme ilustrado na figura 2.3, o primeiro passo no LVM é projetar o espaço original (externo) no espaço reduzido das variáveis latentes (interno). Note que a matriz de regressão no espaço externo possui n_x colunas, e no espaço interno n_{lv} . n_{lv} é uma escolha de projeto e normalmente escolhe-se um valor inferior a n_x assumindo que há colinearidade nos dados de identificação.

O espaço das variáveis latentes é definido pelos escores \mathbf{T} e \mathbf{U} . O modelo no espaço

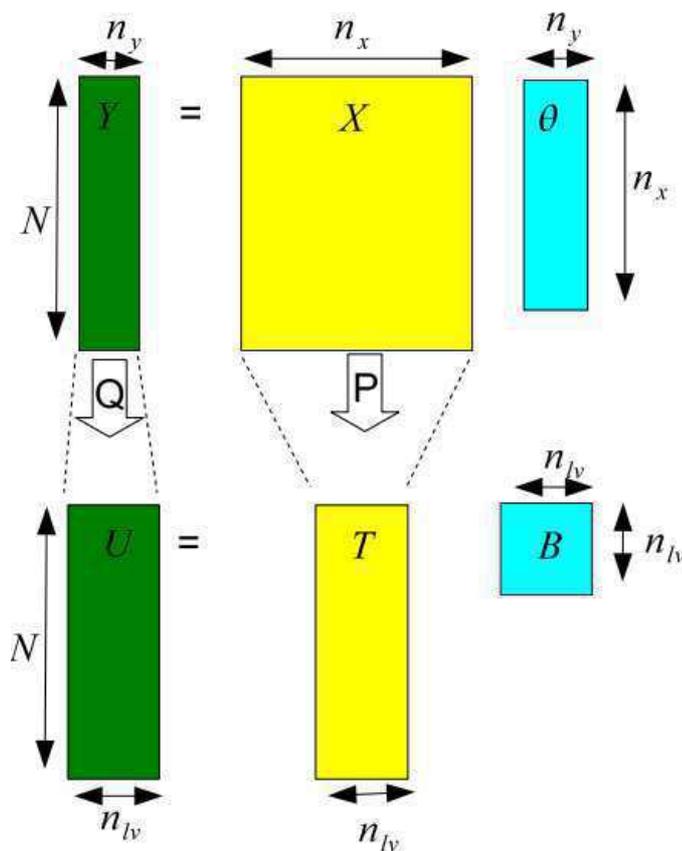


Figura 2.3: LVM: Projeção no espaço das variáveis latentes

interno pode ser obtido como:

$$\mathbf{U} = \underbrace{\mathbf{T}\mathbf{B}}_{\triangleq \hat{\mathbf{U}}} + \text{resíduos} \stackrel{LS}{\Rightarrow} \mathbf{B} = (\mathbf{T}^T\mathbf{T})^{-1}\mathbf{T}^T\mathbf{U} \quad (2.21)$$

Note que agora a identificação pode ser realizada no espaço interno, \mathbf{B} é obtido com base em \mathbf{U} e \mathbf{T} . Uma das propriedades deste método é que as colunas de \mathbf{T} são ortogonais, com isso a obtenção de \mathbf{B} não será afetada pela correlação presente nos vetores colunas de \mathbf{X} .

A fim de escolher o melhor valor para o número de variáveis latentes n_{lv} a serem usadas no espaço interno, introduz-se o índice MSEP (Mean Squared Error of Prediction)

$$MSEP = \frac{1}{N} \sum_{k=1}^N \|\mathbf{y}_f(k) - \hat{\mathbf{y}}_f(k)\|_F^2$$

Dado um conjunto de dados de identificação e um preditor, é possível calcular o MSEP. Para validar um preditor a partir do MSEP, um conjunto de dados de validação diferente daquele utilizado na identificação é geralmente utilizado. Também é possível utilizar o conjunto de dados de identificação e aplicar a validação cruzada (TROP SHA; GRAMA-

TICA; GOMBAR, 2003). De acordo com (PLA, 2012), para um dado conjunto de dados de identificação, o número de parâmetros no modelo afeta diretamente a decomposição do MSEF entre os erros de polarização e de variância. Verifica-se na figura 2.3, a dimensão de $\boldsymbol{\theta}$ é $n_x \times n_y$ e a dimensão de \mathbf{B} é $n_{lv} \times n_{lv}$, assim o ajuste de n_{lv} permite melhorar o MSEF. Além disso, \mathbf{P} e \mathbf{Q} também devem ser considerados uma vez que eles também são ajustados no conjunto de dados de identificação, e isto afeta na redução do número de parâmetros para ajuste.

A estratégia utilizada é avaliar iterativamente o índice MSEF para diversos valores de n_{lv} . Inicia-se com n_{lv} pequeno e conforme ele aumenta, MSEF diminui. Escolhe-se o n_{lv} como o menor valor que fornece uma redução significativa do MSEF.

Assim, se a identificação for realizada no espaço das variáveis latentes: \mathbf{Y} e \mathbf{X} são projetados sob um espaço reduzido (n_{lv}), e o modelo é ajustado neste espaço interno. A diferentes abordagens para definir a projeção: PLS, Regressão em Componentes Principais (PCR), regressão por redução de posto (RRR) e regressão potencial (KIERS; A.K., 2007). De acordo com a comparação entre os vários métodos realizada em (KIERS; A.K., 2007), PLS e PCR são particularmente indicados em casos de presença de colinearidade, enquanto nos demais casos é mais indicado utilizar a regressão linear ordinária.

2.4.1 Algoritmo Partial Least Squares (PLS)

As diferentes abordagens para a obtenção das cargas \mathbf{P} e \mathbf{Q} nas equações (2.19) e (2.20) dão origem a diferentes LVM. No PLS, a i -ésima variável latente é obtida tal que a correlação entre os escores de entrada t_i e os escores de saída u_i seja maximizada. Note que t_i e u_i são as i -ésimas colunas de \mathbf{T} e \mathbf{U} respectivamente.

O algoritmo PLS é descrito em detalhes em (HÖSKULDSSON, 1988), onde uma matriz adicional \mathbf{W} é usada junto a \mathbf{P} no algoritmo NIPALS modificado para obter \mathbf{T} com vetores colunas ortogonais. Como explicado em (MARTENS, 2001), \mathbf{T} pode ser obtido de \mathbf{X} como:

$$\mathbf{T} = \mathbf{X} \underbrace{\mathbf{W}(\mathbf{P}^T \mathbf{W})^{-1}}_{\triangleq \mathbf{Z}} \quad (2.22)$$

Com base nas equações (2.19), (2.20), (2.21) e (2.22), o modelo no espaço externo $\boldsymbol{\theta}$ pode ser obtido usando o PLS como:

$$\hat{\mathbf{Y}} = \hat{\mathbf{U}} \mathbf{Q}^T = \mathbf{T} \mathbf{B} \mathbf{Q}^T = \mathbf{X} \underbrace{\mathbf{Z} \mathbf{B} \mathbf{Q}^T}_{\triangleq \boldsymbol{\theta}} \quad (2.23)$$

$$\boldsymbol{\theta} = \mathbf{Z} \mathbf{B} \mathbf{Q}^T \quad (2.24)$$

O algoritmo do método PLS, implementado em (HÖSKULDSSON, 1988) é descrito a seguir:

1. Início: \mathbf{u} recebe a primeira coluna de \mathbf{Y} .
2. $\mathbf{w} = \mathbf{X}^T \mathbf{u} / (\mathbf{u}^T \mathbf{u})$
3. $\mathbf{w} = \mathbf{w} / \|\mathbf{w}\|$
4. $\mathbf{t} = \mathbf{X} \mathbf{w}$
5. $\mathbf{c} = \mathbf{Y}^T \mathbf{t} / (\mathbf{t}^T \mathbf{t})$
6. $\mathbf{c} = \mathbf{c} / \|\mathbf{c}\|$
7. $\mathbf{u} = \mathbf{Y} \mathbf{c} / (\mathbf{c}^T \mathbf{c})$
8. Testa se $\|\mathbf{u}_i - \mathbf{u}_{i-1}\| \leq \text{Tolerância}$, caso positivo prosseguir para o passo 9; caso negativo voltar ao passo 2
9. $\mathbf{p} = \mathbf{X}^T \mathbf{t} / (\mathbf{t}^T \mathbf{t})$
10. $\mathbf{l} = \mathbf{Y}^T \mathbf{u} / (\mathbf{u}^T \mathbf{u})$
11. $\mathbf{b} = \mathbf{u}^T \mathbf{t} / (\mathbf{t}^T \mathbf{t})$
12. $\mathbf{X} = \mathbf{X} - \mathbf{t} \mathbf{p}^T$
13. $\mathbf{Y} = \mathbf{Y} - \mathbf{b} \mathbf{t} \mathbf{c}^T$
14. Parar o algoritmo caso $\|\mathbf{X}\| = 0$ (ou outro critério de parada, como por exemplo número de iterações). Caso contrário, voltar ao passo 1.

2.5 Regularização

Considere o modelo de regressão linear usual: dados n_i preditores x_1, \dots, x_{n_i} , a resposta predita da saída y é dada por:

$$\hat{y} = \theta_0 + x(1)\theta_1 + \dots + x(n_i)\theta_{n_i}$$

Ao aplicar uma técnica de ajuste de modelo, será produzido um vetor de coeficientes $\theta = [\theta_0, \theta_1, \dots, \theta_{n_i}]$ do sistema. Por exemplo, ao aplicar a técnica dos mínimos quadrados ordinária, as estimativas são obtidas ao minimizar a soma do quadrado da diferença entre os dados observados e os dados previstos pelo modelo. Existem diversos critérios

para avaliar a qualidade de um modelo, no geral os dois aspectos citados a seguir são considerados importantes:

- **Precisão do modelo** é o grau de fidelidade das predições do modelo quando dados futuros são aplicados. Isto pode ser mensurado a partir dos os índices de avaliação de modelo apresentados na seção 4.2.
- **Seleção de recursos** é uma propriedade especialmente desejável em sistemas MIMO, onde existem um grande número de preditores. Esta propriedade consiste no processo de seleção de um subconjunto de recursos (preditores) relevantes para a construção do modelo. A hipótese principal das técnicas de seleção de recurso é a existência de recursos irrelevantes ou redundantes nos dados de identificação. Recursos irrelevantes não fornecem nenhuma informação útil no cálculo da regressão e recursos redundantes são aqueles que não fornecem mais informações do que os recursos atualmente selecionados, ou seja há correlação na matriz de regressão. Os benefícios desta propriedade são:
 - Predição: Melhorar o desempenho da predição ao eliminar características ruidosas e/ou atenuar o problema de sobreajustamento (*overfitting*) via uma redução de dimensão.
 - Eficiência: Reduz consumo de memória, tempo de aprendizado, tempo de processamento. Facilita a aquisição de dados futuros ao reduzir a quantidade dos dados a serem coletados.
 - Compreensão dos dados: Identificar os fatores relevantes permite que se tenha um maior entendimento acerca das relações entre as variáveis do processo.

O erro médio quadrado (MSE) é sem dúvida o mais importante critério utilizado para avaliar o desempenho de um preditor ou um estimador (A distinção sutil entre preditores e estimadores é que as variáveis aleatórias são preditas e variáveis constantes são estimadas). O MSE também é útil para transmitir os conceitos de polarização, precisão e exatidão na estimação estatística. Dado um conjunto de entradas x , o modelo obtido pelo método de identificação irá prever as saídas $\hat{y}(x)$. O MSE pode então ser calculado como:

$$MSE = E (\|y - \hat{y}\|^2) = \frac{1}{N} \sum_{i=1}^N (y(i) - \hat{y}(i))^2 \quad (2.25)$$

No Teorema 1 é mostrado que o MSE pode ser decomposto em termos que refletem a influência da variância e da polarização na estimação.

Na regressão linear pelo método dos mínimos quadrados obtém-se a estimação dos parâmetros ao minimizar a soma dos erros quadráticos. Porém quando existe correlação

entre os preditores, a minimização da função de custo definida na técnica dos mínimos quadrados ordinária pode resultar em coeficientes de regressão excessivamente grandes, uma vez que este método tem dificuldade em quantificar a relação entre uma saída e qualquer número de preditores altamente correlacionados. Como consequência, ocorre uma inflação dos parâmetros da regressão linear estimados. Os métodos de regressão regularizada, também conhecidos como métodos de penalização ou métodos de "encolhimento", adicionam um termo de penalidade na função de custo que irá restringir a magnitude dos coeficientes estimados.

A seguir serão apresentadas três técnicas de regularização, a regressão Ridge, Lasso e Elastic Net. Esta última foi escolhida para implementar o algoritmo de busca linear EN-PH proposto. O primeiro passo nas técnicas de regularização subsequentes é normalizar as variáveis (dependentes e independentes). Isto é realizado subtraindo de cada elemento o valor da média e em seguida dividindo por seu desvio padrão. A normalização pode criar uma dificuldade na notação, uma vez que seria necessário indicar de alguma forma se as variáveis nas equações foram normalizadas ou não. Com o intuito de manter a clareza será assumido que todas as variáveis utilizadas nas técnicas de regularização foram normalizadas e se omite qualquer indicação na notação.

Teorema 1 *Considere um conjunto de vetores de entrada x_1, \dots, x_N e um vetor de saídas medidas y_i associado a cada x_i . Assumindo que a saída medida é afetada por um ruído $y_i = f(x_i) + \epsilon$, onde $f(x_i)$ é a função entrada/saída real e o ruído ϵ é tem média zero e variância σ^2 . O erro médio quadrático das saídas preditas $\hat{y} = \hat{f}(x)$ consiste na combinação da influência da polarização e da variância das estimativas:*

$$MSE = E(\|y - \hat{y}\|^2) = Bias(\hat{y}(x))^2 + Var(\hat{y}(x)) + \sigma^2 \quad (2.26)$$

Onde, $Var(\hat{y})$ é a matriz de covariância de \hat{y} e $Bias(\hat{y}) = E(\hat{y}) - y$.

Prova. De acordo com (HöSKULDSSON, 1988), o MSE pode ser decomposto entre a soma do quadrado da polarização do erro e a variância residual do erro. Uma vez que o operador $E(\cdot)$ é linear, é suficiente provar para o caso SISO (a extensão para o caso MIMO é realizada inserindo o operador $E(\cdot)$ no somatório de cada saída).

$$\begin{aligned} E[(y - \hat{y})^2] &= E[(y - f + f - \hat{y})^2] \\ &= E[(y - f)^2] + E[(f - \hat{y})^2] + 2E[(f - \hat{y})(y - f)] \\ &= E[(y - f)^2] + E[(f - \hat{y})^2] + 2(E[f\hat{y}] + E[yf] - E[y\hat{y}] - E[f^2]) \end{aligned}$$

Neste primeiro passo foi aumentada a soma com dois termos que se cancelam. Também foi usada expansão binomial e a linearidade da esperança. Observe que $y - f = \epsilon$, assim

o primeiro termo é $E[\epsilon^2] = \sigma^2$. Verifica-se então que o último termo é anulado:

- $E[y] = f$ e f é determinístico, assim $E[yf]$ simplifica para f^2 que cancela com o termo $E[f^2] = f^2$.
- $E[y\hat{y}] = E[(f + \epsilon)\hat{y}] = E[f\hat{y}] + E[\epsilon\hat{y}]$; \hat{y} é independente de ϵ e o último tem média zero, anulando este termo. Assim $E[f\hat{y}] - E[y\hat{y}] = 0$.

O resultado deste primeiro passo é:

$$E[(y - \hat{y})^2] = \sigma^2 + E[(f - \hat{y})^2]$$

Aplicando o mesmo procedimento no último termo da equação anterior teremos:

$$\begin{aligned} E[(f - \hat{y})^2] &= E[(f - E[\hat{y}] + E[\hat{y}] - \hat{y})^2] \\ &= E[(f - E[\hat{y}])^2] + E[(E[\hat{y}] - \hat{y})^2] + 2E[(E[\hat{y}] - \hat{y})(f - E[\hat{y}])] \end{aligned}$$

O último termo é anulado de forma análoga ao passo anterior. Finalmente, o termo resultante é:

$$\begin{aligned} E[(y - \hat{y})^2] &= E[(f - E[\hat{y}])^2] + E[(E[\hat{y}] - \hat{y})^2] + \sigma^2 \\ &= Bias(\hat{y})^2 + Var(\hat{y}) + \sigma^2 \end{aligned}$$

■

2.5.1 Regressão Ridge

Para controlar a variância das estimativas, utiliza-se o recurso de regularizar os coeficientes da estimação, i.e. limita-se o crescimento dos coeficientes. A restrição Ridge, ou L2, adiciona a penalidade da soma do quadrado dos parâmetros de regressão à função de custo: (HOERL; KENNARD, 1970)

$$\text{minimizar } \sum_{i=1}^n (y(i) - \hat{y}(i))^2 \text{ s.t. } \sum_{j=1}^p \theta_j^2 \leq t \quad (2.27)$$

Uma outra forma de expressar a restrição Ridge é a partir da soma penalizada do quadrado dos parâmetros de regressão (OSBORNE; PRESNELL; TURLACH, 2000):

$$\begin{aligned} SSE_{L2} &= \sum_{i=1}^n (y(i) - \hat{y}(i))^2 + \lambda \sum_{j=1}^p \theta_j^2 \\ &= (\mathbf{y} - \mathbf{X}\boldsymbol{\theta})^T (\mathbf{y} - \mathbf{X}\boldsymbol{\theta}) + \lambda \|\boldsymbol{\theta}\|_2^2 \end{aligned} \quad (2.28)$$

A terminologia "L2" representa a penalidade de segunda ordem (i.e., quadrática) utilizada na estimativa dos parâmetros. Esta penalidade apenas irá permitir que as estimativas dos parâmetros tornem-se grandes se houver uma redução proporcional na SSE. $\lambda \geq 0$ é um parâmetro de ajuste, conhecido como parâmetro de penalidade, ele controla o grau de penalidade aplicada. Este método tende a "encolher" as estimativas correlacionadas para próximo de 0 conforme λ aumenta (KUHN; JOHNSON, 2013).

Uma vez que a SSE_{L2} é convexa, ela possui solução única. Derivando e igualando a zero obtém-se:

$$\frac{\delta SSE_{L2}}{\delta \boldsymbol{\theta}} = -2\mathbf{X}^T(\mathbf{y} - \mathbf{X}\boldsymbol{\theta}) + 2\lambda\boldsymbol{\theta} = 0$$

Com isso, a solução de ridge é dada por:

$$\boldsymbol{\theta}_\lambda^{ridge} = (\mathbf{X}^T\mathbf{X} + \lambda I)^{-1}\mathbf{X}^T\mathbf{y} \quad (2.29)$$

Note que:

- Quando $\lambda = 0$, Obtem-se a estimação linear padrão $\boldsymbol{\theta}_{\lambda=0}^{ridge} = \boldsymbol{\theta}^{LS}$
- Quando $\lambda = \infty$, Tem-se $\boldsymbol{\theta}^{ridge} = 0$
- Para outros valores de λ , pondera-se a ação entre as duas ideias: ajustar o modelo linear de \mathbf{y} em \mathbf{X} e encolher os coeficientes.

A inclusão do parâmetro de encolhimento λ torna o problema não singular, mesmo no caso de $\mathbf{X}^T\mathbf{X}$ ser não inversível. Esta foi a motivação original para a regressão ridge (HOERL; KENNARD, 1970). Note que a solução é indexada pelo parâmetro λ , ou seja, para cada valor de λ obtém-se uma solução diferente. Assim, ele é responsável por controlar a magnitude dos coeficientes estimados.

Uma propriedade interessante que a regressão ridge possui é a de *seleção de grupo*. Se dois preditores são fortemente correlacionados, seus coeficientes devem ter valores próximos. Se algumas das variáveis forem idênticas, elas deverão possuir o mesmo coeficiente. Ridge desempenha bem a característica de seleção de grupo, porém não é capaz de eliminar preditores irrelevantes.

A decomposição em valores singulares (SVD) da matriz de entradas centralizada $\mathbf{X} \in \mathbb{R}^{N \times n_i}$ é dada por $\mathbf{X} = \mathbf{U}\mathbf{D}\mathbf{V}^T$. ela fornece informações interessantes da natureza da regressão ridge. Aqui $\mathbf{U} \in \mathbb{R}^{N \times n_i}$ e $\mathbf{V} \in \mathbb{R}^{n_i \times n_i}$ são matrizes ortogonais, em que as colunas de \mathbf{U} e \mathbf{V} são chamados de vetores singulares à esquerda e à direita de \mathbf{X} respectivamente. $\mathbf{D} \in \mathbb{R}^{n_i \times n_i}$ é uma matriz diagonal cujos elementos da diagonal $d_1 \geq d_2 \geq \dots \geq d_p \geq 0$ são os valores singulares de \mathbf{X} . Se um ou mais valores $d_j = 0$, \mathbf{X} é singular.

Utilizando a decomposição SVD podemos escrever o vetor ajustado pela solução dos mínimos quadrados como:

$$\begin{aligned}\mathbf{X}\boldsymbol{\theta}^{LS} &= \mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y} \\ &= \mathbf{U}\mathbf{U}^T\mathbf{y}\end{aligned}$$

Após algumas simplificações. Note que $\mathbf{U}^T\mathbf{y}$ são as coordenadas de \mathbf{y} com relação à base ortonormal \mathbf{U} .

As soluções ridge são dadas por:

$$\begin{aligned}\mathbf{X}\boldsymbol{\theta}^{ridge} &= \mathbf{X}(\mathbf{X}^T\mathbf{X} + \lambda\mathbf{I})^{-1}\mathbf{X}^T\mathbf{y} \\ &= \mathbf{U}\mathbf{D}(\mathbf{D}^2 + \lambda\mathbf{I})^{-1}\mathbf{D}\mathbf{U}^T\mathbf{y} \\ &= \sum_{j=1}^{n_i} u_j \frac{d_j^2}{d_j^2 + \lambda} u_j^T \mathbf{y}\end{aligned}$$

onde u_j são as colunas de \mathbf{U} . Note que como $\lambda \geq 0$, o termo $d_j^2/(d_j^2 + \lambda) \leq 1$. Assim como nos mínimos quadrados, ridge computa as coordenadas de \mathbf{y} com respeito à base ortonormal de \mathbf{U} . Em seguida ele atenua (encolhe) estas coordenadas pelo fator $d_j^2/(d_j^2 + \lambda)$. Assim, uma maior atenuação será realizada sobre as coordenadas do vetor da base com menor d_j^2 .

2.5.2 Regressão Lasso

Embora a regressão Ridge encolha alguns coeficientes correlacionados para próximo de 0, ele não alcança o zero absoluto para nenhum valor da penalidade. Assim, mesmo que algumas estimativas de parâmetros tornem-se negligenciavelmente pequenas, este modelo não conduz à *seleção de recursos*.

Uma alternativa popular para a regressão Ridge é o modelo de *encolhimento mínimo absoluto e seleção de operadores*, também chamado de *lasso* (TIBSHIRANI, 1996). No método lasso, considera-se o seguinte problema de otimização L_1 :

$$\min \sum_{i=1}^n (y(i) - \hat{y}(i))^2 \quad s.t. \quad \sum_{j=1}^P |\theta_j| \leq t \quad (2.30)$$

Embora pareça uma mudança pequena, as implicações práticas são significantes. Uma consequência da natureza da penalização dos valores absolutos dos coeficientes é que fazendo t suficientemente pequeno, alguns dos parâmetros serão, de fato, anulados. Assim, o lasso fornece modelos que utilizam regularização para simultaneamente melhorar o desempenho das predições e conduzir à *seleção de recursos*. Se t for escolhido maior que

$t_0 = \sum_1^p |\theta_j^{LS}|$, então as estimativas lasso serão $\boldsymbol{\theta}^{lasso} = \boldsymbol{\theta}^{LS}$. Para valores menores de t a solução dos mínimos quadrados será encolhida por algum fator. A função de custo na equação (2.30) é equivalente à função de custo penalizada da forma: (OSBORNE; PRESNELL; TURLACH, 2000)

$$\begin{aligned} SSE_{L_1} &= \sum_{i=1}^n (y(i) - \hat{y}(i))^2 + \lambda \sum_{j=1}^P |\theta_j| \\ &= (\mathbf{y} - \mathbf{X}\boldsymbol{\theta})^T (\mathbf{y} - \mathbf{X}\boldsymbol{\theta}) + \lambda \|\boldsymbol{\theta}\|_1 \end{aligned} \quad (2.31)$$

Finalmente os coeficientes lasso são obtidos a partir de:

$$\boldsymbol{\theta}_\lambda^{lasso} = \arg \min_{\boldsymbol{\theta}} (\mathbf{y} - \mathbf{X}\boldsymbol{\theta})^T (\mathbf{y} - \mathbf{X}\boldsymbol{\theta}) + \lambda \|\boldsymbol{\theta}\|_1 \quad (2.32)$$

Novamente existe o parâmetro de ajuste λ que controla a quantidade de regularização. Porém, devido à restrição L_1 ser não linear, $\boldsymbol{\theta}_\lambda^{lasso}$ não possui uma solução fechada como na regressão ridge. Sua implementação envolve técnicas de programação quadrática. O algoritmo *Least Angle Regression* (LAR) apresentado em (EFRON et al., 2004) é considerado uma versão mais "democrática" da regressão *forward stepwise* e fornece um algoritmo eficiente para o cálculo do traço lasso completo. O algoritmo LAR é descrito nos passos a seguir:

Algoritmo Least Angle Regression (LAR)

1. Normaliza-se os preditores \mathbf{X} para que tenham média zero e variância unitária e \mathbf{y} centralizado.
2. Inicia-se com o resíduo $r = y - \bar{y}$ e $\theta_1 = \theta_2 = \dots = \theta_p = 0$.
3. Encontra-se o preditor x_j mais correlacionado com r .
4. Move θ_j de 0 em direção a seus coeficientes dos mínimos quadrados $\langle x_j, r \rangle$, até que algum outro competidor x_k possua o mesmo nível de correlação com o resíduo atual que x_j possui. Se um coeficiente não-nulo atingir o zero, remover sua variável do conjunto ativo de variáveis e recalculer a direção dos mínimos quadrados conjuntos.
5. Move θ_j e θ_k na direção definida por seus coeficientes dos mínimos quadrados conjuntos do resíduo atual em (x_j, x_k) , até que algum outro competidor x_l possua o mesmo nível de correlação com o resíduo atual.
6. Repetir o procedimento anterior até que todos os p preditores tenham sido incluídos. após $\min(N - 1, p)$ passos, a solução será alcançada.

Por fim, é interessante notar que a regressão lasso não realiza seleção de grupo como a ridge. Se um grupo de preditores for altamente correlacionado lasso tende a escolher apenas um e encolher os demais para zero. Porém o lasso é capaz de eliminar variáveis irrelevantes.

2.5.3 Regressão Elastic-Net

Embora o lasso tenha mostrado sucesso em muitas aplicações, ele possui algumas limitações. Considere os três seguintes cenários:

- No caso de haver mais coeficientes que amostras, $p > N$, o lasso irá selecionar no máximo N variáveis antes de saturar, graças à natureza do problema da otimização convexa. Este é uma limitação para um método de seleção de variáveis. Além disso, o lasso não é bem definido a menos que a restrição dos coeficientes na norma L_1 seja menor que um certo valor.
- Se há um grupo de variáveis sob as quais exista um alto grau de correlação, então o lasso tende a selecionar apenas uma das variáveis do grupo e não se importa com qual das variáveis será selecionada.
- Para situações usuais, $N > P$, se existe alta correlação entre os preditores, tem sido observado empiricamente que o desempenho da regressão lasso são inferiores aos da regressão ridge (TIBSHIRANI, 1996).

O *elastic net* é uma nova técnica de regularização que atua de forma semelhante ao lasso, nas situações em que os problemas citados anteriormente não são presenciadas, e além disso remove qualquer degeneração e comportamentos abruptos causados por fortes correlações. O elastic net faz simultaneamente a seleção automática de variáveis e atenuação contínua dos regressores. Esta técnica foi comparada a uma rede de pesca extensível que retém "todos os peixes grandes" (ZOU; HASTIE, 2005).

A técnica de regularização elastic net combina as penalidades do tipo ridge e lasso, apresentadas anteriormente. A função de custo penalizada é da forma:

$$SSE_{Enet} = \frac{1}{2N} \sum_{i=1}^N (y(i) - \hat{y}(i))^2 + \lambda \left(\frac{(1 - \alpha^{EN})}{2} \|\boldsymbol{\theta}\|_2^2 + \alpha^{EN} |\boldsymbol{\theta}| \right) \quad (2.33)$$

Novamente, graças à não linearidade da penalidade sobre valor absoluto dos parâmetros, não haverá uma solução de forma fechada, e os parâmetros são obtidos a partir da minimização:

$$\boldsymbol{\theta}_\lambda^{EN} = \arg \min_{\boldsymbol{\theta}} (\mathbf{y} - \mathbf{X}\boldsymbol{\theta})^T (\mathbf{y} - \mathbf{X}\boldsymbol{\theta}) + \lambda \left(\frac{(1 - \alpha^{EN})}{2} \|\boldsymbol{\theta}\|_2^2 + \alpha^{EN} |\boldsymbol{\theta}| \right) \quad (2.34)$$

onde $\lambda \geq 0$ é o parâmetro de complexidade e $0 \leq \alpha^{EN} \leq 1$ é o parâmetro responsável por fazer com que o elastic net atue de forma intermediária entre os comportamentos ridge e lasso. Note que o elastic net torna-se o lasso quando $\alpha^{EN} = 1$ e torna-se o ridge quando $\alpha^{EN} = 0$. Conforme α^{EN} aumenta de 0 a 1, para um dado λ a esparsidade de θ_{λ}^{EN} (i.e. o número de coeficientes iguais a zero) aumenta de 0 até a esparsidade obtida no lasso.

A vantagem deste método é possibilitar uma regularização eficiente via a penalidade mais agressiva do tipo ridge com a qualidade de seleção de recursos da penalidade lasso. O elastic net encoraja o efeito de seleção de grupo visto na regressão ridge e ao mesmo tempo remove variáveis irrelevantes como no lasso. Em (ZOU; HASTIE, 2005) é mostrado que este modelo irá lidar com grupos de preditores altamente correlacionados de forma mais eficiente.

Existem diversos algoritmos que implementam a solução da equação (2.34). Neste trabalho será usado o pacote de algoritmos rápidos de código aberto Glmnet que realiza a estimação dos parâmetros de modelos lineares regularizados via lasso e elastic-net. Ele utiliza a técnica de coordenada descendente cíclica computada com o caminho de regularização (QIAN et al., 2013).

2.5.4 Obtenção de λ

Em todos os métodos de regularização apresentados foi visto que uma solução diferente é obtida para cada valor de λ utilizado. A questão que ainda não foi definida é dado um conjunto de dados de identificação, como escolher o λ ideal. Retoricamente esta questão foi abordada a partir dos caminhos de regressão e em seguida verificou-se que a melhor solução é utilizar a validação cruzada.

Caminhos das regressões

Hoerl e Kennard, os inventores da regressão ridge (HOERL; KENNARD, 1970), utilizaram um gráfico intitulado *traço ridge* para ajudar a escolher um valor apropriado para λ . Estes traços são construídos traçando os componentes de θ_{λ}^{ridge} nas ordenadas e os valores de λ nas abscissas. Ao estudar o gráfico gerado, o analista escolheria um valor para λ para o qual os coeficientes de regressão estivessem estabilizados. No geral os coeficientes variam bastante para pequenos valores de λ e em seguida se estabilizam. Além disso, deve-se escolher o menor valor possível para λ (o que introduzirá menor polarização) a partir do qual os coeficientes permanecerão constantes. Este tipo de gráfico também pode ser gerado para as regressões lasso e elastic net, e uma análise semelhante pode ser empregada.

Em (ZOU; HASTIE, 2005) é apresentado um exemplo da classificação para seleção

de genes em dados de leucemia que consistiam em 7129 genes e 72 amostras. No conjunto de dados de treinamento existiam 38 amostras divididas entre leucemia do tipo 1 e 2. Na figura 2.4 é possível comparar os caminhos ridge, lasso e elastic net. Aqui, as abscissas

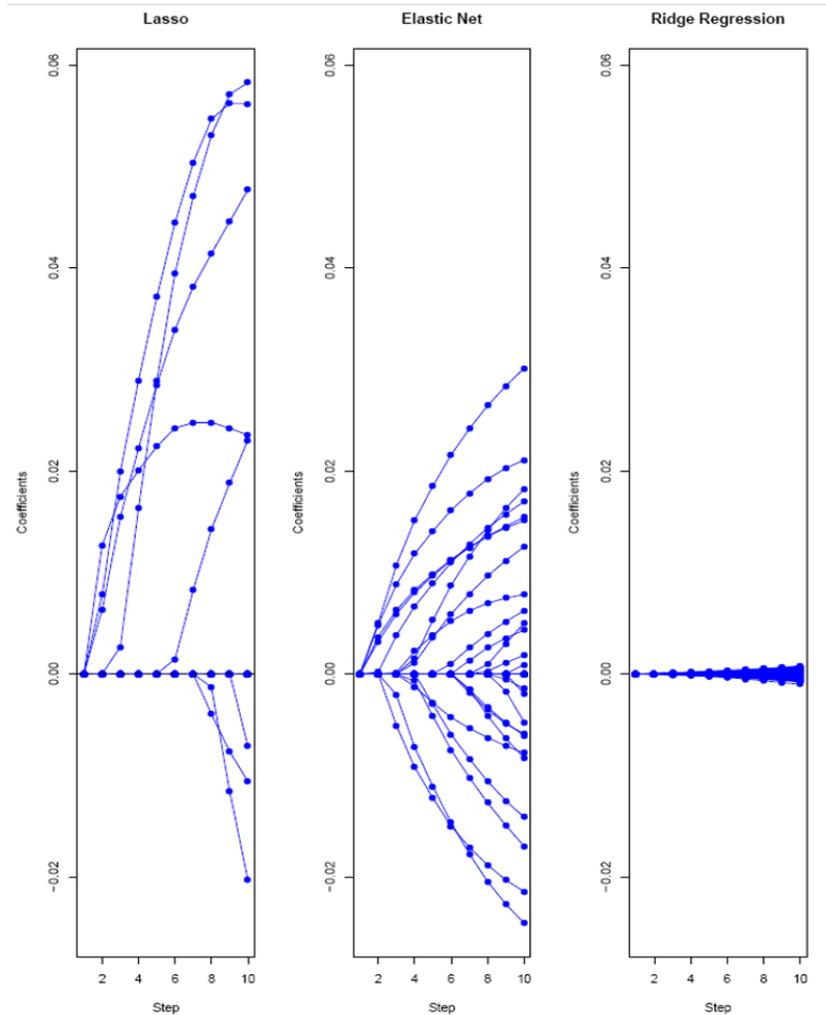


Figura 2.4: caminhos das regressões para o exemplo da leucemia

indicam a magnitude dos coeficientes estimados e as ordenadas mostram apenas os 10 primeiros passos (valores para λ). A partir destes gráficos é possível visualizar o comportamento dos três métodos. No método lasso, vários coeficientes são zerados devido à presença de correlação, porém percebe-se que suas magnitudes não foram tão atenuadas como nos demais métodos. No traço ridge, todos os coeficientes são fortemente atenuados, percebe-se o efeito da seleção de grupo onde os preditores correlacionados obtiveram coeficientes de magnitude semelhante. No traço elastic net foi utilizado $\alpha = 0,2$ e apresenta um caminho intermediário entre o lasso e ridge, ainda é possível ver o efeito da seleção de grupo neste traço.

Validação Cruzada

A estratégia de utilizar os caminhos das regressões para a escolha de λ foi duramente criticada devido à falta de objetividade (BURT; FRANK; BEATTIE, 1987). Para resolver este problema, tornou-se prática padrão utilizar a validação cruzada para selecionar λ . De acordo com (TROPSHA; GRAMATICA; GOMBAR, 2003), a validação cruzada é uma técnica de avaliação da capacidade de generalização de um modelo, a partir de um conjunto de dados. Busca-se estimar o quão preciso é este modelo na prática, i.e. avaliar o seu desempenho para um novo conjunto de dados. Assim, deve-se identificar o modelo $\hat{f}(\cdot)$ a partir de um *conjunto de treinamento* e avaliar o desempenho de suas previsões em novos dados, chamados *conjunto de teste*. Quando o número de observações é suficiente, separa-se os dados disponíveis nos conjuntos de treinamento e de teste.

A abordagem mais comum, escolhida para ser usada no presente trabalho, é a validação cruzada K-fold descrita a seguir:

- Particionar o conjunto de treinamento T em K subconjuntos mutuamente exclusivos de igual tamanho.
 - Obtendo $T = (T_1, T_2, \dots, T_K)$
 - Escolhas padrões de K são $K = 5$ e $K = 10$
- Para cada $k = 1, 2, \dots, K$, ajustar o modelo $\hat{f}_k^{(\lambda)}(x)$ ao conjunto de treinamento com exceção do k-ésimo conjunto T_k .
- Computar os valores dos parâmetros ajustados para as observações em T_k , baseado no conjunto de treinamento que excluiu este subconjunto.
- Computar o erro da validação cruzada (EVC) para o k-ésimo subconjunto:

$$(EVC)_k^{(\lambda)} = |T_k|^{-1} \sum_{(x,y) \in T_k} (y - \hat{f}_k^{(\lambda)}(x))^2 \quad (2.35)$$

- O modelo terá o erro de validação cruzada geral dado por:

$$(EVC)^{(\lambda)} = K^{-1} \sum_{k=1}^K (EVC)_k^{(\lambda)}$$

- Seleciona-se λ^* que apresentar o menor $(EVC)^{(\lambda)}$.
- Computa-se o modelo escolhido $\hat{f}_k^{(\lambda^*)}(x)$ no conjunto de treinamento completo $T = (T_1, T_2, \dots, T_K)$.

- Aplica-se $\hat{f}_k^{(\lambda^*)}(x)$ no conjunto de teste e avalia-se o erro de teste.

Na figura 2.5 apresenta-se um exemplo de curva de validação cruzada para a regressão elastic net.

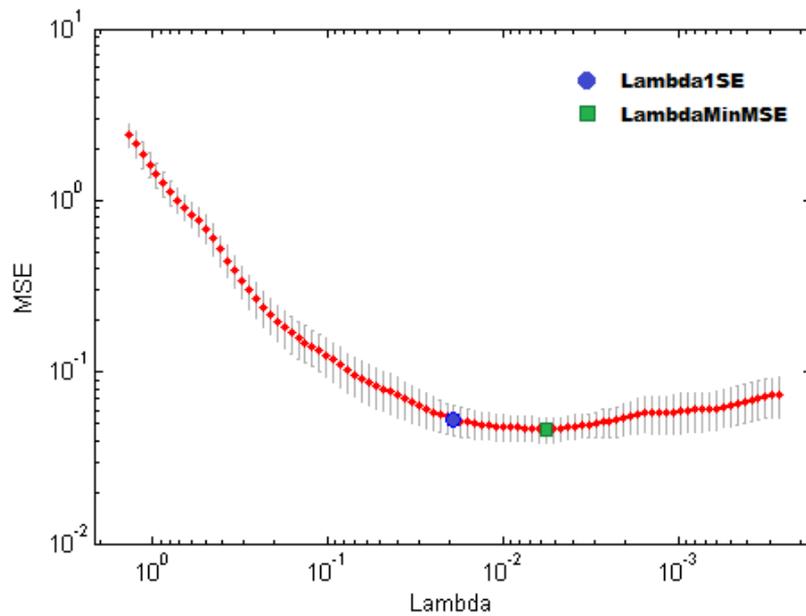


Figura 2.5: Exemplo de um gráfico de validação cruzada para a regularização

Na figura as abscissas são os valores do MSE e as ordenadas os valores de λ . Também são destacados dois pontos importantes da curva: LambdaMinMSE é o valor de λ que fornece o menor MSE; Lambda1SE é o maior valor de λ para o qual o MSE está distante de um desvio padrão de seu mínimo.

2.6 Conclusão

Neste capítulo foram apresentados os principais conceitos dos métodos de identificação relevantes ao controle preditivo. A abordagem paramétrica foi dada a partir da estrutura de modelo ARX, que foi reescrita para ser inserida na equação da regressão linear. Em seguida foi apresentada a diferença fundamental entre os métodos de identificação tradicionais e os MRI, onde o primeiro baseia-se no preditor de um passo à frente e o segundo deve contemplar todo o horizonte de predição futuro. Também foram descritos os passos necessários para a minimização da função de custo de múltiplos passos à frente.

Foi realizada uma apresentação geral dos LVM e apontadas as principais vantagens em realizar a identificação do modelo no espaço das variáveis latentes. Em seguida, apresentou-se o método de variáveis latentes conhecido como PLS. Este método realiza a transformação nas variáveis de forma a maximizar a correlação entre os escores de entrada e de saída.

Por fim foram apresentados neste capítulo os métodos de regularização e suas principais propriedades. Foi visto que ao introduzir uma penalidade no vetor de parâmetros é possível realizar uma troca entre a polarização e a variância do modelo estimado. O método de regressão Ridge utiliza uma penalização quadrática do vetor de parâmetros, o que fornece uma regularização mais agressiva. Na regressão Lasso a penalidade é dada pelo módulo do vetor de parâmetros, isto acrescenta a propriedade de seleção de recursos na regularização. O método de regressão Elastic-Net utiliza a penalidade quadrática e também do módulo do vetor de parâmetros, assim comporta-se de forma intermediária entre o método Ridge e Lasso possibilitando maior eficiência.

Capítulo 3

Identificação de sistemas baseada em otimização numérica

3.1 Introdução

Neste capítulo, a identificação de sistemas relevante ao controle preditivo será abordada sob a luz da otimização numérica. Primeiramente os parâmetros dos modelos serão estimados com base no algoritmo de busca linear iterativo chamado Least Squares - Prediction Horizon (LS-PH), que atualiza os valores dos parâmetros aplicando os mínimos quadrados para minimizar a função de custo de múltiplos passos à frente (J_{LRPI}). Em seguida este algoritmo será estendido dando origem a dois novos métodos, o PLS-PH e o EN-PH.

O algoritmo PLS-PH proposto em (PLA, 2012) é um método paramétrico MRI capaz de identificar processos que apresentam correlação entre as variáveis. A ideia fundamental do algoritmo é reduzir o fenômeno da colinearidade presente na matriz de regressores realizando a identificação no espaço das variáveis latentes.

O método EN-PH proposto neste trabalho também teve como motivação abordar os problemas causados quando há colinearidade na matriz de regressores. Porém ao invés de utilizar os LVM como foi feito no método PLS-PH, o EN-PH baseia-se nos métodos de regressão linear regularizados. No método proposto, uma busca linear é executada visando encontrar os parâmetros que minimizam a função de custo J_{LRPI} , aplicando em cada iteração a técnica Elastic-Net para encontrar a nova direção do vetor de busca.

3.2 Algoritmo Least Squares - Prediction Horizon (LS-PH)

Foi visto que J_{LRPI} é uma expressão quadrática de uma função não linear em $\boldsymbol{\theta}$, por isso deve-se recorrer à otimização não convexa para minimizá-lo. LS-PH (Least Squares Prediction Horizon) é um método de otimização numérica de busca linear baseado numa aproximação de J_{LRPI} diferente da usada em Taylor. Antes de definir a aproximação vamos reescrever J_{LRPI} na forma matricial

$$J_{LRPI} = \sum_{j=1}^{n_f} \sum_{k=1}^N \|\mathbf{e}(k+j)\|_F^2 = \|\mathbf{E}_a\|_F^2 \quad (3.1)$$

onde,

$$\mathbf{E}_a = \begin{bmatrix} \mathbf{E}_{a_1} \\ \vdots \\ \mathbf{E}_{a_{n_f}} \end{bmatrix} \quad ; \quad \mathbf{E}_{a_j} = \begin{bmatrix} \mathbf{e}(1+j) \\ \vdots \\ \mathbf{e}(N+j) \end{bmatrix}, \quad \forall j \in [1, 2, \dots, n_f]$$

\mathbf{E}_{a_j} são matrizes de dimensões $N \times n_0$, e contêm os erros de identificação baseados nas predições em $k+j$ com informações disponíveis até o instante k . A matriz de erro de predição de um passo à frente é obtida para $n_f = 1$, com isso $\mathbf{E}_a = \mathbf{E}_{a_1} = \mathbf{E}$. Cada submatriz \mathbf{E}_{a_j} é obtida de:

$$\mathbf{E}_{a_j} = \mathbf{Y}_{a_j} - \mathbf{X}_{a_j} \boldsymbol{\theta} \quad (3.2)$$

onde,

$$\mathbf{Y}_{a_j} = \begin{bmatrix} \mathbf{y}(j+1) \\ \vdots \\ \mathbf{y}(j+N) \end{bmatrix} \quad ; \quad \mathbf{X}_{a_j} = \begin{bmatrix} \mathbf{x}_j \\ \vdots \\ \mathbf{x}(j+N+1) \end{bmatrix}, \quad \forall j \in [1, 2, \dots, n_f] \quad (3.3)$$

A matriz de erro de identificação global $\mathbf{E}_a \in \mathbb{R}^{N \cdot n_f \times n_0}$ é dada por:

$$\mathbf{E}_a = \mathbf{Y}_a - \mathbf{X}_a \boldsymbol{\theta} \quad (3.4)$$

sendo,

$$\mathbf{Y}_a = \begin{bmatrix} \mathbf{Y}_{a_1} \\ \vdots \\ \mathbf{Y}_{a_{n_f}} \end{bmatrix} \quad ; \quad \mathbf{X}_a = \begin{bmatrix} \mathbf{X}_{a_1} \\ \vdots \\ \mathbf{X}_{a_{n_f}} \end{bmatrix} \quad (3.5)$$

Assim, o índice de custo pode ser expresso como:

$$J_{LRPI}(\boldsymbol{\theta}) = \|\mathbf{Y}_a - \mathbf{X}_a \boldsymbol{\theta}\|_F^2$$

Uma vez que as matrizes \mathbf{X}_{a_j} , contidas em \mathbf{X}_a dependem de $\boldsymbol{\theta} \forall j > 1$, o problema é não linear em $\boldsymbol{\theta}$. Porém é possível aproximar J_{LRPI} para uma expressão linear em $\boldsymbol{\theta}$ na vizinhança do valor atual $\boldsymbol{\theta}_k$ definindo $\mathbf{X}_a|_k$ como sendo \mathbf{X}_a computado a partir do valor $\boldsymbol{\theta}_k$:

$$\begin{aligned} \tilde{J}_{LRPI}(\boldsymbol{\theta}_k + \mathbf{p}) &= \|\mathbf{Y}_a - \mathbf{X}_a|_k(\boldsymbol{\theta}_k + \mathbf{p})\|_F^2 \\ &= \text{tr}((\mathbf{Y}_a - \mathbf{X}_a|_k(\boldsymbol{\theta}_k + \mathbf{p}))^T(\mathbf{Y}_a - \mathbf{X}_a|_k(\boldsymbol{\theta}_k + \mathbf{p}))) \\ &= \text{tr}(\mathbf{Y}_a^T \mathbf{Y}_a) - \text{tr}(\mathbf{Y}_a^T \mathbf{X}_a|_k(\boldsymbol{\theta}_k + \mathbf{p})) \dots \\ &\quad - \text{tr}((\boldsymbol{\theta}_k + \mathbf{p})^T \mathbf{X}_a|_k^T \mathbf{Y}_a) + \text{tr}((\boldsymbol{\theta}_k + \mathbf{p})^T \mathbf{X}_a|_k^T \mathbf{X}_a|_k(\boldsymbol{\theta}_k + \mathbf{p})) \end{aligned} \quad (3.6)$$

Assim como na abordagem Newtoniana, a direção de busca p é obtida igualando a primeira derivada da função de custo \tilde{J}_{LRPI} a zero:

$$\begin{aligned} \frac{\delta \tilde{J}_{LRPI}}{\delta(\boldsymbol{\theta}_k + \mathbf{p})} &= \frac{\delta \text{tr}(\mathbf{Y}_a^T \mathbf{Y}_a)}{\delta(\boldsymbol{\theta}_k + \mathbf{p})} - \frac{\delta \text{tr}(\mathbf{Y}_a^T \mathbf{X}_a|_k(\boldsymbol{\theta}_k + \mathbf{p}))}{\delta(\boldsymbol{\theta}_k + \mathbf{p})} - \frac{\delta \text{tr}((\boldsymbol{\theta}_k + \mathbf{p})^T \mathbf{X}_a|_k^T \mathbf{Y}_a)}{\delta(\boldsymbol{\theta}_k + \mathbf{p})} \dots \\ &\quad + \frac{\delta \text{tr}((\boldsymbol{\theta}_k + \mathbf{p})^T \mathbf{X}_a|_k^T \mathbf{X}_a|_k(\boldsymbol{\theta}_k + \mathbf{p}))}{\delta(\boldsymbol{\theta}_k + \mathbf{p})} \\ &= 0 - (\mathbf{Y}_a^T \mathbf{X}_a|_k)^T - \mathbf{X}_a|_k^T \mathbf{Y}_a + (\mathbf{X}_a|_k^T \mathbf{X}_a|_k + (\mathbf{X}_a|_k^T \mathbf{X}_a|_k)^T)(\boldsymbol{\theta}_k + \mathbf{p}) \\ &= -2\mathbf{X}_a|_k^T \mathbf{Y}_a + 2\mathbf{X}_a|_k^T \mathbf{X}_a|_k^T (\boldsymbol{\theta}_k + \mathbf{p}) \end{aligned} \quad (3.7)$$

$$\begin{aligned} \frac{\delta \tilde{J}_{LRPI}}{\delta(\boldsymbol{\theta}_k + \mathbf{p})} = 0 &= -2\mathbf{X}_a|_k^T \mathbf{Y}_a + 2\mathbf{X}_a|_k^T \mathbf{X}_a|_k^T (\boldsymbol{\theta}_k + \mathbf{p}) \\ (\boldsymbol{\theta}_k + \mathbf{p}) &= (\mathbf{X}_a|_k^T \mathbf{X}_a|_k^T)^{-1} \mathbf{X}_a|_k^T \mathbf{Y}_a \\ \mathbf{p}_k &= (\mathbf{X}_a|_k^T \mathbf{X}_a|_k^T)^{-1} \mathbf{X}_a|_k^T \mathbf{Y}_a - \boldsymbol{\theta}_k \end{aligned} \quad (3.8)$$

O vetor de busca \mathbf{p}_k é dado pela solução dos LS para $\mathbf{X}_a|_k$ e \mathbf{Y}_a menos o vetor de parâmetros atual, daí o nome do algoritmo LS-PH. O próximo passo é calcular o comprimento do passo α_k :

$$\alpha_k = \min_{\alpha} J_{LRPI}(\boldsymbol{\theta}_k + \alpha_k \mathbf{p}_k)$$

Uma busca linear exata de α_k é dispendiosa, e apenas requerer um decréscimo em J_{LRPI} não garante convergência global (PLA, 2012); Daí o interesse em uma busca linear inexata para determinar α_k . As condições de parada na busca linear inexata deve ser uma troca entre alcançar uma queda substancial de J_{LRPI} e não levar muito tempo fazendo a escolha. Quando a primeira derivada da função de custo está disponível, há um conjunto de condições de parada bastante difundido, tais quais as condições de Goldstein e Wolfe

(GOLDSTEIN, 1967). Entretanto não há uma concordância geral para parar a busca quando a derivada não for disponível.

A busca linear adotada é baseada na aproximação quadrática de J_{LRPI} :

$$J_{LRPI}(\boldsymbol{\theta}_k + \alpha \mathbf{p}_k) \approx a + b\alpha + c\alpha^2 \quad (3.9)$$

Note que o range de valores válidos para α é $\alpha \in [0, 1]$ como foi provado na Proposição 1. Com isso, a busca linear inexata será dada nos seguintes passos:

1. Inicialmente, J_{LRPI} é avaliado para três valores de $\alpha = [0, 0.5, 1]$
2. Os três valores de α e J_{LRPI} quando aplicados na equação (3.9), formarão um sistema de três equações e três incógnitas. Resolvendo o sistema obtêm-se os valores a , b e c .
3. Verifica-se dentre os valores de α o que minimiza J_{LRPI} . Se este valor não estiver no range $\alpha \in [0, 1]$, ele será descartado. Neste caso α será tomado como o ponto médio entre os valores anteriores de α se a parábola tiver a concavidade para cima; ou entre o valor anterior de α e 0 ou 1 caso contrário.
4. J_{LRPI} é avaliado para o novo valor de α obtido no passo 3.
5. Os menores dois valores de J_{LRPI} obtidos nas iterações anteriores mais o valor obtido no passo 4 são usados para estimar a , b e c novamente. À seguir o algoritmo retorna ao passo 3.

O procedimento anterior é repetido enquanto houver mudanças significativas de J_{LRPI} , que pode ser verificado na condição de parada à seguir:

$$|J_{LRPI}(\boldsymbol{\theta}_k + \alpha_{k-1} \mathbf{p}_k) - J_{LRPI}(\boldsymbol{\theta}_k + \alpha_k \mathbf{p}_k)| > 0.001 J_{LRPI}(\boldsymbol{\theta}_k) \quad (3.10)$$

Utilizando \mathbf{p}_k obtido na equação (3.8) e α_k com o procedimento de busca descrito acima, o novo valor de $\boldsymbol{\theta}_{k+1}$ é obtido da mesma forma que em Taylor:

$$\boldsymbol{\theta}_{k+1} = \boldsymbol{\theta}_k + \alpha_k \mathbf{p}_k \quad (3.11)$$

O algoritmo LS-PH é iterado enquanto a seguinte equação for satisfeita:

$$(J_{LRPI}(\boldsymbol{\theta}_k) - J_{LRPI}(\boldsymbol{\theta}_{k+1})) > 0.001 \quad (3.12)$$

O LS-PH é um método otimização de busca linear que apresenta os seguintes benefícios comparados às abordagens de Newton e Quasi-Newton:

- Não necessita obter a derivada de J_{LRPI} explicitamente nem utilizar aproximações. Isto é especialmente importante em problemas MIMO, onde $\boldsymbol{\theta}$ possui dimensão elevada.
- A direção de busca \mathbf{p}_k é obtida resolvendo um problema de regressão linear. Até então a forma fechada dos LS foi utilizada, porém qualquer abordagem de regressão linear pode ser usada. Na segunda parte da descrição do algoritmo PLS-PH, LS é substituído pelo PLS para trazer o potencial dos LVM para a minimização de J_{LRPI}

Proposição 1 *O range de valores válidos para α é $\alpha \in [0, 1]$*

Prova.

- α deve ser maior que zero uma vez que \mathbf{p}_k é uma direção descendente.
- Substituindo \mathbf{p}_k da equação (3.8) na equação (3.11) e assumindo que $\alpha = 1$:

$$\begin{aligned}
 \boldsymbol{\theta}_{k+1} &= \boldsymbol{\theta}_k + \alpha \mathbf{p}_k \\
 &= \boldsymbol{\theta}_k + \alpha [(\mathbf{X}_a|_k^T \mathbf{X}_a|_k)^{-1} \mathbf{X}_a|_k \mathbf{Y}_a - \boldsymbol{\theta}_k] \\
 &= \boldsymbol{\theta}_k + (\mathbf{X}_a|_k^T \mathbf{X}_a|_k)^{-1} \mathbf{X}_a|_k \mathbf{Y}_a - \boldsymbol{\theta}_k \\
 &= (\mathbf{X}_a|_k^T \mathbf{X}_a|_k)^{-1} \mathbf{X}_a|_k \mathbf{Y}_a
 \end{aligned}$$

Assim, para $\alpha = 1$, o valor de $\boldsymbol{\theta}_{k+1}$ obtido é o mínimo da aproximação \tilde{J}_{LRPI} (na vizinhança de $\mathbf{X}_a|_k$). Tem-se o melhor caso quando assume-se que $\tilde{J}_{LRPI} = J_{LRPI}$, com essa suposição o valor mínimo será atribuído para $\alpha = 1$.

■

3.3 Algoritmo Partial Least Squares - Prediction Horizon (PLS-PH)

Quando existe correlação entre o conjunto de dados de identificação, os métodos de busca comentados acima enfrentam problemas na inversão de \mathbf{H}_k ; ou no caso do LS-PH, na inversão de $\mathbf{X}_a|_k^T \mathbf{X}_a|_k$. Para lidar com este problema, no algoritmo PLS-PH, a solução pelos mínimos quadrados na equação (3.8) do vetor de busca \mathbf{p}_k é substituída pela solução

do PLS obtida em (2.22) e (2.24). Assim, o novo vetor de busca será dado por:

$$\begin{aligned}
 \mathbf{p}_k &= (\mathbf{X}_{a|k}^T \mathbf{X}_{a|k})^{-1} \mathbf{X}_{a|k}^T \mathbf{Y}_a - \boldsymbol{\theta}_k \\
 &= \mathbf{ZBQ}^T - \boldsymbol{\theta}_k \\
 &= \underbrace{\mathbf{W}(\mathbf{P}^T \mathbf{W})^{-1}}_{\mathbf{ZBQ}} - \boldsymbol{\theta}_k
 \end{aligned} \tag{3.13}$$

onde \mathbf{W} , \mathbf{P} , \mathbf{B} e \mathbf{Q} são obtidos aplicando o algoritmo PLS em $\mathbf{X}_{a|k}$ e \mathbf{Y}_a . A vantagem do método PLS-PH é utilizar PLS para minimizar J_{LRPI} , possibilitando realizar a identificação paramétrica também no caso de dados de identificação mal condicionados.

3.4 Algoritmo Elastic Net - Prediction Horizon (EN-PH)

O método EN-PH proposto neste capítulo teve como motivação abordar os problemas causados quando há colinearidade na matriz de regressores. Porém ao invés de utilizar os LVM como foi feito no método PLS-PH em (PLA, 2012), o EN-PH baseia-se nos métodos de regressão linear regularizada. No método proposto, uma busca linear é executada visando encontrar os parâmetros que minimizam a função de custo J_{LRPI} , aplicando em cada iteração a técnica Elastic-Net para encontrar a nova direção do vetor de busca.

O método EN-PH proposto neste trabalho, assim como o PLS-PH, pode ser dividido em duas partes. A primeira parte é formada pelo algoritmo LS-PH apresentado na seção 3.2, que consiste em uma busca linear cujo a equação do vetor de busca baseia-se na solução dos mínimos quadrados. Porém diferente do PLS-PH, onde o termo da solução dos mínimos quadrados foi substituído pela solução do PLS na equação (3.8) do vetor de busca \mathbf{p}_k , o EN-PH irá encontrar a nova direção de busca \mathbf{p}_k aplicando o método de regularização elastic net sob matrizes de dados de identificação. O método de regularização foi implementado a partir do pacote de código aberto Glmnet, que são algoritmos eficientes capazes de obter os regressores lineares a partir dos caminhos Lasso ou Elastic Net (QIAN et al., 2013).

Para escolher o melhor valor do parâmetro de penalização λ na equação (2.34), utilizou-se o método de validação cruzada K-fold descrito na seção (2.5.4), no qual a saída predita $\hat{f}_k^{(\lambda)}(x) = \boldsymbol{\theta}_\lambda^{EN} x$ e o número de folds escolhido foi $K = 10$. Esta é um valor bastante comum na validação cruzada K-fold e sua escolha foi justificada ao comparar o desempenho dos métodos variando K de 1 a 10.

Conforme descrito anteriormente, o método elastic net faz um balanceamento entre o comportamento ridge e lasso a partir do valor de α^{EN} escolhido. Também foi proposta

neste trabalho uma rotina de busca iterativa para encontrar o melhor valor de α para os dados de identificação que estiverem sendo analisados. A estratégia consiste nos seguintes passos:

1. Aplica-se o método de regularização elastic net implementado no pacote Glmnet, utilizando os dados $\mathbf{X}_{a|k}$ e \mathbf{Y}_a da equação (3.8) para todos os valores de α^{EN} no range $\alpha^{EN} \in [0, 1]$ com o passo de 0,005. Após diversos testes, foi visto que este valor de passo forneceu um desempenho satisfatório e não sobrecarregou o processamento do algoritmo.
2. Armazena-se em um vetor os valores dos erros médios da validação cruzada obtidos para $\lambda = \lambda_{1se}$.
3. Em seguida, busca-se no vetor armazenado no passo anterior qual valor ótimo de alfa, α_{opt}^{EN} , sendo este o que fornece o menor erro médio da validação cruzada.
4. Aplica-se novamente o algoritmo elastic net sob os dados de identificação utilizando $\alpha = \alpha_{opt}^{EN}$ e $\lambda = \lambda_{1se}$, obtendo assim os coeficientes $\boldsymbol{\theta}^{EN}$.

Uma vez obtidos os coeficientes $\boldsymbol{\theta}^{EN}$, o vetor de busca do método EN-PH será dado por:

$$\begin{aligned} \mathbf{p}_k &= \underbrace{(\mathbf{X}_{a|k}^T \mathbf{X}_{a|k})^{-1} \mathbf{X}_{a|k}^T \mathbf{Y}_a}_{EN} - \boldsymbol{\theta}_k \\ &= \boldsymbol{\theta}_{\lambda}^{EN} - \boldsymbol{\theta}_k \end{aligned} \quad (3.14)$$

A partir da expressão do vetor de busca modificado, algoritmo de busca linear prossegue com os passos descritos na seção 3.2, onde calcula-se o comprimento do passo α pela busca inexata e em seguida, itera-se os parâmetros a partir da equação (3.11). Também foram utilizados os mesmos critérios de parada.

Em seguida é apresentado o pseudo-código do algoritmo EN-PH:

1. Seleciona-se α usando \mathbf{X} e \mathbf{Y} no método Elastic Net a partir de uma busca iterativa. (Varia-se α no intervalo $[0, 1]$ com o passo de 0.01 e será escolhido o valor que fornecer o menor MSE)
2. Seleciona-se λ utilizando a técnica de validação cruzada K-fold com X , Y e $K = 10$.
3. $\boldsymbol{\theta}_k$ é obtido aplicando α e λ fornecidos nos passos anteriores.
4. Forma-se \mathbf{Y}_a e obtem-se $\mathbf{X}_{a|k}$ utilizando $\boldsymbol{\theta}$.

5. Obtem-se \mathbf{p}_k utilizando a equação (3.14). Os valores de α e λ devem ser recalculados a cada iteração de forma idêntica à descrita nos passos 1 e 2.
6. Inicializa-se $\vec{\alpha} = [0 \quad 0.5 \quad 1]$
7. Estima-se a, b e c na equação (3.9) com os valores inicializados no passo anterior.
8. Encontrar o ponto onde a primeira derivada da equação (3.9) é zerada.

$$\alpha' = \frac{-b}{2c}$$

9. Decidir o novo valor de α_k : Se $c > 0$ (parábola para cima)
 - Se $0 < \alpha' < 1$

$$\alpha_k = \alpha'$$
 - Caso contrário
 - Se $\alpha' \leq 0$

$$\alpha_k \text{ recebe o ponto médio entre os pontos } 0 \text{ e } \min_{\alpha}(\vec{\alpha}), s.t. \alpha > 0.$$
 - Caso contrário

$$\alpha_k \text{ recebe o ponto médio entre os pontos } \max_{\alpha}(\vec{\alpha}), s.t. \alpha < 1. \text{ e } 1$$
 - Caso contrário
 - Se $\alpha' > \vec{\alpha}(1, 1)$

$$\alpha_k \text{ recebe o ponto médio entre os pontos } 0 \text{ e } \min_{\alpha}(\vec{\alpha}), s.t. \alpha > 0.$$
 - Caso contrário

$$\alpha_k \text{ recebe o ponto médio entre os pontos } \max_{\alpha}(\vec{\alpha}), s.t. \alpha < 1. \text{ e } 1$$
10. O valor de $\vec{\alpha}$ com o maior J_{LRPI} é removido e α_k é inserido em $\vec{\alpha}$.
11. Se a equação (3.10) mantiver-se verdadeira, repetir a partir do passo 7, caso contrário ir para o passo 12.
12. α_k é selecionado como o valor em $\vec{\alpha}$ com o menor J_{LRPI} .
13. $\theta_{k+1} = \theta_k + \alpha_k \mathbf{p}_k$
14. Finalizar caso o critério de parada da equação (3.12) for atingido; Caso contrário, $k = k + 1$ e voltar ao passo 3.

3.5 Conclusão

Neste capítulo foram visto três algoritmos de busca que possuem estratégias semelhantes, diferenciando-se entre si com o método utilizado para encontrar a direção de busca dos parâmetros estimados. O LS-PH é o algoritmo de busca linear que utiliza a solução dos mínimos quadrados para atualizar a direção de busca; Já o método PLS-PH estende o LS-PH ao substituir a solução dos mínimos quadrados pela solução obtida com o algoritmo PLS. Em (PLA, 2012) verificou-se que o desempenho do método PLS-PH foi superior aos demais métodos comparados, especialmente em presença de correlação entre preditores. Por fim o método proposto EN-PH foi definido aplicando o método de regularização Elastic-Net para atualizar a direção do vetor de busca.

Capítulo 4

Avaliação das técnicas de identificação - Simulação/Experimento

4.1 Introdução

Neste capítulo são apresentados exemplos de identificação de sistemas a fim de avaliar o algoritmo EN-PH proposto. Esta avaliação será realizada com base em uma análise comparativa do desempenho preditivo dos modelos que foram obtidos utilizando os seguintes métodos de identificação:

- **LS** é a abordagem mais simples onde apenas o índice J_{OSAPI} é minimizado. A comparação com o LS pode justificar a utilização de algoritmos mais sofisticados que minimizam J_{LRPI} , uma vez que o modelo destina-se a ser usado no MPC.
- **EN** Utiliza-se o pacote Glnet para resolver o problema de otimização definido em (2.34) e obter os parâmetros pela regularização. A comparação com o método EN é interessante para comprovar os benefícios de sofisticar a solução da regularização a partir do método de busca linear proposto.
- **LM** (Levenberg Marquardt) é considerado um método de otimização numérica de região de confiança. Neste método MRI é utilizado o algoritmo LM implementado no toolbox de otimização do Matlab[®] para minimizar J_{LRPI} . Este método foi inserido a fim de analisar o desempenho do novo algoritmo em comparação com outra técnica MRI.
- **PLS-PH** é o método de otimização numérica de busca linear que minimiza o J_{LRPI} utilizando o PLS, para realizar MRI mesmo em casos de dados de identificação mal-condicionados.

- **EN-PH** é o método de otimização numérica de busca linear apresentado neste capítulo. Ele baseou-se no algoritmo PLS-PH, porém ao invés de usar os métodos de variáveis latentes, o EN-PH utiliza os métodos de regressão regularizados para lidar com a colinearidade na matriz de regressores.

Foram considerados dois exemplos de simulação e um experimental. No primeiro exemplo os métodos de identificação foram aplicados ao modelo da coluna de destilação de água/metanol reportada em (WOOD; BERRY, 1973). Em seguida, o desempenho dos modelos foi comparado na identificação do processo da golfada utilizando o modelo não-linear descrito em (JAHANSHAH, 2013). Por último uma planta térmica de escala laboratorial foi identificada e os resultados comparados. A comparação dos desempenhos dos métodos foi realizada utilizando três índices de desempenho descritos na seção seguinte.

4.2 Indicadores de validação dos modelos

Três índices de validação foram utilizados para avaliar o desempenho preditivo dos modelos obtidos por cada técnica de identificação. Os índices descritos a seguir foram calculados usando experimentos de Monte Carlo, que é a técnica de repetir as simulações com novas entradas aleatórias um elevado número de vezes, e em seguida tomar a média dos resultados dos experimentos. Com isso pretende-se ganhar uma maior representabilidade das entradas na validação dos modelos.

4.2.1 $J_{LRPI_{EV}}$

Este índice é usado para avaliar o desempenho dos modelos em termos de MRI. $J_{LRPI_{EV}}$ equivale ao valor de J_{LRPI} calculado na equação (2.17) aplicando os dados de validação externos.

4.2.2 R^2

O indicador R^2 é uma matriz $n_f \times n_o$ contendo os valores dos coeficientes de determinação para dados externos. Cada elemento em R^2 é obtido usando a equação (4.1). R^2 é calculado para visualizar a evolução da performance dentro do horizonte de predição.

$$R_s^2(j) = 1 - \frac{E_{a_{js}}^T E_{a_{js}}}{(Y_{a_{js}} - E\{Y_{a_{js}}\})^T (Y_{a_{js}} - E\{Y_{a_{js}}\})} \quad \forall j \in [1, n_f], \forall s \in [1, n_o] \quad (4.1)$$

Onde $E_{a_{js}}$ é a coluna s da matriz E_{a_j} da equação (3.2); $Y_{a_{js}}$ é a coluna s de Y_{a_j} da equação (3.3); e $E\{\cdot\}$ é o operador da média.

4.2.3 Coeficientes de Theil

Em (THEIL, 1959) foram apresentados os indicadores que baseiam-se no erro de predição, ou seja, na diferença entre os pares de valores observados e estimados. Nestes índices, a soma do quadrado do erro de predição é decomposta em vários componentes possibilitando avaliar separadamente diferentes aspectos estatísticos da predição.

Índice padrão

U , também chamado coeficiente de desigualdade de Theil, fornece uma medida normalizada do erro relativo que reduz o efeito de grandes erros, e sua equação é apresentada a seguir:

$$U = \frac{\sqrt{\frac{1}{N} \sum_{k=1}^N (\hat{y}(k) - y(k))^2}}{\sqrt{\frac{1}{N} \sum_{k=1}^N \hat{y}^2(k)} + \sqrt{\frac{1}{N} \sum_{k=1}^N y^2(k)}} \quad (4.2)$$

onde, \hat{y} é o valor predito (estimado); y é o valor observado (medido); N é o número de observações (medições). O coeficiente padrão está no intervalo $U \in [0, 1]$. Quanto mais próximo de zero, mais preciso o modelo.

Índice de polarização

U_B , também chamado grau de precisão ou exatidão, que é dado por:

$$U_B = \frac{(\bar{\hat{y}} - \bar{y})^2}{\frac{1}{N} \sum_{k=1}^N (\hat{y}(k) - y(k))^2} \quad (4.3)$$

onde, $\bar{\hat{y}}$ é a média dos valores estimados; \bar{y} é a média dos valores observados. O índice de polarização está no intervalo $U_B \in [0, 1]$. Quanto mais próximo de zero, menos polarizada será a estimativa.

Índice de variância

U_V , também chamado grau de consistência, é dado por:

$$U_V = \frac{(\sigma_{\hat{y}} - \sigma_y)^2}{\frac{1}{N} \sum_{k=1}^N (\hat{y}(k) - y(k))^2} \quad (4.4)$$

onde, $\sigma_{\hat{y}}$ é o desvio padrão dos valores estimados; σ_y é o desvio padrão dos valores medidos. O índice de variância também está no intervalo $U_V \in [0, 1]$. Quanto mais próximo de zero for U_V , menor será variação entre os valores das estimativas e das medições.

O **Índice de covariância** U_C , também chamado grau de associação, é dado por:

$$U_C = \frac{2(1 - \rho)\sigma_{\hat{y}}\sigma_y}{\frac{1}{N} \sum_{k=1}^N (\hat{y}(k) - y(k))^2} \quad (4.5)$$

onde, ρ é o coeficiente de correlação dado por:

$$\rho = \frac{1}{\sigma_{\hat{y}}\sigma_y N} \sum_{k=1}^N (\hat{y}(k) - \bar{\hat{y}})(y(k) - \bar{y})$$

O índice de covariância também está no intervalo $U_V \in [0, 1]$. Porém, Quanto mais próximo de um for U_V , mais próximos serão os valores das estimativas e das medições (mais correlacionados), e assim melhor será a estimativa.

4.3 Coluna de destilação

O modelo da coluna de destilação de metanol/água, reportada em (WOOD; BERRY, 1973) é um sistema MIMO típico com fortes interações entre as variáveis controladas. As entradas do sistema são as taxas dos fluxos de vapor de refluxo (u_1) e do refeedor (u_2). As saídas do processo são as composições dos produtos no topo (y_1) e na base (y_2). As taxas do fluxo de alimentação (ξ_1) e da composição da alimentação (ξ_2) são considerados como perturbações no sistema. O diagrama esquemático da coluna de destilação é apresentado na figura 4.1:

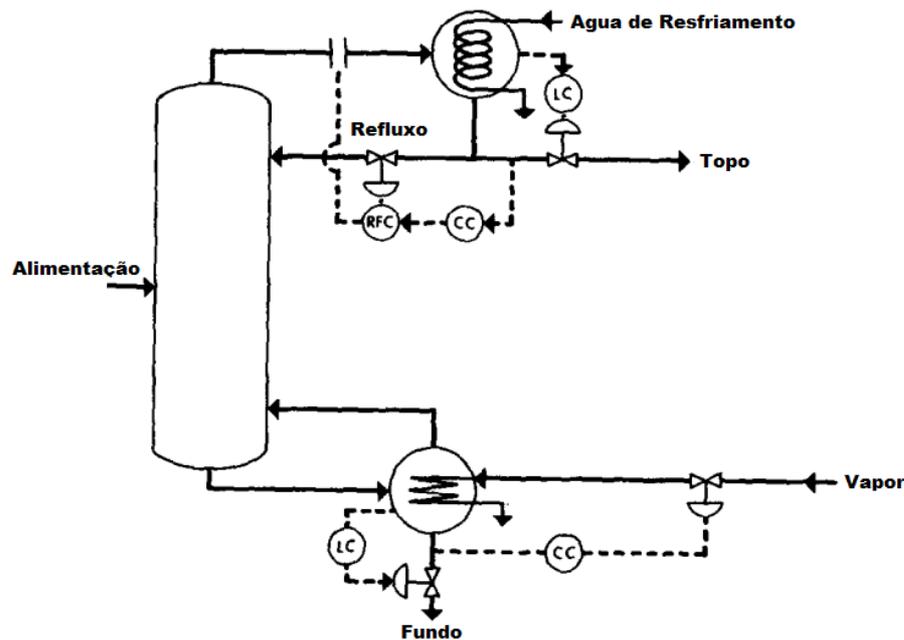


Figura 4.1: Diagrama esquemático da coluna de destilação

Os atrasos presentes no processo foram inseridos a partir da aproximação Pade de

segunda ordem. O processo discretizado com $T_s = 1$ minuto pode ser expresso como:

$$\begin{aligned} \begin{pmatrix} y_1(k) \\ y_2(k) \end{pmatrix} &= \begin{pmatrix} \frac{0,06z^{-1}+0,63z^{-2}+0,072z^{-3}}{1-0,93z^{-1}-0,013z^{-2}-0,0023z^{-3}} & \frac{0,033z^{-1}+0,087z^{-2}-0,58z^{-3}}{1-1,57z^{-1}+0,72z^{-2}-0,13z^{-3}} \\ \frac{0,2z^{-1}-0,63z^{-2}+0,52z^{-3}}{1-2,18z^{-1}+1,58z^{-2}-0,39z^{-3}} & \frac{0,051z^{-1}+0,12z^{-2}-0,85z^{-3}}{1-1,55z^{-1}+0,71z^{-2}-0,13z^{-3}} \end{pmatrix} \begin{pmatrix} u_1(k) \\ u_2(k) \end{pmatrix} \\ &+ \begin{pmatrix} \frac{0,1z^{-1}-0,3z^{-2}+0,23z^{-3}}{1-2,29z^{-1}+1,74z^{-2}-0,45z^{-3}} & \frac{0,0037z^{-1}-0,011z^{-2}+0,0087z^{-3}}{1-2,28z^{-1}+1,72z^{-2}-0,44z^{-3}} \\ \frac{0,003z^{-1}-0,095z^{-2}+0,25z^{-3}}{1-1,65z^{-1}+0,84z^{-2}-0,16z^{-3}} & \frac{0,0052z^{-1}-0,014z^{-2}+0,01z^{-3}}{1-2,34z^{-1}+1,83z^{-2}-0,48z^{-3}} \end{pmatrix} \begin{pmatrix} \xi_1(k) \\ \xi_2(k) \end{pmatrix} \end{aligned} \quad (4.6)$$

Os sinais de excitação para o processo escolhidos foram sinais randômicos Gaussianos (RGS) e foram gerados utilizando a função *idinput* do Matlab[®] com $BAND = [0, 0.01]$. A amplitude do ruído branco foi escolhida para que o SNR seja aproximadamente 10dB para ambas as saídas. O conjunto de dados de identificação gerados é apresentado na figura 4.2, estes sinais foram normalizados antes de sua utilização nos algoritmos de identificação.

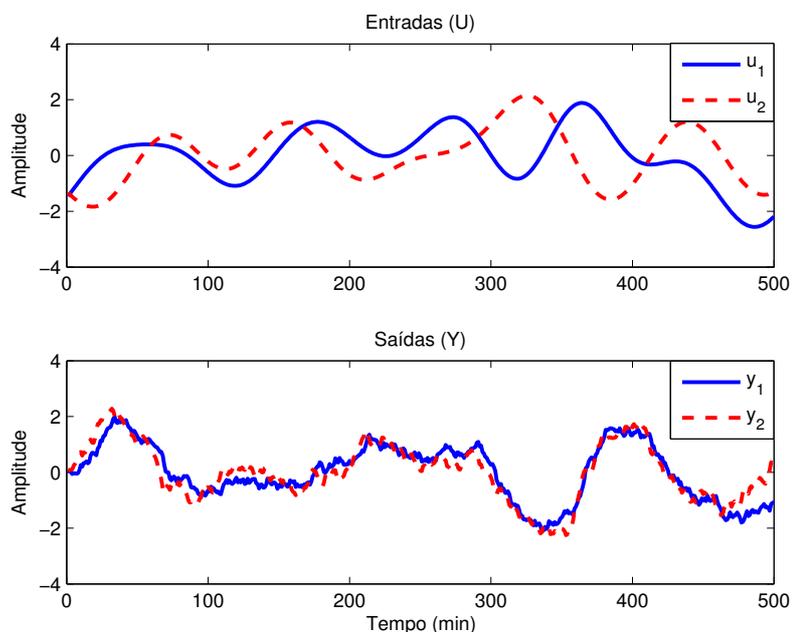


Figura 4.2: Conjunto de dados de identificação da coluna de destilação

Foram utilizados três modelos MIMO ARX, de ordens $n_a = n_b = 2, 3, 4$ para aproximar o processo. O horizonte de predição foi definido como $n_f = 30$. Após a obtenção dos parâmetros θ por cada método de identificação, é aplicada uma nova entrada aleatória RGS de validação a fim de avaliar o desempenho do preditor.

Nas figuras (4.3),(4.4) e (4.5) são apresentadas as curvas dos indicadores de validação $R_{30 \times 2}^2$, que consiste no valor de R^2 obtido aplicando a equação (4.1) para cada saída e varrendo j por todo o intervalo de predição $j \in [0, 30]$. Os valores do índice $J_{LRPI_{EV}}$ para cada método de identificação também foram calculados e inseridos nas legendas das figuras.

Verifica-se nas curvas de $R_{30 \times 2}^2$ das figuras (4.3),(4.4) e (4.5) que o algoritmo EN-PH

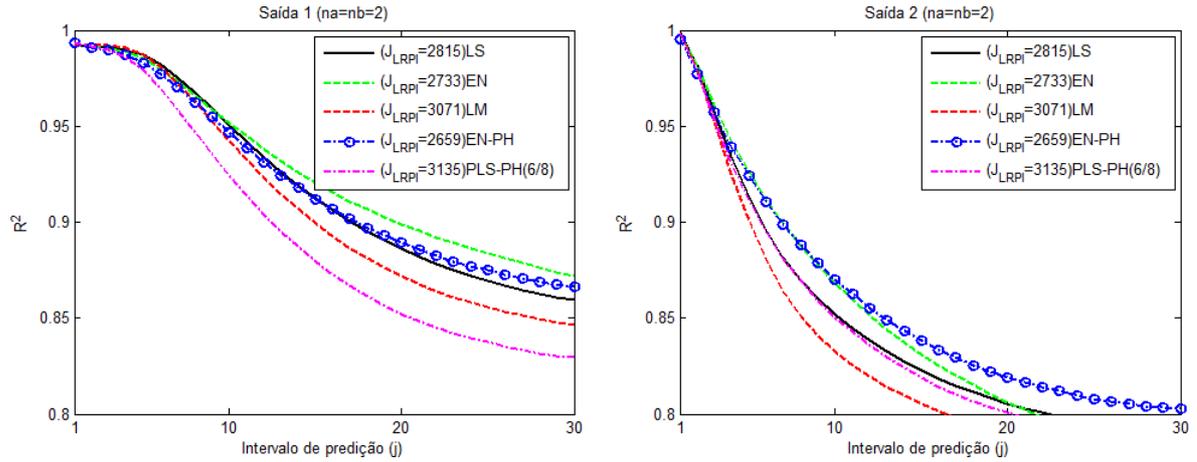


Figura 4.3: Curvas de R^2 para modelo ordem 2 - Coluna de destilação

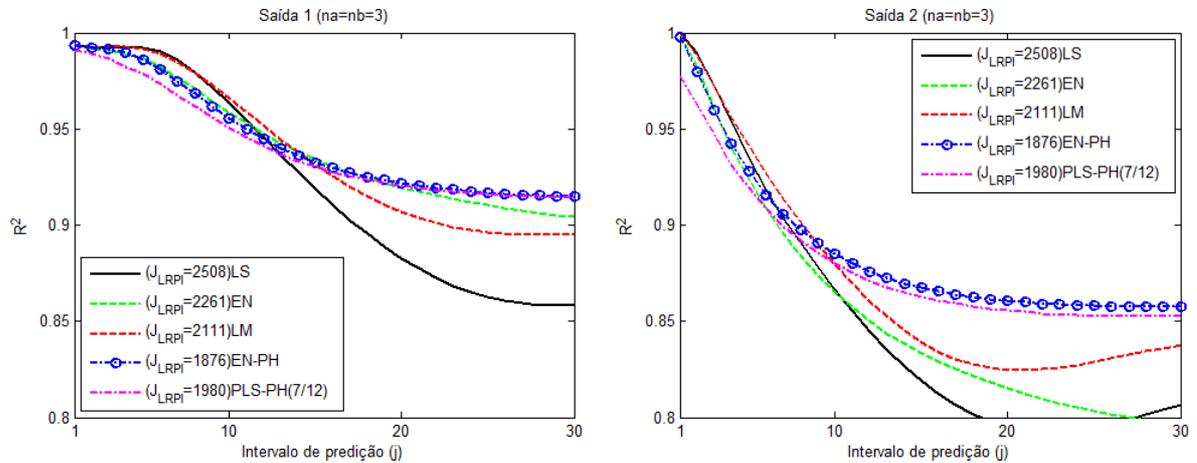


Figura 4.4: Curvas de R^2 para modelo ordem 3 - Coluna de destilação

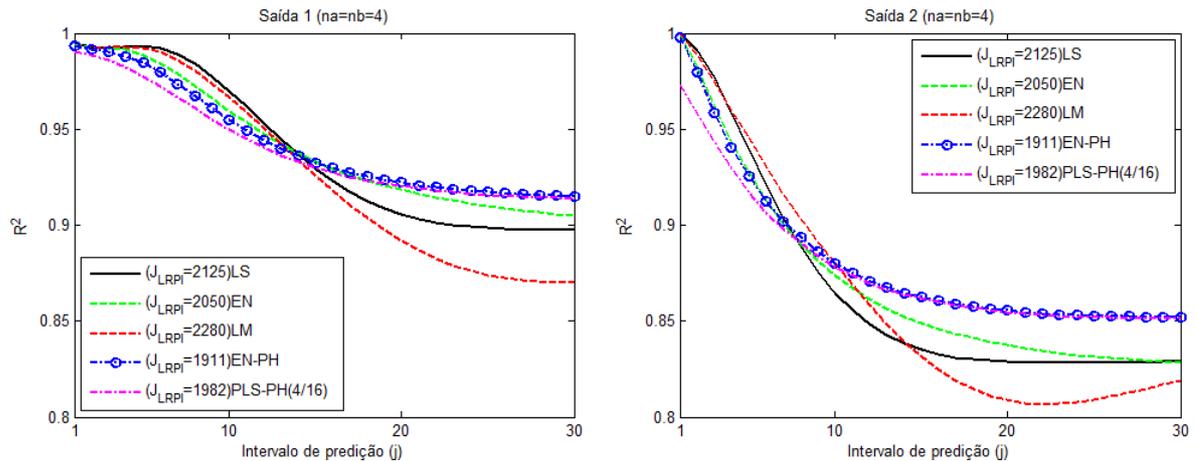


Figura 4.5: Curvas de R^2 para modelo ordem 4 - Coluna de destilação

apresentou o melhor desempenho entre os métodos de identificação analisados, para as ordens dos modelos sugeridas. O melhor ajuste do modelo foi obtido para $n_a = n_b = 3$, onde $J_{LRPI_{EV}} = 1876$. Nota-se que, no geral, o método LS é superado por todos os

Tabela 4.1: Índices de Theil de validação - Coluna de destilação

| Saída 1 | | | | | | |
|-----------|---------|----------------|----------------|----------------|----------------|----------------|
| n_a/n_b | Índice | LS | EN | LM | EN-PH | PLS-PH |
| 2 | U | 0,13225 | 0,11999 | 0,11623 | 0,11368 | 0,12219 |
| | U_b | 0,00588 | 0,01264 | 0,01322 | 0,02108 | 0,01609 |
| | U_v | 0,04434 | 0,04314 | 0,02067 | 0,01573 | 0,02116 |
| | U_c | 0,94978 | 0,94422 | 0,96611 | 0,96319 | 0,96275 |
| 3 | U | 0,17892 | 0,12228 | 0,12629 | 0,11890 | 0,12402 |
| | U_b | 0,04234 | 0,02005 | 0,01782 | 0,02247 | 0,02131 |
| | U_v | 0,03481 | 0,02380 | 0,02069 | 0,02165 | 0,02703 |
| | U_c | 0,92285 | 0,95616 | 0,96149 | 0,95588 | 0,95166 |
| 4 | U | 0,12359 | 0,11597 | 0,13636 | 0,10889 | 0,11223 |
| | U_b | 0,04951 | 0,01330 | 0,05069 | 0,02358 | 0,01801 |
| | U_v | 0,02725 | 0,06147 | 0,03473 | 0,02271 | 0,02288 |
| | U_c | 0,92324 | 0,92523 | 0,91458 | 0,95371 | 0,95911 |
| Saída 2 | | | | | | |
| n_a/n_b | Índices | LS | EN | LM | EN-PH | PLS-PH |
| 2 | U | 0,21538 | 0,18165 | 0,18019 | 0,17956 | 0,18788 |
| | U_b | 0,00730 | 0,01258 | 0,01520 | 0,02120 | 0,01775 |
| | U_v | 0,12542 | 0,15273 | 0,05301 | 0,05296 | 0,09093 |
| | U_c | 0,86728 | 0,83469 | 0,93179 | 0,92584 | 0,89132 |
| 3 | U | 0,24926 | 0,18485 | 0,20680 | 0,17750 | 0,18027 |
| | U_b | 0,04732 | 0,01671 | 0,02271 | 0,01825 | 0,01953 |
| | U_v | 0,04387 | 0,03373 | 0,03965 | 0,03044 | 0,03242 |
| | U_c | 0,90881 | 0,94956 | 0,93764 | 0,95131 | 0,94805 |
| 4 | U | 0,18113 | 0,18214 | 0,19286 | 0,15928 | 0,16290 |
| | U_b | 0,05593 | 0,01594 | 0,05833 | 0,03142 | 0,02738 |
| | U_v | 0,03508 | 0,08976 | 0,03601 | 0,02615 | 0,04137 |
| | U_c | 0,90900 | 0,89430 | 0,90567 | 0,94243 | 0,93125 |

métodos MRI como já era esperado, uma vez que os últimos são técnicas mais sofisticadas e foram criadas almejando a previsibilidade do modelo múltiplos passos à frente.

Uma comparação mais detalhada sobre algumas propriedades dos modelos de regressão obtidos neste exemplo é realizada com base nos coeficientes de Theil apresentados na tabela 4.1, onde os melhores resultados para cada índice e em todas as ordens foram destacados em negrito.

Verifica-se na tabela 4.1 que o método EN-PH forneceu os melhores índices de Theil padrões U para todas as ordens de modelo analisadas, o que era de se esperar uma vez que este índice pode ser visto como uma normalização do $J_{LRPI_{EV}}$. O efeito da utilização da regularização pode ser verificado ao comparar os índices U_b e U_v . No geral, o método EN-PH forneceu estimativas com polarização ligeiramente maior que o PLS-PH, porém a variância dessas estimativas foi reduzida. Esta troca entre polarização e variância é a propriedade principal da regularização e, como pôde ser comprovado neste exemplo, é

justificada por fornecer um erro médio quadrático geral menor. Por fim, o método EN-PH forneceu, no geral, estimativas com maior associação entre as saídas estimadas e as medidas, o que pode ser observado a partir dos valores do índice U_c .

4.4 Regime de fluxo de golfadas em prospecção offshore de petróleo

Neste segundo exemplo os métodos de identificação foram aplicados ao modelo não linear do sistema riser em regime de golfadas descrito em (JAHANSHAHI, 2013) a fim de obter um modelo linear que descreve este processo. A seção inicia-se com uma discussão acerca das características físicas do escoamento multifásico na prospecção de petróleo *offshore*. Em seguida o modelo não linear é descrito em detalhes. Por fim são apresentados os resultados da identificação do processo a partir dos métodos sugeridos.

4.4.1 Prospecção de petróleo offshore

Na prospecção de petróleo *offshore*, ver figura 4.6, as plataformas petrolíferas são conectadas aos poços submarinos através de tubulações. Denomina-se risers a tubulação vertical e oleodutos as horizontais. Nestes sistemas, as tubulações representam custos significativos. Além disso, muitos critérios como por exemplo profundidade, temperatura e geometria devem ser considerados durante o projeto (STORKAAS; SKOGESTAD, 2003). No apêndice B é apresentado o fenômeno da golfada em detalhes.

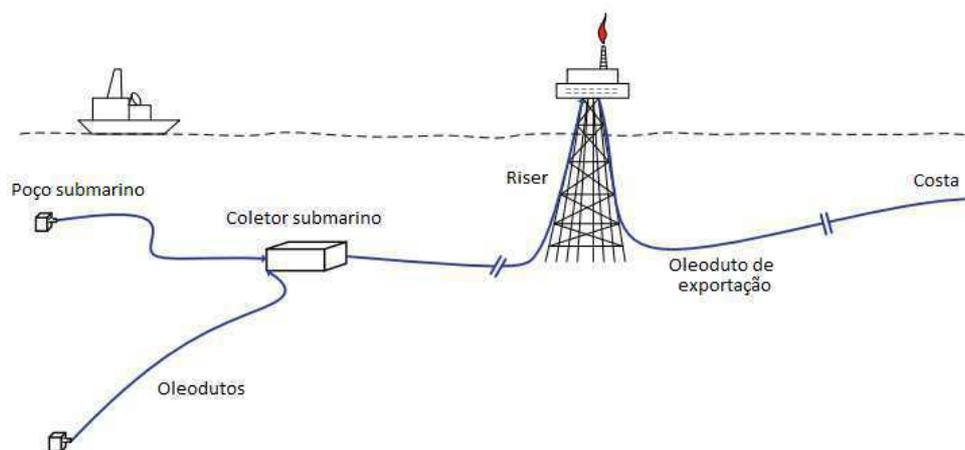


Figura 4.6: Representação de um sistema de oleodutos e riser

4.4.2 Identificação a partir do simulador

A configuração do sistema riser usada neste exemplo foi baseada na estrutura utilizada em (JAHANSHAH, 2013). Nesta configuração, o diâmetro do oleoduto é 0,12m e seu comprimento é 4300m. A partir da entrada do processo, os primeiros 2000m do oleoduto consistem de uma tubulação horizontal, e nos 2300m restantes há uma inclinação descendente com um ângulo de 1°. Esta inclinação faz com que a tubulação tenha uma descida de 40,14m de altura em relação ao ponto de entrada, criando assim um ponto baixo, ou ponto de acumulação, no final do oleoduto. A partir deste ponto baixo inicia a tubulação do riser, que consiste numa tubulação vertical com 300m de comprimento e 0,1m de diâmetro. Do topo do riser segue uma tubulação horizontal com 100m de comprimento e com o mesmo diâmetro que a tubulação do riser, esta tubulação liga o riser à válvula de saída.

A alimentação do sistema (fluxo multifásico extraído dos poços de petróleo) possui um fluxo nominalmente constante de 9kg/s. O fluxo da mistura contém uma contribuição da fase líquida com $\omega_{l,in} = 8,64kg/s$ e da fase gasosa com $\omega_{g,in} = 0,36kg/s$. A pressão do separador (P_s) após a válvula de saída, é considerada nominalmente constante com valor de 50,1bar. Com estas considerações, a abertura da válvula de saída representa o único grau de liberdade de controle do sistema (variável manipulada). Para o presente caso de estudo, o valor crítico da abertura relativa da válvula de saída é $Z_1^* = 5\%$, este é o valor onde ocorre a transição entre o regime de fluxo estável não-oscilatório e o regime de golfadas.

A resposta dinâmica do modelo não linear, descrita na seção B.2, foi utilizada para obter os dados de identificação. Inicialmente foi testado aplicar um sinal de excitação RGS para a abertura da válvula Z , com diversas combinações de bandas e amplitudes, porém os índices de custo de identificação obtidos para esta excitação foram muito altos e insatisfatórios. Verificou-se em seguida, que melhores resultados poderiam ser obtidos neste processo utilizando o sinal binário pseudo-randômico (PRBS), com $band = [0\ 0\ 03]$ e $levels = [0,06\ 0,065]$ (Irã variar a abertura da válvula binariamente entre 6% e 6,5% de forma aleatória). O sinal foi gerado utilizando a função *idinput* do Matlab® e é apresentado na figura 4.7. Assim, a matriz de dados de entrada u construída contém um vetor coluna para cada entrada. Os dois primeiros vetores colunas têm valor constante com os valores dos fluxos $\omega_{g,in}$ e $\omega_{l,in}$ e o terceiro vetor coluna é o sinal PRBS aplicado à válvula.

As saídas do sistema são: P_1 , a pressão na base do riser; P_2 , a pressão no topo do riser; ω_{out} , fluxo da mistura na saída do sistema. Em (JAHANSHAH, 2013), o modelo é simulado para uma abertura constante da válvula Z . A fim de obter a resposta do sistema a um vetor de dados de identificação da entrada Z , o algoritmo foi ligeiramente modificado

adicionando o vetor de entrada e o da escala temporal da entrada. Esta técnica é usada para interpolar o valor do vetor Z para cada instante de tempo da integração do modelo. A saída do sistema para os sinais de entrada $\omega_{l,in} = 8,64\text{kg/s}$, $\omega_{g,in} = 0,36\text{kg/s}$ e o sinal Z apresentado na figura 4.7 é apresentada na figura 4.8.

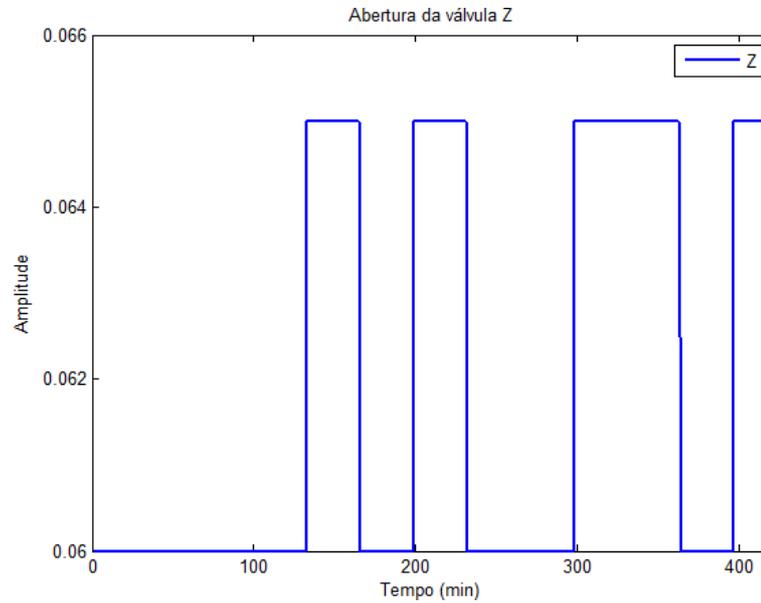


Figura 4.7: Variação da abertura da válvula de saída Z

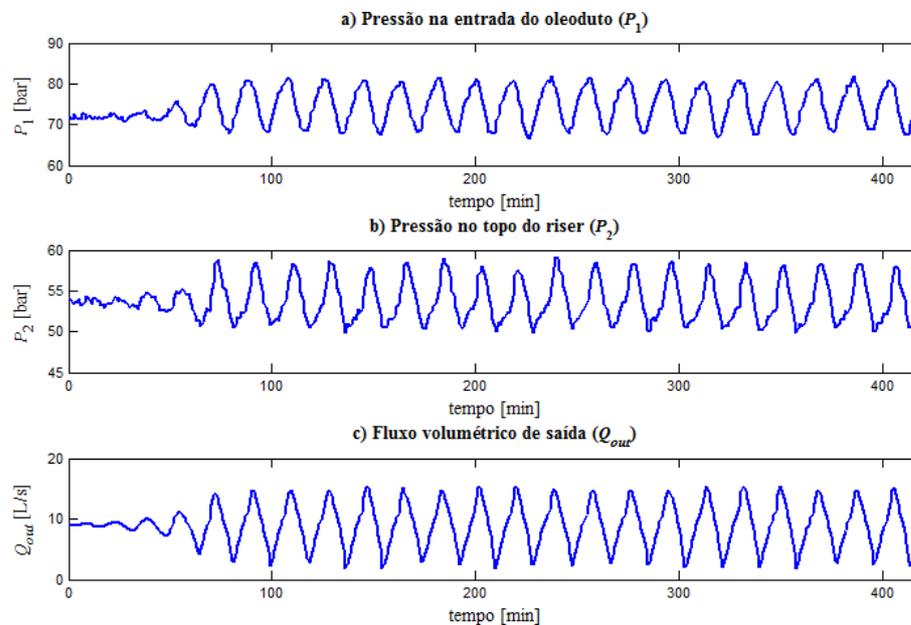


Figura 4.8: Saídas da simulação do modelo da golfada para os dados de identificação

Foi inserido um ruído branco em cada saída, a relação sinal ruído usada foi de aproximadamente 30dB. Os sinais foram normalizados antes de sua utilização nos algoritmos de identificação. Foram utilizados três modelos MIMO ARX, de ordens $n_a = n_b = 2, 3, 4$ para aproximar o processo. O horizonte de predição foi definido como $n_f = 20$. Após a

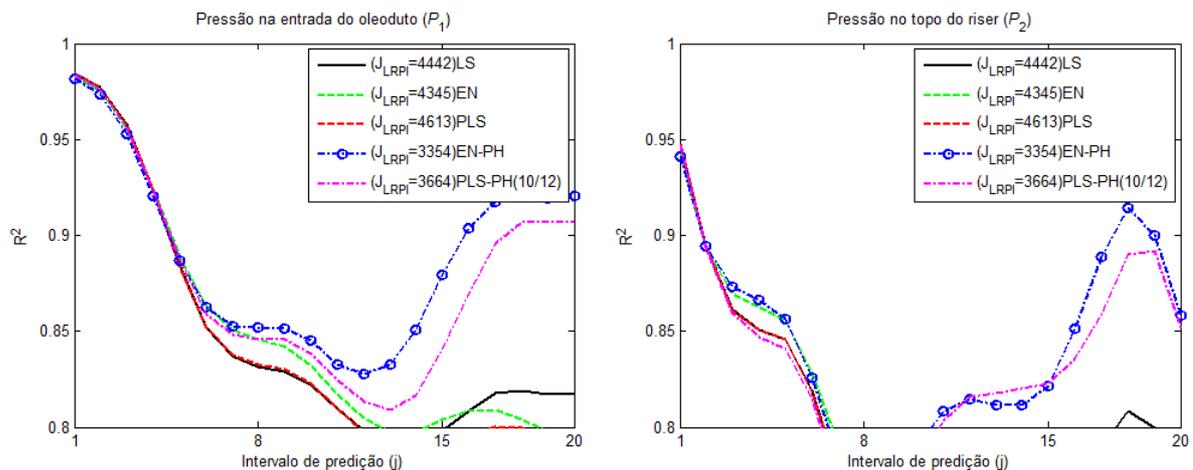
obtenção dos parâmetros θ por cada método de identificação, é aplicada uma nova entrada Z aleatória PRBS de validação a fim de avaliar o desempenho do preditor.

O método de identificação LM utilizado no exemplo anterior não obteve sucesso para identificar o processo da golfada. Na execução do algoritmo ocorreu um erro de memória e foram gerados índices $J_{LRPI_{EV}}$ da ordem de 10^{36} . Sendo assim, este método foi excluído da análise comparativa. A fim de não reduzir a diversidade de técnicas comparadas, o método de identificação PLS apresentado anteriormente, foi utilizado como substituto do método LM.

Nas figuras 4.9, 4.10 e 4.11 são apresentadas as curvas dos indicadores de validação $R^2_{20 \times 3}$, que consiste no valor de R^2 obtido aplicando a equação (4.1) para cada saída e varrendo j por todo o intervalo de predição $j \in [0, 20]$. Os valores do índice $J_{LRPI_{EV}}$ para cada método de identificação também foram calculados e inseridos nas legendas das figuras.

Verifica-se nas curvas de $R^2_{20 \times 3}$ das figuras 4.9 a 4.11, que o algoritmo EN-PH, no geral, apresentou o melhor desempenho entre os métodos de identificação analisados. O melhor ajuste do modelo foi obtido quando $n_a = n_b = 4$, onde $J_{LRPI_{EV}} = 2897$. Nota-se que os índices $R^2_{20 \times 3}$ não decrescem conforme j aumenta neste exemplo, o motivo para este fenômeno está na característica oscilatória das saídas observadas.

Uma comparação mais detalhada sobre algumas propriedades dos modelos de regressão obtidos neste exemplo é realizada com base nos coeficientes de Theil apresentados na tabela 4.2.



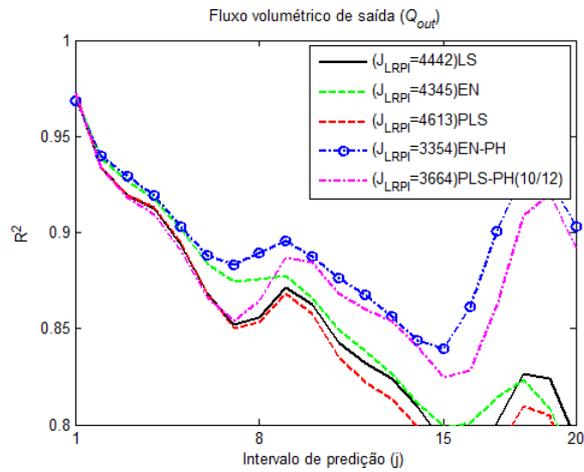


Figura 4.9: Curvas de R^2 para modelo ordem 2 - Golfada

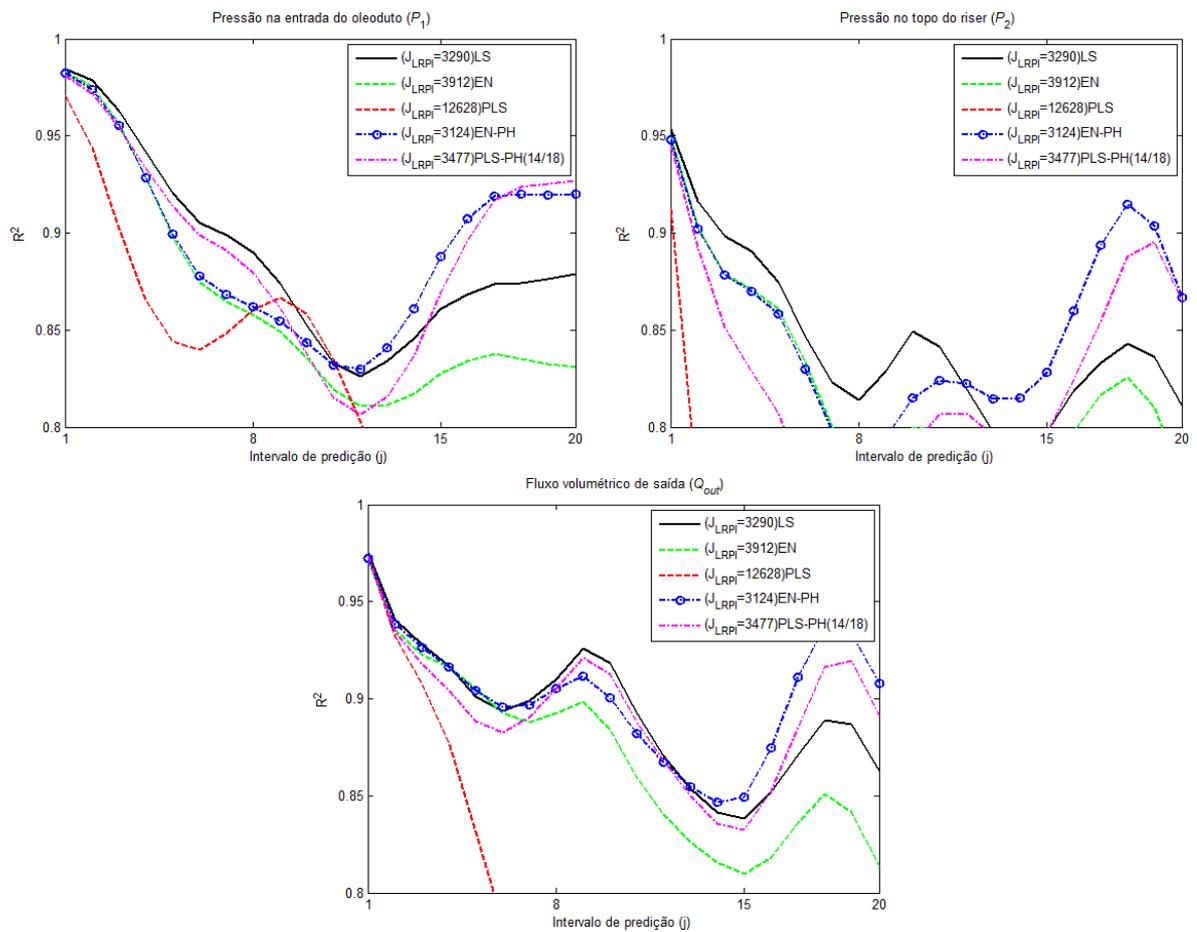


Figura 4.10: Curvas de R^2 para modelo ordem 3 - Golfada

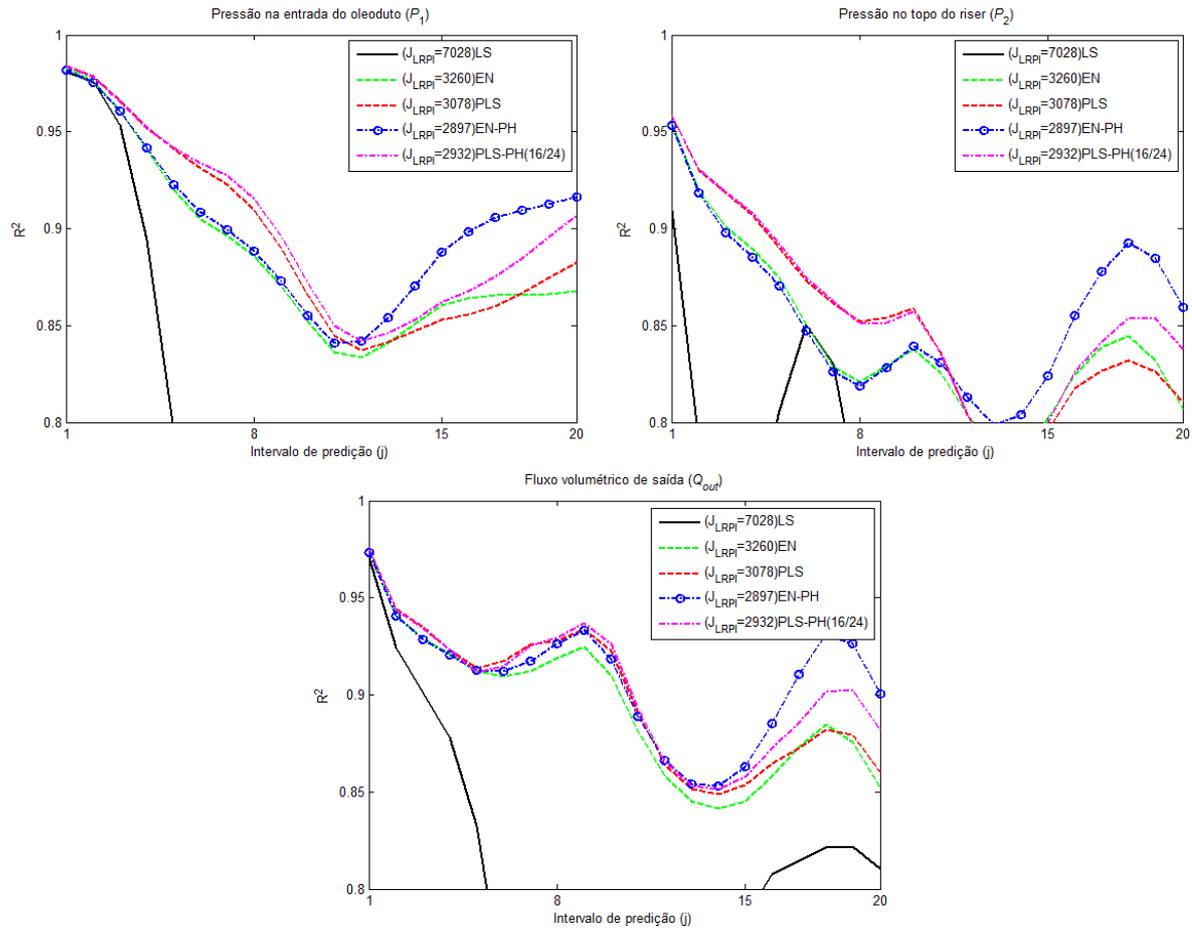


Figura 4.11: Curvas de R^2 para modelo ordem 4 - Golfada

A partir da tabela 4.2 verifica-se que em todos os casos método o EN-PH forneceu um modelo com menor índice padrão e de variância de Theil. Em alguns casos o método EN-PH também apresentou menor índice de polarização, porém na maioria dos casos ele obteve uma polarização ligeiramente maior que os demais métodos. No geral, o índice de correlação do método proposto indicou que ele forneceu uma maior correlação entre as saídas estimadas e medidas neste exemplo.

Tabela 4.2: Índices de Theil de validação - Golfada

| Saída 1 | | | | | | |
|-----------|---------|----------------|----------------|----------------|----------------|----------------|
| n_a/n_b | Índice | LS | EN | PLS | EN-PH | PLS-PH |
| 2 | U | 0,20728 | 0,21240 | 0,21192 | 0,16669 | 0,17629 |
| | U_b | 0,03025 | 0,04403 | 0,00745 | 0,05861 | 0,02698 |
| | U_v | 0,18320 | 0,31825 | 0,20248 | 0,01782 | 0,02366 |
| | U_c | 0,78657 | 0,63776 | 0,79009 | 0,92357 | 0,94936 |
| 3 | U | 0,17078 | 0,19476 | 0,22587 | 0,16135 | 0,16142 |
| | U_b | 0,04847 | 0,05217 | 0,06279 | 0,06713 | 0,08004 |
| | U_v | 0,10772 | 0,21019 | 0,01766 | 0,01624 | 0,00970 |
| | U_c | 0,84382 | 0,73767 | 0,91955 | 0,91662 | 0,91625 |
| 4 | U | 0,25637 | 0,17281 | 0,16381 | 0,15325 | 0,15362 |
| | U_b | 0,00004 | 0,06596 | 0,04857 | 0,07488 | 0,04556 |
| | U_v | 0,00212 | 0,15856 | 0,09431 | 0,01299 | 0,01739 |
| | U_c | 0,99784 | 0,77550 | 0,85713 | 0,91213 | 0,93705 |
| Saída 2 | | | | | | |
| n_a/n_b | Índices | LS | EN | PLS | EN-PH | PLS-PH |
| 2 | U | 0,24040 | 0,24717 | 0,24607 | 0,20321 | 0,20992 |
| | U_b | 0,00032 | 0,00229 | 0,00037 | 0,00407 | 0,00302 |
| | U_v | 0,15454 | 0,30420 | 0,16904 | 0,03061 | 0,03803 |
| | U_c | 0,84515 | 0,69355 | 0,83061 | 0,96532 | 0,95896 |
| 3 | U | 0,20648 | 0,22764 | 0,32615 | 0,19385 | 0,20159 |
| | U_b | 0,00082 | 0,00292 | 0,06070 | 0,00450 | 0,01281 |
| | U_v | 0,11850 | 0,18531 | 0,45312 | 0,01662 | 0,03242 |
| | U_c | 0,88069 | 0,81179 | 0,48623 | 0,97888 | 0,97371 |
| 4 | U | 0,25968 | 0,20770 | 0,19924 | 0,18934 | 0,19030 |
| | U_b | 0,01975 | 0,00360 | 0,00068 | 0,00460 | 0,00051 |
| | U_v | 0,01701 | 0,15114 | 0,10180 | 0,02190 | 0,03140 |
| | U_c | 0,96324 | 0,84528 | 0,89754 | 0,97350 | 0,96810 |
| Saída 3 | | | | | | |
| n_a/n_b | Índices | LS | EN | PLS | EN-PH | PLS-PH |
| 2 | U | 0,20422 | 0,20483 | 0,21028 | 0,16204 | 0,17392 |
| | U_b | 0,00105 | 0,00004 | 0,00214 | 0,00028 | 0,00013 |
| | U_v | 0,16830 | 0,28463 | 0,18397 | 0,00863 | 0,02031 |
| | U_c | 0,83067 | 0,71537 | 0,81391 | 0,99109 | 0,97956 |
| 3 | U | 0,16949 | 0,19134 | 0,33400 | 0,15552 | 0,16535 |
| | U_b | 0,00089 | 0,00009 | 0,00411 | 0,00024 | 0,00440 |
| | U_v | 0,12454 | 0,22251 | 0,00713 | 0,00559 | 0,04010 |
| | U_c | 0,87458 | 0,77743 | 0,96476 | 0,99417 | 0,95551 |
| 4 | U | 0,25264 | 0,16938 | 0,16215 | 0,14916 | 0,15356 |
| | U_b | 0,18003 | 0,00011 | 0,00102 | 0,00022 | 0,00127 |
| | U_v | 0,00751 | 0,17039 | 0,10015 | 0,00701 | 0,01559 |
| | U_c | 0,81246 | 0,82952 | 0,89885 | 0,99277 | 0,98315 |

4.5 Sistema de temperatura baseado em módulos Peltier

Neste terceiro exemplo, os métodos de identificação foram comparados em uma aplicação prática. O sistema de escala laboratorial utilizando módulos Peltier para controle de temperatura consiste em dois elementos peltier que podem ser acionados separadamente, formando um processo acoplado de duas entradas e duas saídas (TITO). Como a montagem é acoplada termicamente, é possível analisar as malhas diretas e cruzadas. Também pode ser feito um estudo de perturbações de carga. Na figura 4.12 é apresentada uma foto da montagem do sistema.



Figura 4.12: Foto da montagem do sistema.

4.5.1 Módulo peltier

Os módulos ou pastilhas peltier, ou ainda pastilhas termoelétricas, funcionam pelo princípio do efeito peltier. Normalmente estes módulos são constituídos de um material semicondutor, o telureto de bismuto, sendo fortemente dopado para a criação de material tipo n e material tipo p. Ao ligar a fonte de tensão à célula peltier (figura 4.13), se estabelece um fluxo de cargas do material semicondutor P para o semicondutor N. Este fluxo cria uma diferença de temperatura entre as superfícies da junção. Assim, uma superfície irá aquecer, enquanto a outra resfria-se.

Normalmente estas células são combinadas eletricamente em série, e termicamente em paralelo. Estes vários agrupamentos em série, são interligados de forma a compor o módulo peltier, ver Figura 4.14.

Na presente montagem da planta térmica, foram utilizados dois módulos peltier modelo CP1.0-127-06 do fabricante Melcor. A temperatura da superfície quente varia de 25°C a 50°C, a potência consumida é de 25,7W a 29,1W e uma corrente máxima de 3A.

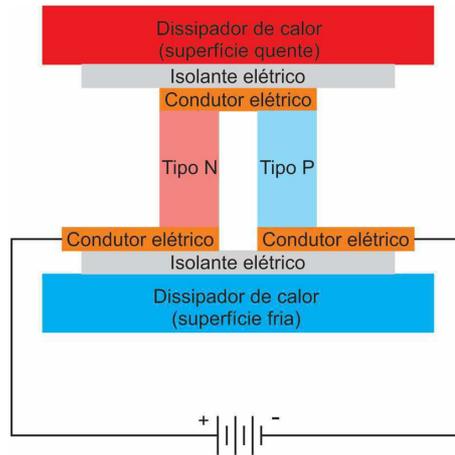


Figura 4.13: Diagrama da célula peltier

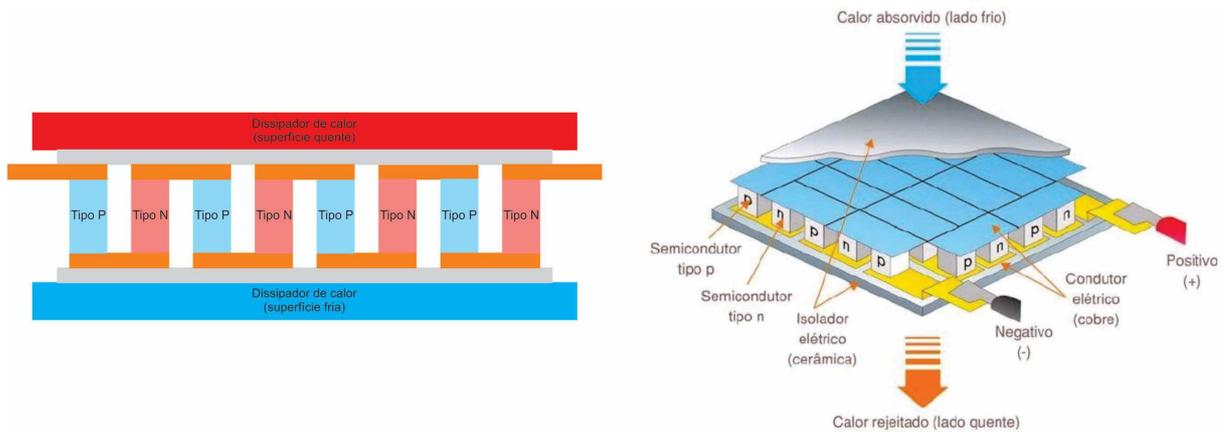


Figura 4.14: Diagrama estrutural do módulo peltier

Cada módulo possui uma face acoplada ao dissipador de calor de um cooler, com o objetivo de melhorar a troca de calor com o ambiente. A outra face de cada módulo está acoplada termicamente a uma peça cilíndrica de alumínio sólido, composta por três discos de diâmetros diferentes, conforme a Figura 4.15. Nesta peça de alumínio, estão fixados os sensores de temperatura LM35.

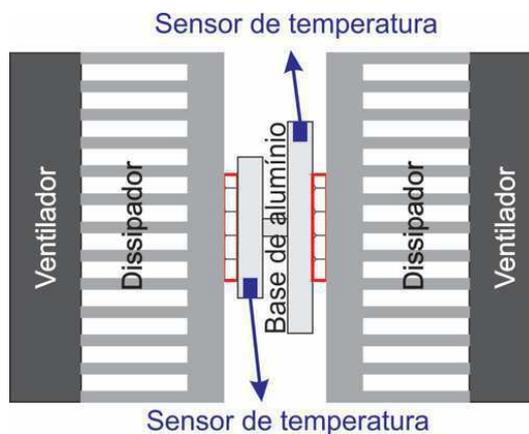


Figura 4.15: Montagem dos módulos peltier

4.5.2 Circuito eletrônico

O kit de desenvolvimento USBizi, fabricado pela GHI Electronics, é responsável por realizar a comunicação com o PC, acionar os drivers de potência e realizar a leitura dos sensores de temperatura. Este kit possui um microcontrolador ARM7 de 32 bits rodando a 72 MHz, assim como vários periféricos do tipo, PWM; USB; GPIO; USART; entradas analógicas; entre outros. Seu *firmware* é desenvolvido sobre a plataforma .NET Micro Framework da Microsoft, utilizando o Visual Studio 2010. Para o acionamento dos módulos peltier, foram utilizados quatro transistores MOSFET canal N, para cada módulo. Estes transistores foram interligados formando um circuito de uma ponte H, possibilitando a inversão no sentido da corrente nos módulos. A inversão da corrente ocasiona uma inversão no sentido de fluxo de calor, assim, a superfície que estava quente passará a esfriar, e vice-versa. O USBizi é interligado ao driver de potência por meio das saídas digitais e do PWM. As saídas digitais são utilizadas para selecionar o sentido da corrente nos módulos, já o PWM, é responsável por realizar o acionamento das chaves eletrônicas, possibilitando o controle da potência transferida para os módulos. O monitoramento da temperatura da carga de alumínio é realizado por dois sensores LM35. Este sensor possui uma saída de tensão analógica, cujo valor é diretamente proporcional ao aumento da temperatura ao qual está submetido, mais precisamente $10\text{mV}/^{\circ}\text{C}$. A saída de cada sensor é conectada a um circuito amplificador de tensão, que tem como objetivo, melhorar o condicionamento do sinal para que ele seja aplicado a entrada analógica do USBizi. Na saída do circuito amplificador, tem-se um sinal analógico de aproximadamente $30\text{mV}/^{\circ}\text{C}$. Na figura 4.16 é apresentado o diagrama de ligação do sistema.

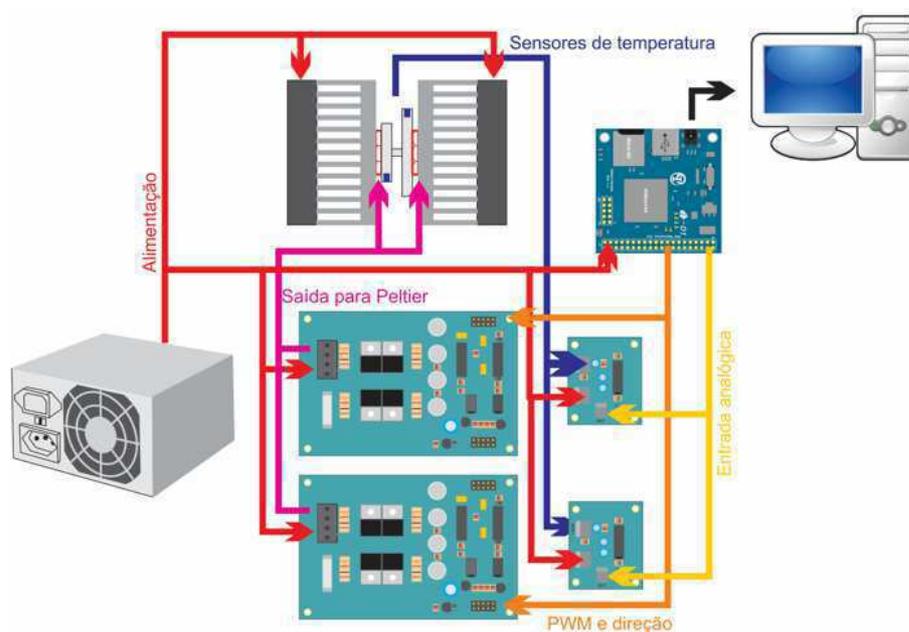


Figura 4.16: Diagrama de ligação do sistema.

4.5.3 Interface de comunicação

Na planta térmica havia sido criada uma versão simples do conhecido protocolo Modbus entre o PC de supervisão e o microcontrolador embarcado na planta didática. As informações escritas e lidas no *software* de supervisão são posteriormente salvas em um arquivo de texto ou disponibilizadas via servidor OPC. O *software* de supervisão foi implementado em C# e ele é responsável pela interface entre a planta desenvolvida e o usuário do sistema. Como explicado previamente, a comunicação entre o supervisor e a planta é realizada via Modbus e entre o supervisor e qualquer outra aplicação, via OPC. A tela principal do *software* de supervisão é apresentada na figura 4.17

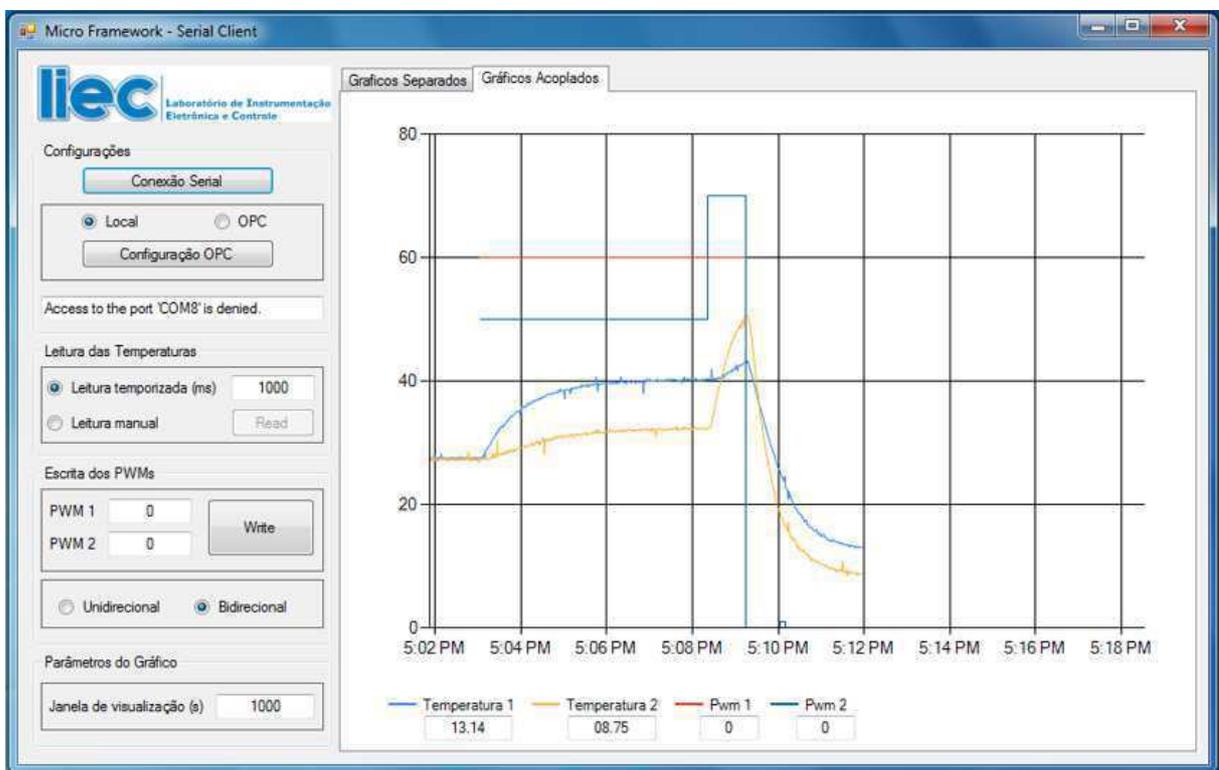


Figura 4.17: Tela do Software supervisorio.

Inicialmente é realizada a configuração da porta serial, através do botão "Conexão Serial", em seguida é escolhido se a manipulação de dados será dada localmente ou via OPC. A leitura das temperaturas pode ser feita de modo assíncrono ou de forma temporizada, através do painel "Leitura das temperaturas". Os valores dos duty cycles dos sinais PWM a serem aplicados nos peltiers são inseridos no painel "Escritas dos PWMs", note que o modo unidirecional aplica somente tensões positivas nos módulos, enquanto o modo bidirecional permite tensões negativas.

4.5.4 Identificação

Os sinais de excitação escolhidos para o processo foram sinais PRBS e foram gerados utilizando a função *idinput* do Matlab[®] com período de amostragem de 0,5s $BAND = [0 \ 0,05]$ e $levels = [50 \ 100]$ (Varia o duty cycle do PWM de 50% a 100%). O conjunto de dados de identificação gerados é apresentado na figura 4.18.

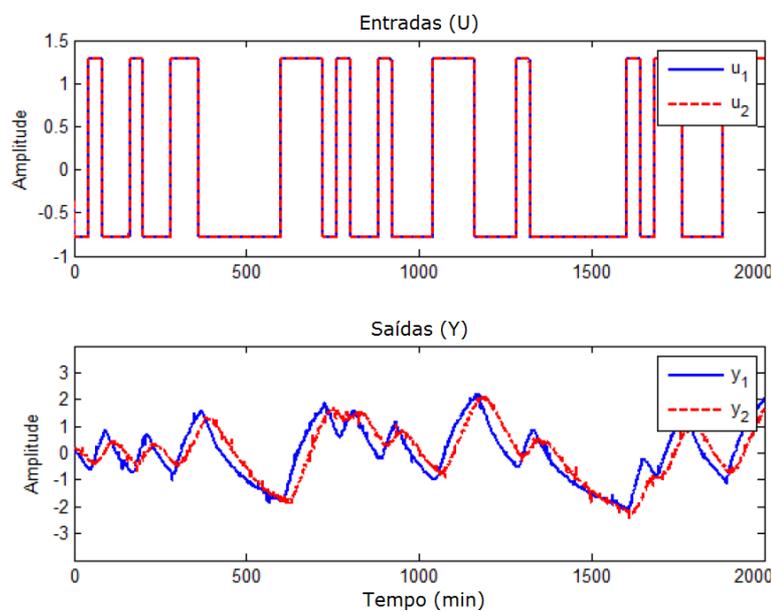


Figura 4.18: Dados de identificação da planta térmica

Estes sinais foram normalizados antes de sua utilização nos algoritmos de identificação. Também foram utilizados três modelos MIMO ARX, de ordens $n_a = n_b = 2, 3, 4$ para aproximar o processo. O horizonte de predição foi definido como $n_f = 30$. Após a obtenção dos parâmetros θ por cada método de identificação, é aplicada uma nova entrada aleatória PRBS de validação a fim de avaliar o desempenho do preditor.

Como no exemplo anterior, o método de identificação LM não obteve sucesso para identificar o processo do sistema de temperatura, gerando índices J_{LRPIEV} da ordem de 10^{31} , por isso ele foi excluído da análise comparativa. O método de identificação PLS apresentado anteriormente, foi utilizado para substituir o método LM.

Nas figuras 4.19, 4.20 e 4.21 são apresentadas as curvas dos indicadores de validação $R_{30 \times 2}^2$, que consiste no valor de R^2 obtido aplicando a equação (4.1) para cada saída e varrendo j por todo o intervalo de predição $j \in [0, 30]$. Os valores do índice J_{LRPIEV} para cada método de identificação também foram calculados e inseridos nas legendas das figuras.

Verifica-se nas curvas de $R_{30 \times 2}^2$ das figuras 4.19, 4.20 e 4.21 que o algoritmo EN-PH apresentou o melhor desempenho entre os métodos de identificação analisados, para todas as ordens dos modelos sugeridas. O melhor ajuste do modelo foi obtido para $n_a = n_b = 4$,

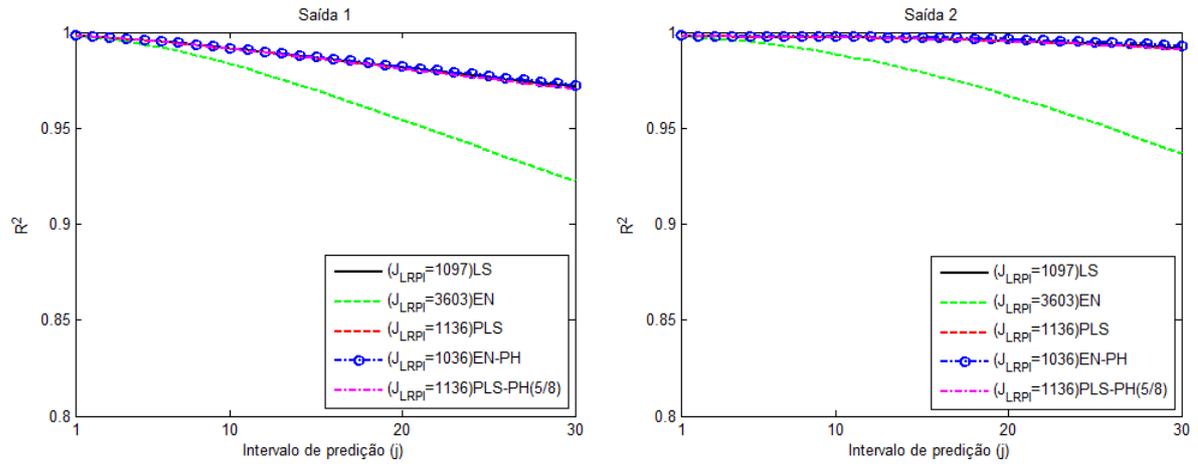


Figura 4.19: Curvas de R^2 para modelo ordem 2 - Planta térmica

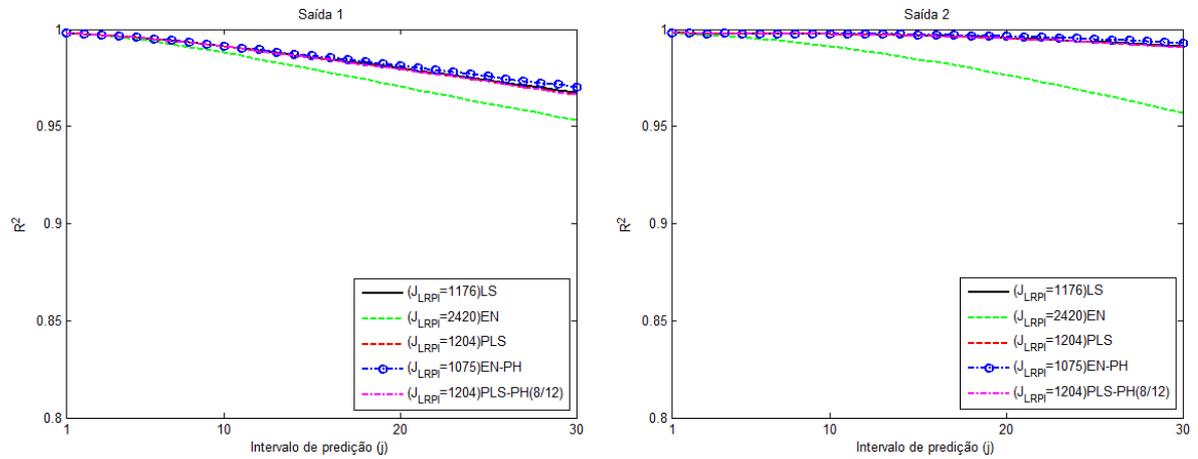


Figura 4.20: Curvas de R^2 para modelo ordem 3 - Planta térmica

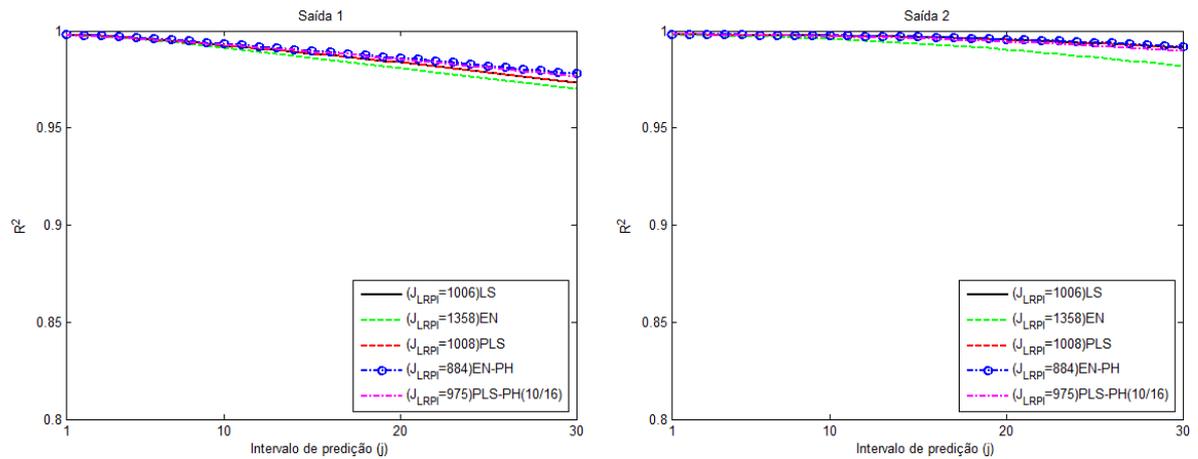


Figura 4.21: Curvas de R^2 para modelo ordem 4 - Planta térmica

onde $J_{LRPI_{EV}} = 884$. Nota-se que em todos os métodos de identificação, o valor de R^2 manteve-se próximo de 1 para todos os valores de j , a razão disto é porque este sensor está conectado ao disco cilíndrico de raio maior, tornando esta saída menos susceptível

ao ruído e a correlação.

Tabela 4.3: Índices de Theil de validação - Planta térmica

| Saída 1 | | | | | | |
|-----------|---------|----------------|----------------|----------------|----------------|----------------|
| n_a/n_b | Índice | LS | EN | LM | EN-PH | PLS-PH |
| 2 | U | 0,05890 | 0,09883 | 0,05985 | 0,05847 | 0,05985 |
| | U_b | 0,00645 | 0,00239 | 0,00626 | 0,00651 | 0,00626 |
| | U_v | 0,02096 | 0,43448 | 0,02330 | 0,01019 | 0,02330 |
| | U_c | 0,97259 | 0,56313 | 0,97044 | 0,98329 | 0,97044 |
| 3 | U | 0,05747 | 0,07142 | 0,05711 | 0,05754 | 0,05711 |
| | U_b | 0,00722 | 0,00485 | 0,00729 | 0,00727 | 0,00729 |
| | U_v | 0,03702 | 0,27615 | 0,03646 | 0,00656 | 0,03646 |
| | U_c | 0,95576 | 0,71900 | 0,95626 | 0,98617 | 0,95626 |
| 4 | U | 0,05583 | 0,06269 | 0,05588 | 0,05163 | 0,05279 |
| | U_b | 0,00805 | 0,00673 | 0,00804 | 0,00946 | 0,00960 |
| | U_v | 0,04902 | 0,23691 | 0,04879 | 0,06526 | 0,09233 |
| | U_c | 0,94292 | 0,75636 | 0,94317 | 0,92528 | 0,89808 |
| Saída 2 | | | | | | |
| n_a/n_b | Índices | LS | EN | LM | EN-PH | PLS-PH |
| 2 | U | 0,03230 | 0,08616 | 0,03292 | 0,02932 | 0,03292 |
| | U_b | 0,00881 | 0,00116 | 0,00851 | 0,01097 | 0,00851 |
| | U_v | 0,07281 | 0,70600 | 0,07367 | 0,13693 | 0,07367 |
| | U_c | 0,91838 | 0,29286 | 0,91782 | 0,85211 | 0,91782 |
| 3 | U | 0,03128 | 0,05525 | 0,03117 | 0,02978 | 0,03117 |
| | U_b | 0,01000 | 0,00311 | 0,01005 | 0,01139 | 0,01005 |
| | U_v | 0,17533 | 0,63925 | 0,17537 | 0,19622 | 0,17537 |
| | U_c | 0,81467 | 0,35765 | 0,81458 | 0,79239 | 0,81458 |
| 4 | U | 0,03142 | 0,04613 | 0,03145 | 0,03058 | 0,03404 |
| | U_b | 0,01040 | 0,00476 | 0,01039 | 0,01114 | 0,00949 |
| | U_v | 0,23704 | 0,59193 | 0,23808 | 0,21965 | 0,28279 |
| | U_c | 0,75256 | 0,40332 | 0,75154 | 0,76921 | 0,70772 |

Na tabela 4.3 verificou-se que o método EN-PH forneceu um modelo com menor índice padrão de Theil, novamente refletindo a associação com deste índice com o $J_{LRPI_{EV}}$. O método EN-PH também alcançou os melhores índices de variância e covariância para a saída 1, porém para a saída 2 o método dos mínimos quadrados o superou ligeiramente nestes quesitos. É interessante verificar que, no geral, os maiores índices de covariância são obtidos nos métodos que apresentam a menor variância.

4.6 Conclusão

Neste capítulo o desempenho do método proposto EN-PH foi analisado e comparado com outros métodos de identificação já difundidos. Os métodos foram aplicados em três exemplos, dois utilizando processos simulados um terceiro exemplo prático. Os métodos foram comparados a partir de três índices de desempenho e verificou-se nos casos analisados que o método proposto EN-PH apresentou resultados superiores. No geral, o método EN-PH forneceu modelos com menor variância da saída estimada e maior covariância entre a saída estimada e a saída medida. Com isso, conclui-se que o EN-PH é uma alternativa bastante atrativa para identificação MRI de processos.

Nos exemplos verificou-se que o algoritmo EN-PH levou um tempo consideravelmente maior para obter os parâmetros que os outros métodos, isto foi atribuído à metodologia utilizada para encontrar o melhor valor para α do método EN-PH. Porém na prática, este tempo não representa grandes problemas uma vez que a identificação do sistema é realizada *offline*.

Capítulo 5

Conclusões e Sugestões de Trabalhos Futuros

5.1 Conclusões

Neste trabalho foram discutidos temas acerca de identificação de sistemas. Em especial, foi enfatizada a importância da identificação relevante ao controle para a obtenção de modelos dos processos mais apropriados para a tarefa de projetar o controlador.

Dada a relevância que o controle preditivo tem obtido na indústria como uma das técnicas de controle avançado mais utilizada atualmente, segundo pesquisa realizada em (QIN; BADGWELL, 2003), foi objetivo deste trabalho investigar as técnicas MRI.

A primeira técnica MRI apresentada foi a de minimizar a função de custo de múltiplos passos à frente utilizando o algoritmo de otimização numérica Levenberg-Marquardt. Porém esta abordagem não obteve sucesso em dois exemplos em que foi aplicada, onde o algoritmo não convergiu resultando em erro no programa.

A segunda técnica MRI vista foi o método PLS-PH criado em (PLA, 2012). O algoritmo PLS-PH mostrou possuir grande potencial em obter melhores modelos para aplicações de controle preditivo e conforme apresentado em (QUACHIO, 2012) também foi capaz de realizar identificação em malha fechada e identificar modelos não lineares.

O método EN-PH desenvolvido nesta dissertação foi implementado aplicando a técnica de regressão regularizada Elastic Net para estender o algoritmo de busca linear e assim, obter os parâmetros do modelo ARX do processo.

Os métodos foram comparados em três exemplos práticos. Em todos os exemplos estudados o método EN-PH forneceu os menores erros médios quadráticos dentre os métodos utilizados. No geral, o desempenho do método EN-PH foi similar ao do método PLS-PH. O método EN-PH obteve uma eficiência ligeiramente maior que o PLS-PH nestes exemplos, e a causa foi atribuída à troca variância-polarização ter sido mais eficiente a partir

dos métodos de regularização do que pelo PLS, como pôde ser observado à partir dos índices de Theil. Nestes índices observou-se que o EN-PH forneceu uma predição com menor variância e ligeiramente maior polarização, o que pelos motivos descritos anteriormente, resultou em um erro médio quadrático inferior ao PLS-PH e aos demais métodos.

5.2 Sugestões de Trabalhos Futuros

Dentre as perspectivas de trabalhos futuros podem ser citadas:

- Aprimorar o algoritmo EN-PH implementando a técnica da validação cruzada para escolher um valor melhor de α na regressão elastic net.
- Estender o método de identificação EN-PH para o caso do sistema em malha fechada;
- Estender o método de identificação EN-PH para a identificação de modelos não lineares com estrutura NARX polinomial;
- Propor diferentes tipos de excitações de forma a manter o processo de identificação o menos invasivo possível;
- Propor técnicas de sintonia de controladores MPC baseando-se nos modelos identificados;
- Avaliar os modelos obtidos tanto no sentido da identificação como no projeto do controlador MPC;

Referências Bibliográficas

- ASTROM, T. B. K. J. Numerical identification of linear dynamic systems for normal operating records. *Proc. 2nd IFAC Symp. Theory of Self-Adaptive Systems, Teddington*, p. 96/111, 1965.
- BAI, Y.; BAI, Q. *Subsea Pipelines and Riser*. Oxford: [s.n.], 2005.
- BELLEMANS, T.; SCHUTTER, B. D.; MOOR, B. D. Model predictive control for ramp metering of motorway traffic - a case study. *Control Engineering Practice*, 14, p. 757 – 767, 2006.
- BOHLIN, T.; ASTROM, K. J. Numerical identification of linear dynamic systems for normal operating records. *2nd IFAC Symp. Theory of Self-Adaptive Systems, Teddington*, p. 96–111, 1965.
- BURT, O. R.; FRANK, M.; BEATTIE, B. R. Prior information and heuristic ridge regression for production function estimation. *Western Journal of Agricultural Economics*, v. 12, p. 135–143, 1987.
- CAMPOS, M. C.; LOUREIRO, P.; B., F. Álvaro. Novas estratégias de controle para a plataforma de petróleo p-55. *Rio Oil & Gas Expo and Conference*, Rio de Janeiro - RJ: IBP - Instituto Brasileiro de Petróleo e Gás, 2006.
- CLARKE, D. W.; MOHTADI, C.; S, P. S. T. GENERALIZED PREDICTIVE CONTROL-PART 1. *Automatica*, 1987.
- CUTLER, C.; RAMAKER, B. Dynamic matrix control - a computer control algorithm. *Proceedings of the joint American control conference*, San Francisco, p. Paper WP5–B, 1980.
- EFRON, B. et al. Least angle regression. *The Annals of Statistics*, v. 32, p. 407–451, 2004.
- GEVERS, M. A decade of progress in iterative process control design: from theory to practice. *Journal of Process Control*, v. 12, p. 519–531, 2002.
- GODHAVN, J.-M.; FARD, M. P.; FUCHS, P. H. New slug control strategies, tuning rules and experimental results. *Journal of Process Control*, v. 15, p. 547 – 557, 2005.
- GOLDSTEIN, A. *Constructive Real Analysis*. Londres: [s.n.], 1967.
- GOPALUNI, R. B.; PATWARDHAN, R. S.; SHAH, S. L. Mpc relevant identification - tuning the noise model. *Journal of Process Control*, 14, p. 699 – 714, 2004.

- HOERL, A.; KENNARD, R. Ridge regression: Biased estimation for nonorthogonal problems. *Tecnometrics*, v. 12, p. 55–67, 1970.
- HROVAT, D. et al. The development of model predictive control in automotive industry: A survey. *IEEE international conference on control applications*, p. 295 – 302, 2012.
- HÖSKULDSSON. Pls regression methods. *Journal of Chemometrics*, v. 2, p. 211–228, 1988.
- ISERMANN, R.; MÜNCHHOF, M. *Identification of Dynamic Systems: An Introduction with Applications*. [S.l.]: Springer, 2011.
- JAHANSHAH, E. *Control Solutions for Multiphase Flow, Linear and nonlinear approaches to anti-slug control*. Tese (Doutorado) — Norwegian University of Science and Technology, Trondheim, Outubro 2013.
- KALMAN, B. L. H. R. E. Effective construction of linear state-variable models from input/output functions. *Regelungstechnik vol. 14*, n. no. 12, p. 545/548, 1966.
- KIERS, H. A. L.; A.K., S. A comparison of various methods for multivariate regression with highly collinear variables. *Statistical Methods and Applications*, v. 16, p. 193–228, 2007.
- KOURO, S. et al. Model predictive control - a simple and powerful method to control power converters. *IEEE Transactions on Industrial Electronics*, 56, p. 1826 – 1838, 2009.
- KUHN, M.; JOHNSON, K. *Applied Predictive Modeling*. New York, USA: [s.n.], 2013.
- LJUNG, L. *System Identification: Theory for the user*. 2nd ed.. ed. [S.l.]: Prentice Hall, 1999.
- MACIEJOWSKI, J. M. *Predictive Control with Constraints*. Englewood Cliffs: Prentice Hall, 2002. 330 p. ISBN 0201398230.
- MARTENS, H. Reliable and relevant modelling of real world data: a personal account of the development of pls regression. *Chemometrics and Intelligent Laboratory Systems*, v. 58, p. 85–95, 2001.
- MEGLIO, F. D.; KAASA, G.; PETIT, N. A first principle model for multiphase slugging flow in vertical risers. *48th IEEE Conference on Decision and Control and 28th Chinese Control Conference*, p. 8244 – 8251, 2009.
- MELAND, A. K. *Stabilization of two-phase flow in risers from reservoirs, Comparison of alternative simple models*. Dissertação (Specialization Thesis) — Norwegian University of Science and Technology, Dezembro 2010.
- MOOR, P. V. O. B. *Subspace Identification for Linear Systems*. [S.l.]: Kluwer Academic Publishers, 1996.
- MORARI, M.; LEE., J. H. Model predictive control: past, present and future. *Computers and Chemical Engineering*, vol. 23, p. 667 – 682, 1999.

- MORE, J. J. The levenberg-marquardt algorithm: implementation and theory. *Numerical Analysis*, ed. G. A. Watson, *Lecture Notes in Mathematics 630*, Springer Verlag, p. 105–116, 1977.
- OSBORNE, M.; PRESNELL, B.; TURLACH, B. On the lasso and its dual. *Journal of Computational and Graphical Statistics*, v. 9(2), p. 319–337, 2000.
- PALADINO, E. *Estudo do Escoamento Multifásico em Medidores de Vazão do Tipo Pressão Diferencial*. Tese (Doutorado) — Universidade Federal de Santa Catarina, Florianópolis - SC, Abril 2005.
- PAYNE, G. G. R. *Dynamic System Identification: Experiment Design and Data Analysis*. [S.l.]: Academic Press, 1977.
- PLA, D. *MPC: Relevant Identification, and Control in the Latent Variable Space*. Tese (Doutorado) — Universidad Politécnica de Valencia, Valencia, Janeiro 2012.
- QIAN, J. et al. Glnet for matlab. disponível em: http://www.stanford.edu/~hastie/glnet_matlab/, Acesso em: 23 fev. 2015, 2013.
- QIN, S. J.; BADGWELL, T. A. A survey of industrial model predictive control technology. *Control Engineering Practice*, v. 11, n. 7, p. 733 – 764, 2003.
- QUACHIO, R. *Análise do algoritmo PLS-PH para identificação de sistemas*. Dissertação (Mestrado) — Escola Politécnica da Universidade de São Paulo, São Paulo, Fevereiro 2012.
- RICHALET, J. et al. Paper: Model predictive heuristic control. *Automatica*, v. 14, n. 5, p. 413–428, set. 1978. ISSN 0005-1098.
- ROSSITER, J. A. *Model-Based Predictive Control: a Practical Approach*. [S.l.]: CRC press, 2003. 344 p.
- SHOOK, D.; MOHTADI, C.; SHAH, S. Identification for long-range predictive error. *IEE proceedings-D*, v. 138, p. 75–84, 1991.
- SILVA, C. B. C. *Medição de Vazão e Propriedades em Escoamento Multifásico Solução Econômica para Diferentes Atividades Industriais*. Rio de Janeiro, 2000.
- SIVERTSEN, H.; SKOGESTAD, S. Anti-slug control experiments on a smallscale two-phase loop. *Proc. European Symposium on Computer Aided Process Engineering*, Barcelona, Spain: Elsevier Science, p. 1021 – 1026, 2005.
- STOICA, T. S. P. *System Identification*. [S.l.]: Prentice Hall, New Jersey, 1989.
- STORKAAS, E.; SKOGESTAD, S. A low-dimensional dynamic model of severe slugging for control design and analysis. *11th International Conference on Multiphase flow*, BHR Group, p. 117 – 133, 2003.
- STORKAAS, E.; SKOGESTAD, S. Controllability analysis of two-phase pipeline-riser systems at riser slugging conditions. *Control Engineering Practice* 15, p. 567–581, 2007.

THEIL, H. Economic forecasts and policy. *The American Economic Review*, v. 49, p. 711–716, 1959.

TIBSHIRANI, R. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society Series B (Methodological)*, v. 58, p. 267–288, 1996.

TROPSHA, A.; GRAMATICA, P.; GOMBAR, V. The importance of being earnest: Validation is the absolute essential for successful application and interpretation of qspr models. *QSAR and Combinatorial Science*, v. 22, p. 69–77, 2003.

WOOD, R.; BERRY, M. Terminal composition control of a binary distillation column. *Chemical Engineering Science*, v. 28, p. 1707–1717, 1973.

ZOU, H.; HASTIE, T. Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society Series B*, v. 67, p. 301–320, 2005.

Apêndice A

Modelo MIMO ARX

Considere o seguinte modelo MIMO ARX ¹:

$$\begin{bmatrix} y_1(k) \\ \vdots \\ y_{n_o}(k) \end{bmatrix} = \begin{bmatrix} \frac{\mathbf{B}_{1,1}}{\mathbf{A}_1} & \dots & \frac{\mathbf{B}_{n_i,1}}{\mathbf{A}_1} \\ \vdots & \ddots & \vdots \\ \frac{\mathbf{B}_{1,n_o}}{\mathbf{A}_{n_o}} & \dots & \frac{\mathbf{B}_{n_i,n_o}}{\mathbf{A}_{n_o}} \end{bmatrix} \begin{bmatrix} u_1(k) \\ \vdots \\ u_{n_i}(k) \end{bmatrix} + \begin{bmatrix} \frac{\xi_1(k)}{\mathbf{A}_1} \\ \vdots \\ \frac{\xi_{n_o}(k)}{\mathbf{A}_{n_o}} \end{bmatrix} \quad (\text{A.1})$$

onde

$$\begin{aligned} \mathbf{B}_{e,s} &= b_{1e,s}z^{-1} + \dots + b_{n_{b_{e,s}}e,s}z^{-n_{b_{e,s}}} \\ \mathbf{A}_{e,s} &= 1 + a_{1s}z^{-1} + \dots + a_{n_{a_s}s}z^{-n_{a_s}} \\ e &\in [1, 2, \dots, n_i] \quad ; \quad s \in [1, 2, \dots, n_o] \end{aligned}$$

Por simplicidade, assume-se que todos os numeradores e denominadores têm a mesma ordem. Caso contrário, podem ser aplicadas operações polinomiais e em seguida fixar alguns valores em zero para transformar a estrutura original em na expressão dada para um modelo MIMO ARX.

A saída s pode ser expressa como:

$$\mathbf{A}_s y_s = [\mathbf{B}_{1,s} \dots \mathbf{B}_{n_i,s}] \begin{bmatrix} u_1(k) \\ \vdots \\ u_{n_i}(k) \end{bmatrix} + \xi_s(k)$$

¹Para facilitar a notação, os polinômios em z^{-1} serão representados apenas por seus nomes, $\mathbf{A}_{n_o}(z^{-1}) \equiv \mathbf{A}_{n_o}$.

$$\begin{aligned}
 y_s &= [\mathbf{B}_{1,s} \dots \mathbf{B}_{n_i,s}] \begin{bmatrix} u_1(k) \\ \vdots \\ u_{n_i}(k) \end{bmatrix} + (1 - \mathbf{A}_s)y_s(k) + \xi_s(k) \\
 &= [\bar{\mathbf{b}}_{1_s} \dots \bar{\mathbf{b}}_{n_{b_s}}] \begin{bmatrix} \mathbf{u}_{k-1}^T \\ \vdots \\ \mathbf{u}_{k-n_b}^T \end{bmatrix} + [-a_{1_s} \dots -a_{n_{a_s}}] \begin{bmatrix} y_s(k-1) \\ \vdots \\ y_s(k-n_a) \end{bmatrix} + \xi_s(k) \\
 &= [\mathbf{u}_{k-1} \dots \mathbf{u}_{k-n_b}, y_s(k-1) \dots y_s(k-n_a)] \begin{bmatrix} \bar{\mathbf{b}}_{1_s}^T \\ \vdots \\ \bar{\mathbf{b}}_{n_{b_s}}^T \\ -a_{1_s} \\ \vdots \\ -a_{n_{a_s}} \end{bmatrix} \xi_s(k)
 \end{aligned}$$

onde:

$$\begin{aligned}
 \bar{\mathbf{b}}_{\alpha_s} &= [b_{\alpha_{1,s}} \dots b_{\alpha_{n_i,s}}], \quad \forall \alpha \in [1, 2, \dots, n_b] \\
 \mathbf{u}_k &= [u_1(k) \dots u_{n_i}(k)]
 \end{aligned}$$

A expressão geral para modelos lineares usada nesta tese pode ser derivada da equação anterior como:

$$\mathbf{y}_k = \mathbf{x}_{k-1} \boldsymbol{\theta} + \boldsymbol{\xi}_k \tag{A.2}$$

onde,

$$\begin{aligned}
 \mathbf{y}_k &= [y_1(k) \dots y_{n_o}(k)] \\
 \mathbf{x}_{k-1} &= [\mathbf{u}_{k-1} \dots \mathbf{u}_{k-n_b}, \mathbf{y}_{k-1} \dots \mathbf{y}_{k-n_a}] \\
 \boldsymbol{\theta} &= \begin{bmatrix} \mathbb{B}^T \\ \mathbb{A}_1 \\ \vdots \\ \mathbb{A}_{n_a} \end{bmatrix} \\
 \mathbb{B} &= \begin{bmatrix} \bar{b}_{1_1} & \dots & \bar{b}_{1_1} \\ \vdots & \ddots & \vdots \\ \bar{b}_{1_{n_o}} & \dots & \bar{b}_{n_b n_o} \end{bmatrix}
 \end{aligned}$$

$$\mathbb{A}_\beta = \begin{bmatrix} -a_{\beta_1} & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & -a_{\beta_{n_o}} \end{bmatrix}, \quad \forall \beta \in [1, 2, \dots, n_a]$$

$$\boldsymbol{\xi}(k) = [\xi_1(k) \dots \xi_{n_o}(k)] \quad (\text{A.3})$$

Apêndice B

Regime de fluxo multifásico de golfadas

A seguir será apresentado o regime de fluxo multifásico da golfada em tubulações risers.

B.1 Escoamento multifásico

Escoamentos multifásicos podem ser encontrados em diversas áreas tecnológicas, como por exemplo, nas indústrias química, alimentícia, de conversão de energia, processamento de materiais e de P&G (PALADINO, 2005). O escoamento multifásico é definido como o fluxo simultâneo de dois ou mais fluidos com propriedades diferentes e imiscíveis. O termo fase está relacionado com o número de interfaces presentes no sistema.

Na tese de (PALADINO, 2005), um sistema multifásico pode ser definido como uma região do espaço onde coexistem dois ou mais fluidos imiscíveis separados por uma interface, podendo esta ser conexa (escoamento estratificado, anular, etc.) ou desconexa (escoamento de bolhas, gotas, etc.) ou uma combinação de ambos os casos, onde a mesma fase pode aparecer em forma contínua ou dispersa (padrão anular com gotas, etc.).

Quando óleo, gás e água fluem simultaneamente em uma tubulação, as três fases podem se distribuir em diferentes configurações devido à diferença de densidades. A maneira com a qual as fases estão distribuídas depende das condições de operação, tais quais velocidade das fases, ângulo de inclinação da tubulação, etc. Na figura B.1 são apresentados alguns padrões para fluxos multifásicos horizontais e a seguir uma breve descrição (SILVA, 2000).

- **Estratificado** (Stratified Flow): caracterizado pela separação gravitacional total das fases líquida e gasosa com escoamento contínuo de cada fase. Acontece em velocidades muito baixas de líquido e gás.
- **Ondulado** (Wavy Flow): com o aumento da velocidade da fase gasosa no escoamento estratificado, ondas são formadas na interface de separação líquido-gás, ge-

rando o regime ondulado.

- **Anular** (Annular Flow): o escoamento anular ocorre com altas vazões de gás, onde o líquido escoar na superfície interna da tubulação, formando um filme com bolhas dispersas, e o gás escoar no centro da tubulação.

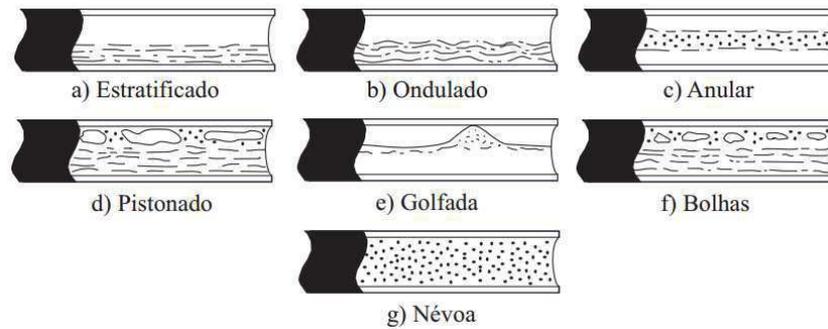


Figura B.1: Padrões de escoamento de fluxo multifásico em uma tubulação horizontal.

- **Pistonado** (Plug Flow): com o aumento da velocidade da fase gasosa no escoamento em bolhas, há o coalescimento, formando bolhas alongadas em forma de balas, que tendem a escoar no topo da tubulação.
- **Golfada** (Slug Flow): aumentando ainda mais a velocidade do gás no escoamento ondulado, as ondas passam a ocupar, em alguns casos, toda a seção transversal da tubulação formando golfadas, que são rapidamente deslocadas dentro da tubulação. Golfadas de gás e líquido se sucedem na tubulação com a de líquido arrastando pequenas bolhas dispersas.
- **Bolhas** (Bubble Flow): devido ao efeito gravitacional a fase gasosa em forma de bolhas discretas, tende a se dispersar no topo da tubulação com a fase líquida contínua.
- **Névoa** (Mist Flow): é caracterizado por velocidades superficiais de gás e líquido muito altas. Todo o líquido encontra-se disperso no núcleo do gás e as gotículas formadas pelo líquido viajam à mesma velocidade superficial do gás.

Os padrões de escoamento vertical mais comuns são apresentados na figura B.2 e descritos na sequência (MELAND, 2010).

- **Bolhas** (Bubble Flow): a fase gasosa se encontra dispersa na fase líquida em forma de bolhas discretas, sendo a fase líquida contínua. Esse escoamento ocorre tipicamente para baixas velocidades superficiais de gás.

- **Golfadas** (Slug Flow): aumentando a velocidade da fase gasosa, as bolhas coalescem e o diâmetro desta nova bolha atinge dimensão similar com a da tubulação. Quando isto ocorre, formam-se bolhas grandes em formato de bala, também denominadas bolhas de Taylor. Com isto, golfadas de gás e líquido se sucedem na tubulação com a golfada de líquido apresentando pequenas bolhas dispersas.
- **Agitado** (Churn Flow): com as velocidades de gás e líquido maiores do que no caso anterior, a quebra das bolhas do escoamento em golfadas conduz a um padrão instável e desordenado, onde existe um movimento oscilatório de líquido para cima e para baixo na tubulação.
- **Anular** (Annular Flow): o líquido escoa na periferia do duto formando um filme com bolhas dispersas, e o gás escoa no centro da tubulação, arrastando gotículas de líquido.

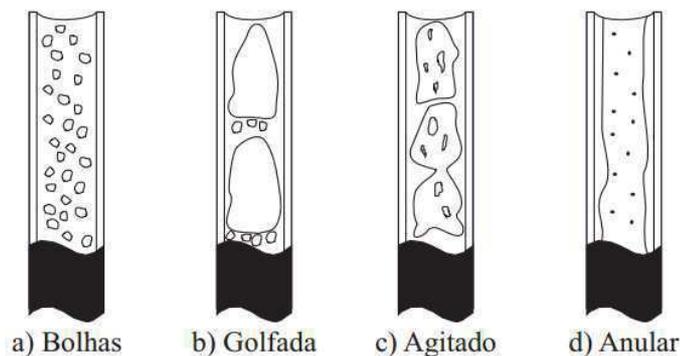


Figura B.2: Padrões de escoamento de fluxo multifásico em uma tubulação vertical.

B.1.1 Regime de golfadas

A instabilidade causada no sistema quando o fluxo assume um regime de golfadas é um dos principais problemas no campo de garantia de escoamento, conforme discutido na seção anterior. Este tipo de fluxo geralmente evidencia-se no fim do ciclo de vida de um poço, quando sua pressão interna e temperatura se reduziram desde quando iniciou sua exploração (SIVERTSEN; SKOGESTAD, 2005). Em tubulações com uma inclinação descendente seguida de um riser existe maior probabilidade de ocorrência de golfadas. Isto possibilita uma acumulação de líquido na entrada do riser, o que causa um bloqueio da passagem de gás para o riser. Em seguida, o gás no oleoduto será comprimido, aumentando sua pressão até que esta seja maior que a pressão hidrostática do riser. Quando isto acontece o gás se expande e lança o líquido que bloqueava o ponto baixo para fora da tubulação. O processo da golfada é tido como cíclico, pois uma vez que o gás foi liberado, a pressão diminui e o líquido volta a se acumular na base.

O mecanismo de formação de golfada no riser é ilustrado na figura B.3 e descrito nos seguintes passos: (STORKAAS; SKOGESTAD, 2003)

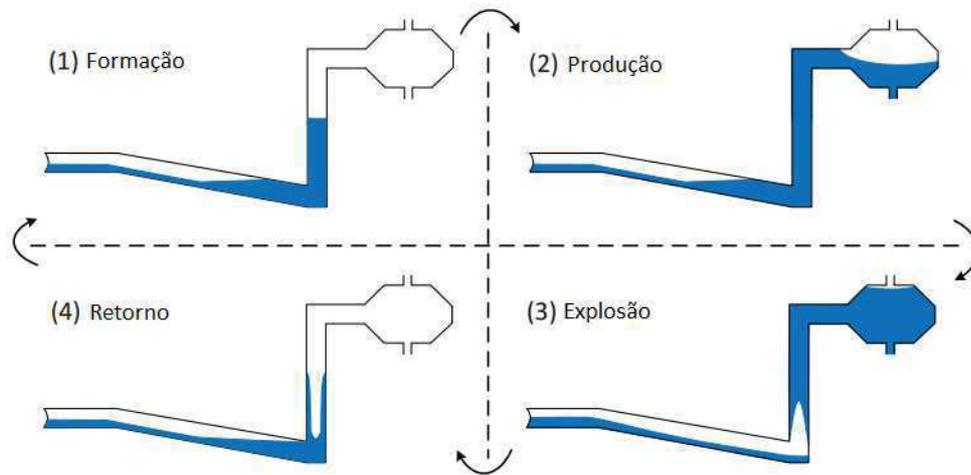


Figura B.3: Fases da golfada no riser.

1. **Formação:** quando as velocidades do gás e líquido são baixas o suficiente, o líquido acumula-se na base do riser devido à ação da gravidade, dando início à golfada.
2. **Produção:** enquanto a pressão hidrostática do líquido no riser for maior que a variação de pressão entre o topo do riser e a seção de alimentação do encanamento, a golfada continua crescendo.
3. **Explosão:** quando a pressão do gás na seção de alimentação do encanamento torna-se maior que o peso da coluna de líquido no riser ocorre a explosão ou seja, o líquido é lançado para fora do riser.
4. **Retorno:** a pressão na seção de alimentação diminui e o líquido volta a se acumular no ponto-baixo, reiniciando o ciclo.

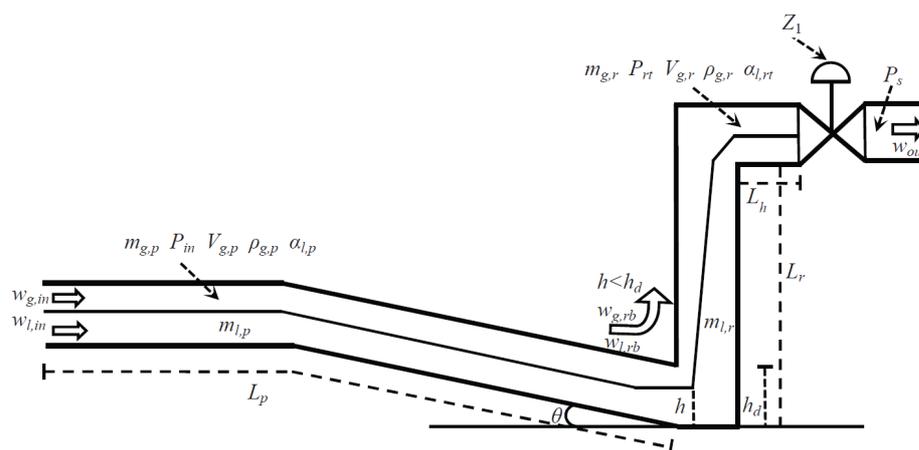
Este regime cíclico ocasiona um fluxo intermitente na saída do riser, o que pode reduzir a capacidade de produção. Problemas maiores, como parada de produção, podem ser ocasionados quando grandes volumes de líquido chegam ao vaso separador, inundando-o. As plataformas de petróleo lidam com grandes quantidades de um produto muito valioso, por isto, uma redução da produção ou pequenas paradas no processo refletem em perdas financeiras substanciais. A título de exemplo, na plataforma P-55 na Bacia de Campos no Rio de Janeiro, uma parada indevida de um trem de produção de 30 minutos resulta numa perda média de 1.875 barris. Considerando o valor do barril de petróleo como sendo US\$60,00, esta perda de produção custaria US\$112.500,00 (CAMPOS; LOUREIRO; B., 2006).

B.2 Modelo de Jahanshahi

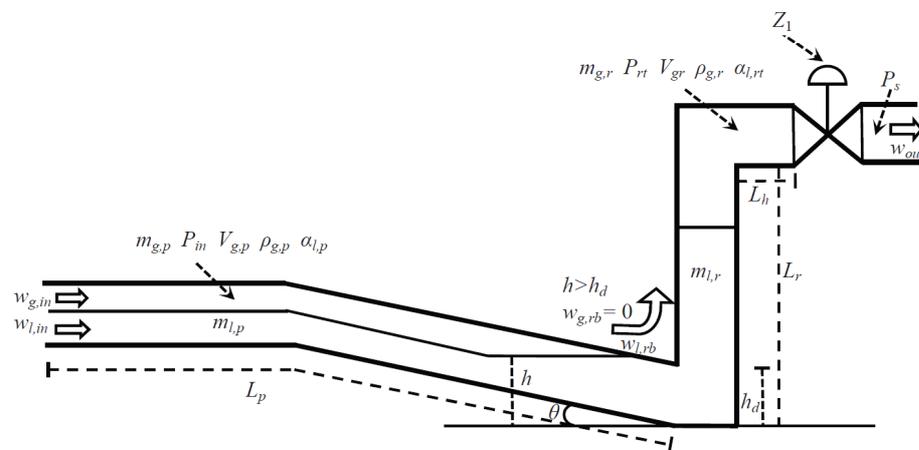
Nesta seção será descrito o modelo com quatro estados e quatro parâmetros de sintonia, do processo da golfada desenvolvido por Jahanshahi em (JAHANSHAH, 2013). Em seguida são apresentados os resultados da simulação do modelo não linear.

A mistura de água e óleo extraída dos poços será considerada como um único componente e será referida a partir daqui como mistura, o outro componente sendo o gás. Neste modelo, Existem duas fases para cada componente: uma fase da mistura na base e no topo do riser; uma fase do gás na base e no topo do riser.

A representação do sistema oleoduto-riser da figura B.4 será utilizada como base para a obtenção das equações do modelo. Nesta figura o sistema é ilustrado em dois momentos distintos, o primeiro é quando o regime de fluxo permite a passagem de gás pelo ponto baixo e o segundo quando o líquido preenche o ponto baixo do sistema bloqueando a passagem de gás.



(a) Representação simplificada do regime desejado



(b) Representação simplificada do bloqueio do líquido que leva a golfada

Figura B.4: Sistema oleoduto-riser

Onde $\omega_{g,in}$ e $\omega_{L,in}$ são os fluxos de massa de gás e líquido na entrada do sistema e ω_{out}

o fluxo de massa da mistura na saída. $\omega_{g,rb}$ e $\omega_{l,rb}$ são os fluxos de gás e líquido no ponto baixo do riser respectivamente. As seguintes propriedades são indicadas diferenciando os componentes e suas fases por dois índices: m é a massa, V o volume, ρ a densidade, α a fração de massa do componente. O primeiro índice referencia o componente, g sendo o gás e l o líquido, e o segundo índice referencia a fase, p indica a fase do componente no oleoduto e r a fase do componente no riser. Além disso, P_{in} é a pressão do gás no oleoduto, P_{rt} é a pressão do gás no topo do riser e P_s a pressão do separador (Pressão após a válvula de saída).

Os demais parâmetros do modelo, juntamente com seus valores padrões, são descritos na tabela B.1. Os quatro parâmetros de sintonia descritos a seguir podem ser usados para ajustar o modelo a dados numéricos ou experimentais de um sistema oleoduto-riser.

- K_h : fator de correção para o nível de líquido no oleoduto
- C_{v1} : constante de produtividade da válvula de saída
- K_g : coeficiente para o fluxo de gás através do ponto baixo
- K_l : coeficiente para o fluxo de líquido através do ponto baixo

Tabela B.1: Parâmetros do sistema oleoduto-riser

| Símbolo | Descrição | Valores | Unidades |
|-----------|---|------------------------|--------------|
| R | constante universal dos gases | 8314 | $J/(kmol.K)$ |
| g | gravidade | 9,81 | m/s^2 |
| μ | viscosidade | $1,426 \times 10^{-4}$ | $Pa.s$ |
| ρ_l | densidade do líquido | 832,2 | kg/m^3 |
| M_g | peso molecular do gás | 20 | gr |
| P_{res} | pressão do reservatório | 320 | bar |
| T_p | temperatura do oleoduto | 337 | K |
| V_p | volume do oleoduto | 48,63 | m^3 |
| D_p | Diâmetro do oleoduto | 0,12 | m |
| L_p | comprimento do oleoduto | 4300 | m |
| T_r | temperatura do riser | 298,3 | K |
| V_r | volume do riser | 3,14 | m^3 |
| D_r | diâmetro do riser | 0,1 | m |
| L_r | comprimento do riser | 300 | m |
| L_h | comprimento da seção horizontal | 100 | m |
| P_s | pressão do separador | 50,1 | bar |
| K_h | fator de correção do nível | 0,7 | - |
| K_g | orifício do fluxo de gás no ponto baixo | $3,49 \times 10^{-2}$ | - |
| K_l | orifício do fluxo de líquido no ponto baixo | $2,81 \times 10^{-1}$ | - |
| C_{v1} | constante da válvula de saída | $1,16 \times 10^2$ | - |

B.2.1 Equações de conservação de massa para oleoduto e riser

As quatro equações diferenciais do modelo advêm diretamente da lei de conservação de massa para o gás e o líquido no oleoduto e no riser: (JAHANSHAH, 2013)

$$\frac{dm_{g,p}}{dt} = \omega_{g,in} - \omega_{g,rb} \quad (\text{B.1a})$$

$$\frac{dm_{l,p}}{dt} = \omega_{l,in} - \omega_{l,rb} \quad (\text{B.1b})$$

$$\frac{dm_{g,r}}{dt} = \omega_{g,rb} - \omega_{g,out} \quad (\text{B.1c})$$

$$\frac{dm_{l,r}}{dt} = \omega_{l,rb} - \omega_{l,out} \quad (\text{B.1d})$$

O New Model criado por Jahanshahi possui quatro variáveis de estado referentes \tilde{A} s diferentes fases do fluido, são elas:

- $m_{g,p}$: massa de gás no oleoduto [kg]
- $m_{l,p}$: massa de líquido no oleoduto [kg]
- $m_{g,r}$: massa de gás no riser [kg]
- $m_{l,r}$: massa de líquido no riser [kg]

B.2.2 Condições de entrada

Nas equações (B.1a) e (B.1b), $\omega_{g,in}$ e $\omega_{l,in}$ são os fluxos de massa de gás e líquido na entrada do sistema. Assume-se que esses fluxos são constantes, porém tal condição de contorno pode ser facilmente modificada para, por exemplo, definir um fluxo de entrada dependente da pressão. (JAHANSHAH, 2013)

B.2.3 Condições de saída

Assume-se que a pressão do separador P_s é constante e que o fluxo de saída da mistura bifásica é dado pela equação simplificada da válvula:

$$\omega_{out} = C_{v1} f(z_1) \sqrt{\rho_{rt} \max(P_{rt} - P_s, 0)} \quad (\text{B.2})$$

Onde $0 \leq z_1 \leq 1$ é a abertura da válvula normalizada e $f(z_1)$ é a equação característica da válvula. Nas simulações serão usadas válvulas lineares, i.e. $f(z_1) = z_1$, porém esta

condição pode ser modificada para outras válvulas. Os fluxos individuais de líquido e gás são calculados por:

$$\omega_{l,out} = \alpha_{l,rt}^m \omega_{out} \quad (\text{B.3})$$

$$\omega_{g,out} = (1 - \alpha_{l,rt}^m) \omega_{out} \quad (\text{B.4})$$

onde $\alpha_{l,rt}^m$, a fração de massa de líquido no topo do riser, vale:

$$\alpha_{l,rt}^m = \frac{\alpha_{l,rt} \rho_l}{\alpha_{l,rt} \rho_l + (1 - \alpha_{l,rt}) \rho_{g,r}} \quad (\text{B.5})$$

O termo no denominador da equação (B.5) é a densidade ρ_{rt} da mistura no topo do riser, que também aparece na equação (B.2). A fração de volume de líquido no riser $\alpha_{l,rt}$ é descrita na equação (B.44).

B.2.4 Modelo do oleoduto

O oleoduto é definido como a tubulação horizontal que estende-se desde o poço até a base do riser. Nele, a fração de volume de líquido é dada por:

$$\alpha_l = \frac{\alpha_l^m / \rho_l}{\alpha_l^m / \rho_l + (1 - \alpha_l^m) / \rho_g} \quad (\text{B.6})$$

A fração média de massa de líquido nesta tubulação vem das condições de contorno na entrada:

$$\bar{\alpha}_{l,p}^m = \frac{\omega_{l,in}}{\omega_{l,in} + \omega_{g,in}} \quad (\text{B.7})$$

Já a fração média de volume de líquido no oleoduto é:

$$\bar{\alpha}_{l,p} = \frac{\bar{\rho}_{g,p} \omega_{l,in}}{\bar{\rho}_{g,p} \omega_{l,in} + \rho_l \omega_{g,in}} \quad (\text{B.8})$$

O cálculo da densidade média de gás $\bar{\rho}_{g,p}$ é feito a partir da pressão nominal (estado estacionário) no oleoduto, assumindo que o gás é ideal:

$$\bar{\rho}_{g,p} = \frac{P_{in,nom} M_g}{RT_p} \quad (\text{B.9})$$

$P_{in,nom}$ por sua vez é dependente de $\bar{\alpha}_{l,p}$ e por isso, é calculado a partir de uma inicialização geral do modelo no estado estacionário. Usando a equação (B.9) e taxas de fluxo de entrada constantes (nominais), resulta em $\bar{\alpha}_{l,p}$ constante.

A área da seção transversal do oleoduto é:

$$A_p = \frac{\pi}{4} D_p^2 \quad (\text{B.10})$$

onde D_p é o diâmetro do oleoduto, com isso o volume do oleoduto será $V_p = A_p L_p$.

Quando o gás e líquido estão distribuídos homoganeamente ao longo do oleoduto, a massa de líquido presente no oleoduto é dada por:

$$\bar{m}_{l,p} = \rho_l V_p \bar{\alpha}_{l,p} \quad (\text{B.11})$$

Com esta consideração, o nível do líquido no ponto baixo do oleoduto será:

$$\bar{h} = K_h h_d \bar{\alpha}_{l,p} \quad (\text{B.12})$$

onde $h_d = D_p / \cos(\theta)$ é a abertura do oleoduto na base do riser, θ a inclinação do oleoduto e K_h um fator de correção que pode ser utilizado para fazer um ajuste fino no modelo.

Supondo que haja um aumento de $\Delta m_{l,p}$ na massa do líquido, um comprimento ΔL do oleoduto será preenchido, apenas com líquido, a partir do ponto baixo. Assim:

$$\Delta m_{l,p} = m_{l,p} - \bar{m}_{l,p} = \Delta L A_p (1 - \bar{\alpha}_{l,p}) \rho_l \quad (\text{B.13})$$

O nível de líquido torna-se $h = \bar{h} + \Delta L \sin(\theta)$ ou ainda:

$$h = \bar{h} + \frac{m_{l,p} - \bar{m}_{l,p}}{A_p (1 - \bar{\alpha}_{l,p}) \rho_l} \sin(\theta) \quad (\text{B.14})$$

Com isso o nível de líquido h pôde ser escrito como uma função da massa de líquido no oleoduto, que é uma variável de estado do modelo. O restante dos parâmetros da equação (B.14) são constantes.

A equação densidade do gás no oleoduto é:

$$\rho_{g,p} = \frac{m_{g,p}}{V_{g,p}} \quad (\text{B.15})$$

onde o volume ocupado pelo gás no oleoduto é:

$$V_{g,p} = V_p - m_{l,p} / \rho_l \quad (\text{B.16})$$

A pressão na entrada do oleoduto, assumindo que o gás é ideal, será:

$$P_{in} = \frac{\rho_{g,p} R T_p}{M_g} \quad (\text{B.17})$$

Apenas a fase líquida é considerada no cálculo da queda de pressão ocasionada pelo atrito no oleoduto, que é dada por:

$$\Delta P_{fp} = \frac{\bar{\alpha}_{l,p} \lambda_p \rho_l \bar{U}_{sl,in}^2 L_p}{2D_p} \quad (\text{B.18})$$

onde λ_p é o fator de atrito do oleoduto, que pode ser computado por uma aproximação explícita da equação de Colebrook-White:

$$\frac{1}{\sqrt{\lambda_p}} = 1,8 \log_{10} \left[\left(\frac{\epsilon/D_p}{3,7} \right)^{1,11} + \frac{6,9}{Re_p} \right] \quad (\text{B.19})$$

aqui, o número de Reynolds é dado por:

$$Re_p = \frac{\rho_l \bar{U}_{sl,in} D_p}{\mu} \quad (\text{B.20})$$

onde μ é a viscosidade do líquido e $\bar{U}_{sl,in}$ é a velocidade superficial do líquido, que é dada por:

$$\bar{U}_{sl,in} = \frac{4\omega_{l,in}}{\pi D_p^2 \rho_l} \quad (\text{B.21})$$

B.2.5 Modelo do riser

O riser é a tubulação vertical que liga a plataforma ao fundo do oceano. O volume total do riser é dado por:

$$V_r = A_r(L_r + L_h) \quad (\text{B.22})$$

onde,

$$A_r = \frac{\pi}{4} D_r^2 \quad (\text{B.23})$$

O volume ocupado pelo gás no riser é:

$$V_{g,r} = V_r - \frac{m_{l,r}}{\rho_l} \quad (\text{B.24})$$

A densidade do gás no riser é obtida por:

$$\rho_{g,r} = \frac{m_{g,r}}{V_{g,r}} \quad (\text{B.25})$$

A pressão no topo do riser é derivada da lei dos gases ideais:

$$P_{rt} = \frac{\rho_{g,r} R T_r}{M_g} \quad (\text{B.26})$$

A fração média de volume de líquido no riser é dada por:

$$\bar{\alpha}_{l,r} = \frac{m_{l,r}}{V_r \rho_l} \quad (\text{B.27})$$

A densidade média da mistura no interior do riser é:

$$\bar{\rho}_{m,r} = \frac{m_{g,r} + m_{l,r}}{V_r} \quad (\text{B.28})$$

A queda de pressão, devido ao atrito, entre as fases no riser é dada por:

$$\Delta P_{fr} = \frac{\bar{\alpha}_{l,r} \lambda_r \bar{\rho}_{m,r} \bar{U}_m^2 (L_r + L_h)}{2D_r} \quad (\text{B.29})$$

O fator de atrito do riser usa a mesma correlação do oleoduto:

$$\frac{1}{\sqrt{\lambda_r}} = -1,8 \log_{10} \left[\left(\frac{\epsilon/D_r}{3,7} \right)^{1,11} + \frac{6,9}{Re_r} \right] \quad (\text{B.30})$$

onde o número de Reynolds no riser é dado por:

$$Re_r = \frac{\bar{\rho}_{m,r} \bar{U}_m D_r}{\mu} \quad (\text{B.31})$$

A velocidade da mistura no riser equivale \tilde{A} soma da velocidade superficial do líquido e do gás:

$$\bar{U}_m = \bar{U}_{sl,r} + \bar{U}_{sg,r} \quad (\text{B.32})$$

onde as velocidades de líquido e gás são respectivamente:

$$\bar{U}_{sl,r} = \frac{\omega_{l,in}}{\rho_l A_r} \quad (\text{B.33})$$

$$\bar{U}_{sg,r} = \frac{\omega_{g,in}}{\rho_{g,r} A_r} \quad (\text{B.34})$$

B.2.6 Modelo do fluxo de gás na base do riser

Conforme ilustrado na figura (B.4), quando o nível de líquido excede a abertura do oleoduto no ponto baixo ($h > h_d$), o líquido irá bloquear a passagem de gás naquele ponto e o fluxo de gás será zero:

$$\omega_{g,rb} = 0, \quad h \geq h_d \quad (\text{B.35})$$

Quando o líquido não está bloqueando o ponto baixo ($h < h_d$), haverá um fluxo de gás do volume $V_{g,p}$ para o $V_{g,r}$ a uma taxa de $\omega_{g,rb}$ [kg/s]. Assume-se que este fluxo tem

a equação simplificada da válvula:

$$\omega_{g,rb} = K_g A_g \sqrt{\rho_{g,p} \Delta P_g}, \quad h < h_d \quad (\text{B.36})$$

onde,

$$\Delta P_g = P_{in} - \Delta P_{fp} - P_{rt} - \bar{\rho}_{m,r} g L_r - \Delta P_{fr} \quad (\text{B.37})$$

A área livre A_g para a passagem de gás pode ser calculada precisamente utilizando funções trigonométricas (STORKAAS; SKOGESTAD, 2003). Porém, por simplicidade, será utilizada a seguinte aproximação quadrática:

$$A_g = A_p \left(\frac{h_d - h}{h_d} \right)^2, \quad h < h_d \quad (\text{B.38a})$$

$$A_g = 0, \quad h < h_d \quad (\text{B.38b})$$

B.2.7 Modelo do fluxo de líquido na base do riser

O fluxo de massa de líquido na base do riser também é descrito pela equação simplificada da válvula:

$$\omega_{l,rb} = K_l A_l \sqrt{\rho_l \Delta P_l} \quad (\text{B.39})$$

onde a variação de pressão do líquido é obtida como:

$$\Delta P_l = P_{in} - \Delta P_{fp} + \rho_l g h - P_{rt} - \bar{\rho}_{m,r} g L_r - \Delta P_{fr} \quad (\text{B.40})$$

e a área da seção transversal da passagem do líquido vale:

$$A_l = A_p - A_g \quad (\text{B.41})$$

B.2.8 Modelo de distribuição das fases na saída da válvula

A fim de calcular os fluxos de massa das fases individualmente em (B.3) e (B.4), a distribuição das fases no topo do riser deve ser conhecida.

A fração de volume de líquido no topo do riser, $\alpha_{l,rt}$, que é usada na equação (B.5), pode ser calculada a partir do modelo de arrastamento proposto em (STORKAAS; SKOGESTAD, 2003), porém suas equações são complicadas. Como alternativa, utiliza-se o fato de que em um oleoduto vertical, existe uma relação aproximadamente linear

entre a pressão e a fração de volume de líquido. Além disso, o gradiente de pressão é assumido constante ao longo do riser para o regime de fluxo desejável (sem golfadas). Consequentemente o gradiente de fração de volume de líquido também será constante, i.e. $\frac{\delta\alpha_{l,r}}{\delta y} = \text{constante}$. Assim, a fração média de volume de líquido no riser é:

$$\bar{\alpha}_{l,r} = \frac{\alpha_{l,rb} + \alpha_{l,rt}}{2} = \frac{m_{l,r}}{V_r \rho_l} \quad (\text{B.42})$$

aqui, $\alpha_{l,rb}$ é o mesmo da equação (B.27) e $\alpha_{l,rt}$ é determinado a partir da área do fluxo de líquido na base do riser:

$$\bar{\alpha}_{l,rb} = \frac{A_l}{A_p} \quad (\text{B.43})$$

Finalmente, a fração de volume de líquido no topo do riser é dada por:

$$\alpha_{l,rt} = 2\bar{\alpha}_{l,r} - \alpha_{l,rb} = \frac{2m_{l,r}}{V_r \rho_l} \quad (\text{B.44})$$