

Identificação de Sistema por Erro de Predição com Modelos Polinomiais

Guilherme de Paoli Beal
Universidade Federal do Rio Grande do Sul

Resumo—Este trabalho busca identificar um sistema por erro de predição, a partir de um conjunto de dados de entrada e saída, com a saída corrompida por ruído de medição. As classes de modelo ARX, ARMAX, Output Error e Box-Jenkins, com diferentes ordens, são avaliadas. As identificações são realizadas pelo pacote `pysid` em linguagem Python. Os resultados são comparados pelo erro quadrático médio de predição e pelo critério de informação de Akaike ajustado para dados com poucas amostras.

Palavras-chave—Identificação por Erro de Predição, `pysid`, ARX, ARMAX, Output Error, Box-Jenkins

I. INTRODUÇÃO

Este trabalho visa à aplicação de identificação de um sistema discreto, linear e invariante no tempo, através do método de erro de predição, ou *Prediction Error Method* (PEM). Uma batelada de dados de um sistema, supostamente desconhecido, é fornecida, em que a saída está contaminada por ruído de medição. A identificação é realizada utilizando diferentes classes de modelo e ordens e os resultados são comparados.

As implementações são desenvolvidas em Python, versão 3.9.12. A identificação por erro de predição utiliza o pacote `pysid` em versão de desenvolvimento 0.1.0. O código deste projeto está publicado em github.com/GuiBeal/system-identification.

II. IDENTIFICAÇÃO POR ERRO DE PREDIÇÃO

Considere um sistema discreto, linear, invariante no tempo, com uma única entrada e uma única saída. A resposta deste sistema é dada por

$$y(t) = G_0(q) u(t) + \nu(t), t \in \mathbb{N}, \quad (1)$$

em que $y(t)$ é o sinal de saída, $G_0(q)$ é a função de transferência do sistema, $u(t)$ é o sinal de entrada, $\nu(t)$ é um ruído de medição desconhecido, q é o operador de avanço — de modo que $q x(t) = x(t+1)$ — e $t \in \mathbb{N}$ é a variável de tempo discreto. O ruído de medição é gerado a partir de ruído branco filtrado, como

$$\nu(t) = H_0(q) e(t),$$

em que H_0 é a função de transferência do filtro e $e(t)$ é ruído branco.

G. Beal, guilherme.beal@ufrgs.br

A. Modelos

O modelo busca representar o sistema real $G_0(q)$ expresso em (1). Em particular, na identificação por erro de predição, o modelo procurar caracterizar também o ruído $\nu(t)$ através da identificação de $H_0(q)$.

Quatro diferentes classes de modelo são exploradas neste trabalho:

- Autorregressivo com Entrada Externa, ou *Autoregressive with Extra Input* (ARX);
- Autorregressivo com Média Móvel e Entrada Externa, ou *Autoregressive Moving Average with Extra Input* (ARMAX);
- Erro na Saída, ou *Output Error* (OE); e
- Box-Jenkins (BJ).

Essas classes são casos particulares de um modelo mais genérico, definido por

$$A(q) y(t) = \frac{B(q)}{F(q)} u(t) + \frac{C(q)}{D(q)} e(t). \quad (2)$$

Os polinômios têm os formatos

$$\begin{aligned} A(q) &= 1 + a_1 q^{-1} + \dots + a_{n_a} q^{-n_a}, \\ B(q) &= q^{-n_k} (b_0 + b_1 q^{-1} + \dots + b_{n_b} q^{-n_b}), \\ C(q) &= 1 + c_1 q^{-1} + \dots + c_{n_c} q^{-n_c}, \\ D(q) &= 1 + d_1 q^{-1} + \dots + d_{n_d} q^{-n_d}, \text{ e} \\ F(q) &= 1 + f_1 q^{-1} + \dots + f_{n_f} q^{-n_f}, \end{aligned}$$

em que n_a , n_b , n_c , n_d e n_f são suas ordens e n_k é o número de atrasos da entrada para a saída. Definindo as funções de transferência

$$\begin{aligned} G(q) &= \frac{B(q)}{A(q) F(q)}, \text{ e} \\ H(q) &= \frac{C(q)}{A(q) D(q)}, \end{aligned}$$

então (2) pode ser reescrita como

$$y(t) = G(q) u(t) + H(q) e(t).$$

Note que $G(q)$ visa a caracterizar o sistema real $G_0(q)$. Por sua vez, o ruído de medição $\nu(t)$ é representado como ruído branco $e(t)$ filtrado por $H(q)$.

Pelas definições dos polinômios,

$$H(\infty) = 1. \quad (3)$$

Outrossim, se o modelo representa um sistema amostrado, então $G(q)$ deve ser estritamente própria — isto é, o grau

de seu denominador é maior que o de seu numerador — o que é garantido com $n_k \geq 1$.

Ressalta-se que este modelo genérico pode apresentar definições e notações diferentes, como em [1]–[3]. A definição aqui apresentada corresponde ao formato dos polinômios retornados pelas funções do pacote `pysid`.

B. Predição

A predição é realizada aplicando

$$\hat{y}(t) = L_u(q) u(t) + L_y(q) y(t) \quad (4)$$

com

$$L_u(q) = \frac{G(q)}{H(q)}, \text{ e}$$

$$L_y(q) = 1 - \frac{1}{H(q)}.$$

Embora isso não seja diretamente evidenciado por (4), o fato de $G(q)$ ser estritamente própria juntamente com (3) garante que a predição $\hat{y}(t)$ no instante $t = t_0$ depende somente de valores de $u(t)$ e $y(t)$ em instantes $t < t_0$ — isto é, a predição é realizado a partir de valores anteriores dos sinais de entrada e saída medidos.

O erro de predição é definido como

$$\varepsilon(t) = y(t) - \hat{y}(t) = \frac{y(t) - G(q) u(t)}{H(q)}.$$

Assim, o erro quadrático médio de predição é definido por

$$J = \frac{1}{N} \sum_{t=1}^N (\varepsilon(t))^2, \quad (5)$$

em que N é o número de amostras preditas.

C. Identificação

A identificação por erro de predição visa, a partir de um conjunto de dados de entrada $u(t)$ e de saída $y(t)$ medidos do processo, a identificar os parâmetros dos polinômios do modelo de forma a minimizar o custo expresso em (5). O algoritmo aplicado neste processo depende da classe do modelo e está fora do escopo deste trabalho. Estes são implementados pelo pacote `pysid`.

D. Classes de Modelos

As quatro classes de modelo são obtidas fixando determinadas ordens no modelo genérico.

1) *ARX*: Neste modelo, há liberdade nas escolhas de n_a , n_b e n_k , enquanto $n_c = n_d = n_f = 0$. Assim, as funções de transferência tornam-se

$$G(q) = \frac{B(q)}{A(q)}, \text{ e}$$

$$H(q) = \frac{1}{A(q)}.$$

Esta classe tem a vantagem de fazer com que (5) seja quadrática nos parâmetros, os quais podem, portanto, ser calculados por mínimos quadrados.

2) *ARMAX*: Este modelo requer a arbitração das ordens n_a , n_b , n_c e n_k , com $n_d = n_f = 0$. Portanto, as funções de transferência são

$$G(q) = \frac{B(q)}{A(q)}, \text{ e}$$

$$H(q) = \frac{C(q)}{A(q)}.$$

3) *Output Error*: Para este modelo são escolhidas as ordens n_b , n_f e n_k , fixando $n_a = n_c = n_d = 0$. Nesse caso, tem-se

$$G(q) = \frac{B(q)}{F(q)}, \text{ e}$$

$$H(q) = 1.$$

Note que este modelo considera que o ruído de medição é ruído branco.

4) *Box-Jenkins*: Finalmente, neste modelo arbitra-se n_b , n_c , n_d , n_f e n_k , com $n_a = 0$. Com isso, obtém-se

$$G(q) = \frac{B(q)}{F(q)}, \text{ e}$$

$$H(q) = \frac{C(q)}{F(q)}.$$

III. CRITÉRIO DE INFORMAÇÃO DE AKAIKE

Para avaliar a qualidade dos modelos identificados será utilizado o critério de informação de Akaike, ou *Akaike Information Criterion* (AIC). Este critério pondera a qualidade de predição junto à ordem do modelo, penalizando critérios de maior ordem. Considerando o erro de predição, o critério é definido por

$$\text{AIC} = N \log(J) + 2k,$$

em que N é o número de amostras preditas, k é o número de parâmetros do modelo e J é o erro médio quadrático de predição definido em (5) [4].

Para conjuntos com poucas amostras, pode-se utilizar um critério adaptado, definido por

$$\text{AICc} = \text{AIC} + \frac{2k(k+1)}{N-k-1}. \quad (6)$$

Note que $N \rightarrow \infty \implies \text{AICc} \rightarrow \text{AIC}$ — isto é, o critério adaptado se aproxima do critério original à medida que cresce o número de amostras considerado.

IV. DADOS

Os dados são apresentados na Figura 1, em que o sinal $u(t)$ é multiplicado por um fator de 10 para melhor visualização. Ambos os sinais $u(t)$ e $y(t)$ contêm 200 amostras. Note que a entrada $u(t)$ é similar a uma onda quadrada.

Na Figura 1, a linha tracejada vertical divide os dados em dois conjuntos. Para a identificação dos modelos, utilizam-se os dados à esquerda desta linha, correspondentes às primeiras 160 amostras. As 40 amostras restantes são aplicadas na validação dos modelos identificados.

A Figura 2 mostra o espectro de amplitude dos dados — isto é, a magnitude da transformada de Fourier dos sinais — em escala logarítmica. Note como ambos os sinais possuem um conteúdo elevado em baixas frequências, seguidos de picos pontuais que decrescem com o aumento da frequência.

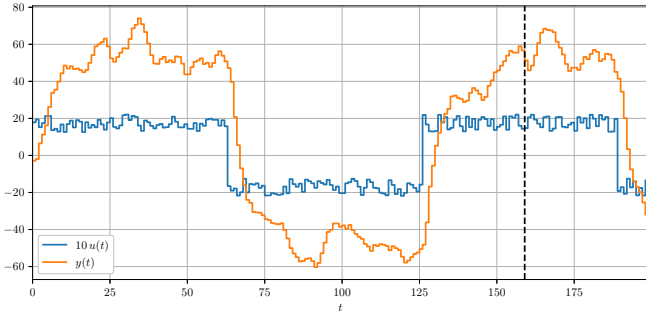


Figura 1. Dados de entrada e saída.

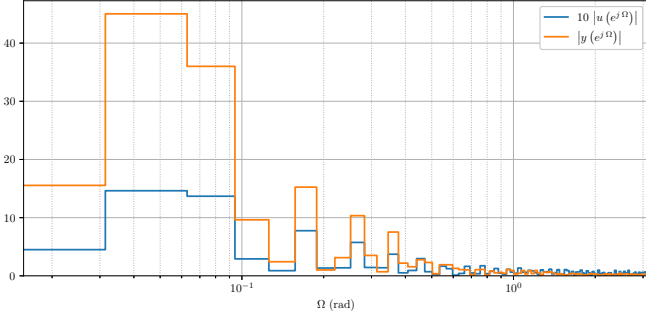


Figura 2. Espectro de amplitude dos dados de entrada e saída.

V. RESULTADOS

Diversas identificações são realizadas, variando as ordens dos polinômios e o atraso entre as quatro classes de modelos. A Tabela I sintetiza os intervalos de variação desses valores para cada modelo. A fim de considerar sistemas amostrados, todos os sistemas tem atraso de pelo menos um amostra — isto é, $n_k \geq 1$.

Tabela I
INTERVALO DE VARIAÇÃO DAS ORDENS DOS POLINÔMIOS

Classe	n_a	n_b	n_c	n_d	n_f	n_k	Total
ARX	[1, 4]	[0, 4]	—	—	—	[1, 4]	80
ARMAX	[1, 4]	[0, 4]	[1, 4]	—	—	[1, 4]	320
OE	—	[0, 4]	—	—	[1, 4]	[1, 4]	80
BJ	—	[0, 4]	[0, 4]	[1, 4]	[1, 4]	[1, 4]	1600
Total							2080

Destaca-se que dentre as 1600 identificações realizadas com a classe BJ 960 apresentaram algum problema de implementação interno ao pacote `pysid`. Nesses casos, os modelos foram descartados. Assim, o total de modelos efetivamente identificados é de 1120.

Para cada identificação são calculados o erro quadrático médio, conforme (5), e o critério de informação de Akaike para pequenos conjuntos de dados, conforme (6). Ambos são realizados tanto para o conjunto de dados de identificação — denotado pelo subíndice i — como para o conjunto de validação — denotado pelo subíndice v .

Para cada uma das quatro classes, a Tabela II apresenta os seis melhores modelos resultantes, com base no critério

de informação de Akaike com os dados de validação — isto é, $AICc_v$. Os modelos são unicamente identificados por um índice, exibidos na primeira coluna.

Comparando a ordem de grandeza de cada critério de qualidade na Tabela II, nota-se que os modelos ARX, ARMAX e BJ apresentam um desempenho consideravelmente melhor que os modelos OE; isto é um indicativo de que o ruído de medição não é branco, conforme supõe esse último modelo. Além disso, para os modelos ARX, ARMAX e BJ, todos as identificações presentes na Tabela II têm $n_k = 1$ — isto é, um atraso da entrada para a saída de apenas uma amostra. Destaca-se, ainda, que entre os modelos ARMAX e BJ presentes na Tabela II todas as identificações têm n_c partindo de seu valor mínimo — 1 para ARMAX e 0 para BJ —, levando à interpretação de que o numerador de $H(q)$ tem ordem pequena.

A Tabela III mostra as 20 melhores identificações, novamente classificadas pelo critério de informação de Akaike adaptado com os dados de validação — $AICc_v$. Note a prevalência de modelos ARX, bem como das características discutidas no parágrafo anterior.

A partir das identificações, são obtidas as funções de transferência do sistema $G(q)$ e do ruído $H(q)$. A fim de compará-las visualmente, as Figuras 3 e 4 apresentam suas respostas em frequência, considerando as identificações na Tabela III.

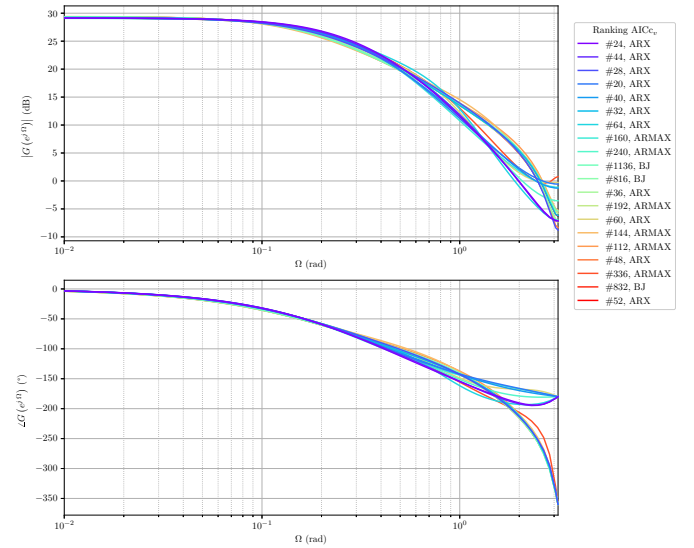


Figura 3. Respostas em frequência de $G(q)$ a partir das identificações na Tabela III.

Note, através da Figura 3, que, entre todas as identificações, o comportamento do sistema $G(q)$ é semelhante nas baixas frequências mas discordante nas altas frequências. Isso está em consonância com o fato de que os dados trazem mais informações nas baixas frequências, conforme visto na Figura 2. Por outro lado, a Figura 4 mostra que o sistema $H(q)$ é mais diverso entre as diferentes identificações.

A. Escolha de Modelos Adequados

Através das observações anteriores, procura-se arbitrar classes de modelo razoáveis para representar o sistema efetivo,

Tabela II
MELHORES RESULTADOS POR CLASSE COM BASE NO CRITÉRIO $AICc_v$

#	Classe	n_a	n_b	n_c	n_d	n_f	n_k	$AICc_v$	$AICc_i$	J_v	J_i
24	ARX	2	1	–	–	–	1	79,463	288,142	5,801	5,750
44	ARX	3	1	–	–	–	1	80,358	289,971	5,556	5,740
28	ARX	2	2	–	–	–	1	80,596	276,503	5,589	5,276
20	ARX	2	0	–	–	–	1	80,820	290,423	6,384	5,910
40	ARX	3	0	–	–	–	1	81,303	291,572	6,074	5,875
32	ARX	2	3	–	–	–	1	82,077	278,289	5,410	5,264
160	ARMAX	2	0	1	–	–	1	82,406	292,387	6,244	5,905
240	ARMAX	3	0	1	–	–	1	82,440	291,698	5,853	5,802
192	ARMAX	2	2	1	–	–	1	83,302	278,685	5,579	5,277
144	ARMAX	1	4	1	–	–	1	83,435	299,330	5,199	5,922
112	ARMAX	1	2	1	–	–	1	83,574	294,450	6,021	5,902
336	ARMAX	4	1	1	–	–	1	84,031	282,669	5,277	5,337
465	OE	–	4	–	–	1	2	203,538	635,576	112,710	49,103
469	OE	–	4	–	–	2	2	205,126	637,238	108,924	48,942
466	OE	–	4	–	–	1	3	208,209	688,227	126,671	68,237
468	OE	–	4	–	–	2	1	210,717	620,915	125,264	44,195
470	OE	–	4	–	–	2	3	211,195	690,408	126,771	68,234
464	OE	–	4	–	–	1	1	211,208	622,115	136,532	45,141
1136	BJ	–	2	0	2	1	1	82,909	276,778	5,524	5,214
816	BJ	–	1	0	2	1	1	83,040	273,646	5,941	5,183
832	BJ	–	1	0	3	1	1	84,041	276,120	5,683	5,193
1456	BJ	–	3	0	2	1	1	84,939	277,505	5,398	5,167
1140	BJ	–	2	0	2	2	1	85,844	278,965	5,521	5,214
848	BJ	–	1	0	4	1	1	86,957	274,874	5,677	5,083

Tabela III
MELHORES RESULTADOS GERAIS COM BASE NO CRITÉRIO $AICc_v$

#	Classe	n_a	n_b	n_c	n_d	n_f	n_k	$AICc_v$	$AICc_i$	J_v	J_i
24	ARX	2	1	–	–	–	1	79,463	288,142	5,801	5,75
44	ARX	3	1	–	–	–	1	80,358	289,971	5,556	5,74
28	ARX	2	2	–	–	–	1	80,596	276,503	5,589	5,276
20	ARX	2	0	–	–	–	1	80,82	290,423	6,384	5,91
40	ARX	3	0	–	–	–	1	81,303	291,572	6,074	5,875
32	ARX	2	3	–	–	–	1	82,077	278,289	5,41	5,264
64	ARX	4	1	–	–	–	1	82,145	288,495	5,419	5,611
160	ARMAX	2	0	1	–	–	1	82,406	292,387	6,244	5,905
240	ARMAX	3	0	1	–	–	1	82,44	291,698	5,853	5,802
1136	BJ	–	2	0	2	1	1	82,909	276,778	5,524	5,214
816	BJ	–	1	0	2	1	1	83,04	273,646	5,941	5,183
36	ARX	2	4	–	–	–	1	83,283	281,484	5,179	5,297
192	ARMAX	2	2	1	–	–	1	83,302	278,685	5,579	5,277
60	ARX	4	0	–	–	–	1	83,403	291,217	5,995	5,784
144	ARMAX	1	4	1	–	–	1	83,435	299,33	5,199	5,922
112	ARMAX	1	2	1	–	–	1	83,574	294,45	6,021	5,902
48	ARX	3	2	–	–	–	1	83,698	278,625	5,634	5,275
336	ARMAX	4	1	1	–	–	1	84,031	282,669	5,277	5,337
832	BJ	–	1	0	3	1	1	84,041	276,12	5,683	5,193
52	ARX	3	3	–	–	–	1	84,774	280,442	5,376	5,263

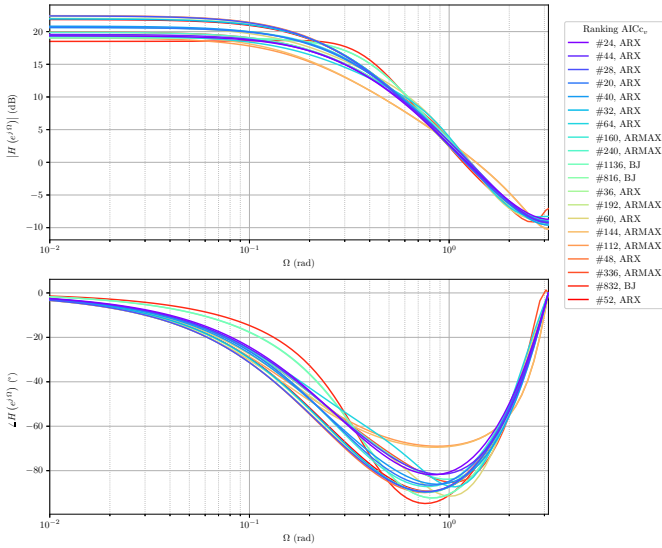


Figura 4. Respostas em frequência de $H(q)$ a partir das identificações na Tabela III.

ainda desconhecido. Os modelos OE são logo descartados devido ao seu desempenho inferior. Ademais, o atraso é arbitrado como $n_k = 1$ pois esse valor prevalece nos modelos bem-classificados.

Conforme já citado, o numerador de $H(q)$ parece ser de ordem pequena. Em particular, com a prevalência de modelos ARMAX com $n_c = 0$ — o que equivale a $C(q) = 1$, sem parâmetros — e de modelos ARX, descarta-se também a classe ARMAX. Restam, portanto, modelos ARX e modelos BJ com $n_c = 0$; em ambos, $n_k = 1$. A diferença entre esses dois casos é que o primeiro requer que os denominadores de $G(q)$ e $H(q)$ sejam iguais, enquanto o segundo permite que sejam independentes.

Finalmente, por apresentar melhores classificações gerais com estrutura mais simples, sugere-se a utilização do modelo ARX. Pelos resultados sugere-se uma ordem n_a entre 2 e 3. A ordem n_b pode ser escolhida entre 0 e 3. Ressalta-se que a arbitração de ordens mais elevadas permite contemplar uma quantidade maior de modelos na classe, ao custo de exigir a identificação de mais parâmetros. Conforme já mencionado, o atraso sugerido é $n_k = 1$.

B. Desempenho do Menor Modelo de Ordem Completa

As funções de transferência que efetivamente geraram os dados são agora reveladas como

$$\begin{aligned} G_0(q) &= \frac{2q^2 + 2q - 1,5}{q^3 - 1,4q^2 + 0,48q} \\ &= \frac{q^{-1}(2 + 2q^{-1} - 1,5q^{-2})}{1 - 1,4q^{-1} + 0,48q^{-2}}, \text{ e} \\ H_0(q) &= \frac{q^3}{q^3 - 1,4q^2 + 0,48q} \\ &= \frac{1}{1 - 1,4q^{-1} + 0,48q^{-2}}. \end{aligned}$$

Observe que os denominadores de $G_0(q)$ e $H_0(q)$ são idênticos. Além disso, no formato de potências negativas, o

numerador de $H_0(q)$ é unitário. Assim, a classe de modelo ARX com $n_a = 2$, $n_b = 2$ e $n_k = 1$ é a menor classe que contém o sistema efetivo.

Note, através das Tabelas II e III, que essa classe corresponde à identificação de índice 28. Além disso, é a terceira melhor classificada tanto entre os modelos ARX como no geral. Seus polinômios são

$$A_{28}(q) = 1 - 1,40683q^{-1} + 0,482612q^{-2}, \text{ e}$$

$$B_{28}(q) = q^{-1}(2,16247 + 1,61084q^{-1} - 1,60164q^{-2}),$$

os quais resultam nas funções de transferência

$$\begin{aligned} G_{28}(q) &= \frac{2,16247q^2 + 1,61084q - 1,60164}{q^3 - 1,40683q^2 + 0,482612q} \\ &= \frac{q^{-1}(2,16247 + 1,61084q^{-1} - 1,60164q^{-2})}{1 - 1,40683q^{-1} + 0,482612q^{-2}}, \text{ e} \end{aligned}$$

$$\begin{aligned} H_{28}(q) &= \frac{q^3}{q^3 - 1,40683q^2 + 0,482612q} \\ &= \frac{1}{1 - 1,40683q^{-1} + 0,482612q^{-2}}. \end{aligned}$$

A Figura 5 mostra a previsão da saída, obtida pela aplicação de (4), a partir do resultado desta identificação, em comparação aos dados de validação. Conforme consta nas Tabelas II e III, o erro quadrático médio associado é 5,589.

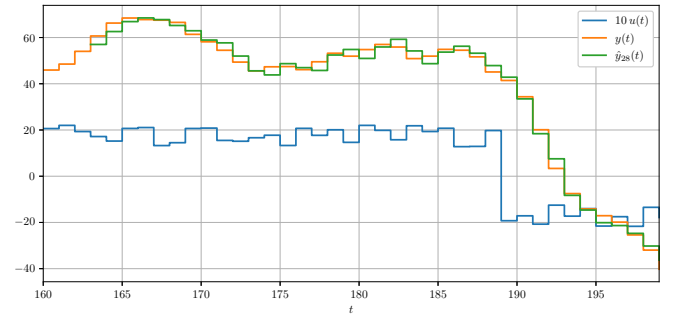


Figura 5. Previsão obtida com o menor modelo de ordem completa.

A comparação com os sistemas efetivos é realizada pelas respostas em frequência das funções de transferência, as quais são exibidas nas Figuras 6 e 7. Observe que a função $G_{28}(q)$ apresenta um ganho estático visivelmente abaixo de $G_0(q)$.

VI. CONCLUSÕES

A determinação da classe do modelo é um desafio ao projetista, ainda mais quando a ordem do sistema a ser identificado é desconhecida. O erro quadrático médio de predição e o critério de informação de Akaike permitem avaliar a qualidade dos modelos obtidos.

Tendo sido realizada uma única identificação por classe, sempre a partir do mesmo conjunto de dados de identificação e validação, os resultados têm um fator considerável de aleatoriedade. A aleatoriedade poderia ser reduzida com a divisão do conjunto de dados em múltiplos subconjuntos de identificação e validação — procedimento este denominado validação cruzada — e posterior comparação dos vários resultados obtidos para cada classe.

REFERÊNCIAS

- [1] L. Ljung, *System identification: theory for the user*, 2nd ed. Upper Saddle River, NJ, Estados Unidos da América: Prentice Hall PTR, 1999.
- [2] L. A. Aguirre, *Introdução à identificação de sistemas: técnicas lineares e não-lineares aplicadas a sistemas reais*, 3rd ed. Belo Horizonte, MG, Brasil: Editora UFMG, 2007.
- [3] The MathWorks, Inc., “What are polynomial models?” Natick, MA, Estados Unidos da América, 2022. [Online]. Disponível em: <https://www.mathworks.com/help/ident/ug/what-are-polynomial-models.html>
- [4] T. Söderström and P. Stoica, *System Identification*. Hemel Hempstead, Reino Unido: Prentice Hall, 1989.

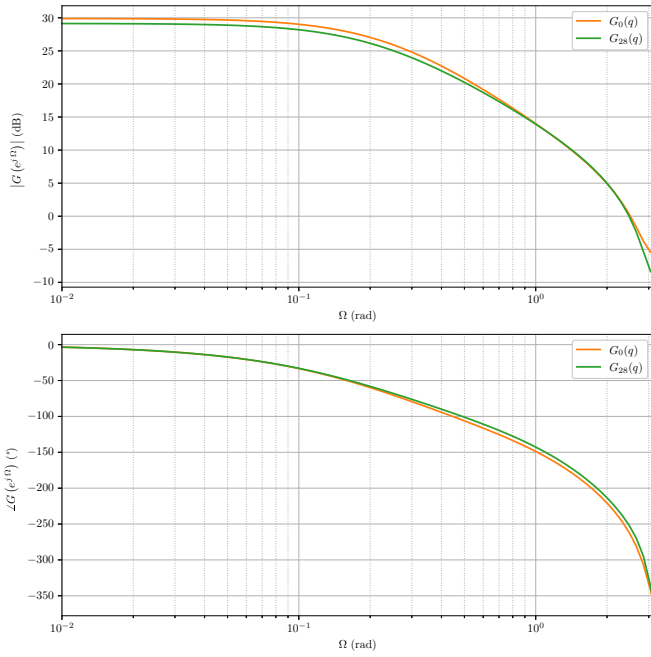


Figura 6. Resposta em frequências de $G(q)$ a partir da identificação com o menor modelo de ordem completa.

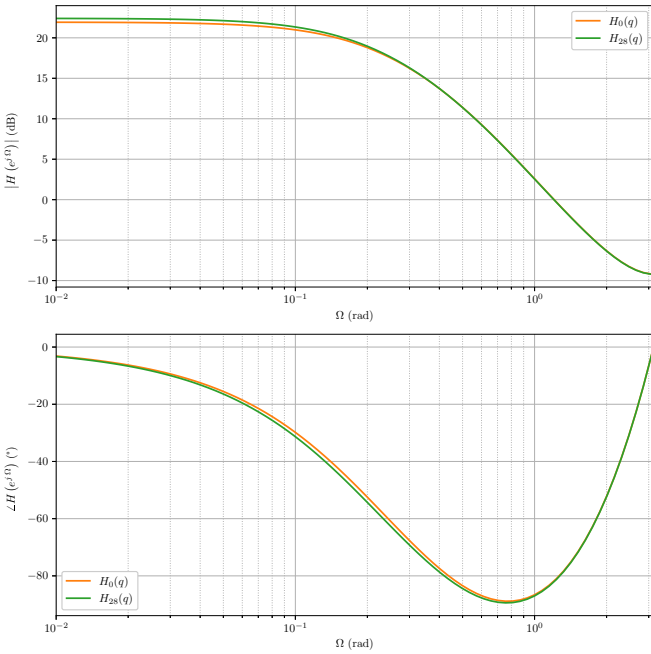


Figura 7. Resposta em frequências de $H(q)$ a partir da identificação com o menor modelo de ordem completa.