DISSERTATION

Research on Order Selection in Model-Free Predictive Control

モデルフリー予測制御の次数選択に関する研究

Graduate School of
Natural Science & Technology
Kanazawa University

Division of Electrical Engineering
and Computer Science

Student ID: 2124042007
Name: Pratvittaya Jiravit
Chief advisor: Prof. Shigeru Yamamoto

Year and Month of Submission: January 2024

Abstract

In this thesis, we introduce a two-stage approach for order selection in model-free predictive control method which is utilizes measured input/output data from the storage data directly computes the control input and does not use any mathematical model as in conventional methods. The capability of model-free predictive control has been already demonstrated in nonlinear systems using linear and polynomial regression for data storage. However, identifying the appropriate order that aligns with the actual system order remains a primary challenge, selecting an incorrect order may result in increasing redundant terms, ultimately leading to instability issues. In this study, we employed the Singular Value Decomposition (SVD) order selection technique, combined with the Bayesian Information Criterion (BIC), to identify the appropriate input and output orders of the system as well as the optimal horizon order in predictive control. This combined technique was subsequently applied to determine the appropriate order for model-free predictive control. Our findings confirmed the effectiveness of the proposed method using numerical simulations in both linear and nonlinear systems and then we extend these findings to compare effective of ℓ_1-minimization approach and Singular Value Decomposition (SVD) with the model-free predictive control. Our numerical simulation results indicate that there is no significant difference between the two methods for linear systems, while for nonlinear systems, the approach utilizing ℓ_1-norm minimization shows superior performance.

Contents

List of Figures

List of Tables

Acknowledgement

First of all, I would like to express my sincere gratitude to my supervisor, Professor Shigeru YAMAMOTO, for his unwavering support, insightful comments, and encouragement throughout my Ph.D. study and research. It has been an honor to be under his guidance, and his enthusiasm, coupled with his vast knowledge, has been a constant source of motivation for my academic and research pursuits. I appreciate all his invaluable contributions of time, ideas, and experiences, which have made my journey as a Ph.D. student both productive and stimulating. The laboratory has provided a wealth of experiences, and I would also like to express my gratitude to Associate Prof. Yuki TODO, Prof. Hidetaka NAMBO, and Ichiro JIKUYA for their support, knowledge, and everything in sharing their expertise.

Second, I extend my gratitude to my parents, Pratvittaya JEERAPA and THANAVIT, receive my deepest appreciation and love for their dedication and years of support throughout my study abroad.

Third, I would like to express my appreciation to all members of MOCCOS laboratory for engaging in numerous insightful and great discussions related to this research. Special thanks to my tutor, Mr.Tsukada Kyosuke, for his assistance in daily life. I am also grateful to my friends from Thailand for their substantial support and continuous encouragement throughout my academic journey. I would not have made it this far without their encouragement and understanding, and I am sincerely grateful to each one of them.

Last but not least, I also extend my gratitude to the Ministry of Education, Culture, Sports, Science and Technology (MEXT), Japan, and the authorities of Kanazawa University for awarding me the scholarship that facilitated my studies and for providing invaluable support throughout my academic pursuits.

Pratvittaya Jiravit
Kanazawa University
January 2024

Chapter 1

Introduction

Model-free predictive control operates within a just-in-time (JIT) modeling framework utilizing short-length data from the stored input/output data of the controlled system to estimate future input sequences. To simplify this model, linear algebraic equations, including the ℓ_1-minimization approach, have been employed, and extended to a polynomial regression model. The primary challenge in this approach lies in selecting the appropriate order. The use of an inappropriate order on an unknown system can deteriorate its performance, making the selection of an appropriate order a challenging task. In this thesis, the challenge is to determine the optimal approach for selecting the appropriate order for an unknown system.

1.1 Previous and Related Researches

Model predictive control (MPC) is one of the most popular control techniques in industrial systems, especially in chemical processes [1, 2]. A prerequisite for this method requires a mathematical model that can represent the dynamics of the controlled system and predict its future behavior. In principle, the accuracy performance of the prediction depends on the precision of the mathematical model. Hence, several system identification methods need to be used to obtain a precise mathematical model [3]. However, the dynamic behavior in most industrial processes is nonlinear, rendering the construction of accurate mathematical models difficult.

To overcome such issues, model-free predictive control was proposed by Stenman in 1999 [4], also known as Just-In-Time (JIT) modeling [5, 6, 7, 8], which constantly updates the mathematical model based on online measurements of input/output data. This method utilizes the past and present recorded input/output data of the query point to estimate a local linear model [6, 7]. JIT modeling is also known as model-on-demand [8, 9], lazy learning [10], or instance-based learning [11]. JIT modeling can be applied to a wide range of industrial applications, including steel industry [12, 13, 14, 15], PID parameter tuning [16, 17], soft sensors in industrial chemical processes [18].

Model-free predictive control in a JIT modeling framework was initially proposed as a data-driven control method that is independent of accurate mathematical models [19, 20, 21]. This method uses short-length data from recorded input/output data of the controlled system to estimate future input sequences using locally weighted averaging (LWA) techniques. This method can be used in several applications, such as treating discretized input systems [22], inverted pendulum systems [23, 24], and parallel mechanisms with pneumatic drives [25]. To simplify this model, linear algebraic equations, such as least-norm squares [26], ℓ_1-minimization approach [27], have been applied, as opposed to model-free-predictive control instead of local linearly weighted average (LWA) method [21]. Moreover, the efficiency of the least-norm squares method, ℓ_1-minimization and LWA of model-free-predictive control was compared in Ref. [28], extending the short-length data to a polynomial regression model [29, 30, 31]. Furthermore, polynomial-regression based on model-free-predictive control has also been applied to multi-input multi-output (MIMO) nonlinear systems [32], controlled systems attached by False Data Injection (FDI) [33], and autonomous underwater vehicles [34]. Although this method can be used in several applications, short-length data becomes longer when the order of polynomial increases, rendering some of the data in the systems unnecessary, while causing the performance of estimating future input to be volatile and out of control.

1.2 Motivations and Objectives

Although the approach based on the Bayesian Information Criterion (BIC)[35] can effectively determine appropriate orders, this method has the drawback of considering all conditions. In contrast, the Singular Value Decomposition (SVD) method can proficiently identify the optimal order of input and output, yet it lacks the capacity to determine an appropriate order for horizon in the predictive control. Therefore, the present study aims to combine the strengths of these two techniques to identify the appropriate order of input and output for the system and evaluate an appropriate order for model-free predictive control without having to evaluate all possible orders. This can be achieved by utilizing the SVD technique to identify the appropriate input and output order, while at the same time employing the BIC-based technique to determine the optimal horizon in the predictive control.

1.3 Outline of The Dissertation

This dissertation is organized as follows. Chapter 2, we describe the model-free predictive control in detail. Chapter 3 explains order selection in model-free predictive control through a two-stage approach which is combines the advantages of Singular Value Decomposition (SVD) and Bayesian Information Criterion (BIC). This chapter also describes the algorithmic steps of the two-stage techniques for model-free predictive control, moreover, the effectiveness of this technique using numerical simulations in both linear and nonlinear system will be discussed. Chapter 4 , we extend our investigation to compare the performance of Singular Value Decomposition (SVD) and ℓ_1-minimization approach in optimizing the input of model-free predictive control. Additionally, some simulations will be discussed. In Chapter 5, we present conclusios summarizing the research findings and a plan for future works.

Chapter 2

Model-Free Predictive Control

Consider the discrete-time system

$$y(t) = f(\boldsymbol{x}(t)) + \epsilon(t), \tag{2.1}$$

where $y \in \mathbb{R}$ is the output of the system,

$$\boldsymbol{x}(t) = \begin{bmatrix} y(t-n) \\ \vdots \\ y(t-1) \\ u(t-m) \\ \vdots \\ u(t-1) \end{bmatrix} \in \mathbb{R}^{n+m} \tag{2.2}$$

is the regression vector consisting of the output and input to the system $u \in \mathbb{R}$, and ϵ is the independent and identically distributed (i.i.d.) noise. We assume that n and m are unknown order in the nonlinear function f, respectively. The control objective is to use h-step future input sequence

$$\boldsymbol{u}_{\mathrm{f}}(t) = \begin{bmatrix} u(t) \\ \vdots \\ u(t+h-1) \end{bmatrix} \in \mathbb{R}^{h} \tag{2.3}$$

so that h-step future outputs

$$\boldsymbol{y}_{\mathrm{f}}(t) = \begin{bmatrix} y(t+1) \\ \vdots \\ y(t+h) \end{bmatrix} \in \mathbb{R}^{h} \tag{2.4}$$

can track the desired reference trajectory

$$\boldsymbol{r}(t) \quad = \begin{bmatrix} r(t+1) \\ \vdots \\ r(t+h) \end{bmatrix} \in \mathbb{R}^{h}. \tag{2.5}$$

Assumption 1 There exists a steady-state for a given $\boldsymbol{r}(t)$ that satisfies

$$\hat{y}(t) = r(t) \tag{2.6}$$

when $\epsilon(t) \equiv 0$. In other words, there exist t^* and $u(t)$ that satisfy this condition when $t > t^*$, as follows

$$r(t) = f(\boldsymbol{x}^*(t)), \quad \boldsymbol{x}^*(t) = \begin{bmatrix} r(t-n) \\ \vdots \\ r(t-1) \\ u(t-m) \\ \vdots \\ u(t-1) \end{bmatrix}. \tag{2.7}$$

To achieve the control objective, the model-free predictive control was introduced in Ref. [19]. This method utilizes vectors $\boldsymbol{a}_i$ and $\boldsymbol{c}_i$ $(i = 1, \ldots, N)$ constructed from stored past input and output $\{u(t), y(t)\}$ of the system (2.1). Furthermore, it employs a query vector $\boldsymbol{b}$ that integrates the most recent input trajectory $\boldsymbol{u}_{\mathrm{p}}(t)$, output trajectory $\boldsymbol{y}_{\mathrm{p}}(t)$, and the reference trajectory $\boldsymbol{r}(t)$. These are defined as follows:

$$\begin{aligned}
\boldsymbol{a}_i &= \begin{cases} \begin{bmatrix} \boldsymbol{y}_{\mathrm{p}}(t_i) \\ \boldsymbol{y}_{\mathrm{f}}(t_i) \\ \boldsymbol{u}_{\mathrm{p}}(t_i) \end{bmatrix} \in \mathbb{R}^{n+m-1+h}, & m \geq 2, \\ \begin{bmatrix} \boldsymbol{y}_{\mathrm{p}}(t_i) \\ \boldsymbol{y}_{\mathrm{f}}(t_i) \end{bmatrix} \in \mathbb{R}^{n+h}, & m = 1, \end{cases} && (2.8) \\
\boldsymbol{b} &= \begin{cases} \begin{bmatrix} \boldsymbol{y}_{\mathrm{p}}(t) \\ \boldsymbol{r}(t) \\ \boldsymbol{u}_{\mathrm{p}}(t) \end{bmatrix} \in \mathbb{R}^{n+m-1+h}, & m \geq 2, \\ \begin{bmatrix} \boldsymbol{y}_{\mathrm{p}}(t) \\ \boldsymbol{r}(t) \end{bmatrix} \in \mathbb{R}^{n+h}, & m = 1, \end{cases} && (2.9) \\
\boldsymbol{c}_i &= \boldsymbol{u}_{\mathrm{f}}(t_i) \in \mathbb{R}^h, && (2.10)
\end{aligned}$$

where $\boldsymbol{u}_{\mathrm{f}}(t)$ is specified in (2.3), and $\boldsymbol{y}_{\mathrm{f}}(t)$ is also given in (2.4). Furthermore,

$$\begin{aligned}
\boldsymbol{u}_{\mathrm{p}}(t) &= \begin{bmatrix} u(t-m+1) \\ \vdots \\ u(t-1) \end{bmatrix} \in \mathbb{R}^{m-1}, \; m \geq 2, && (2.11) \\
\boldsymbol{y}_{\mathrm{p}}(t) &= \begin{bmatrix} y(t-n+1) \\ \vdots \\ y(t) \end{bmatrix} \in \mathbb{R}^{n}. && (2.12)
\end{aligned}$$

In earlier studies on model-free predictive control [19, 20, 21], the input sequence $\hat{\boldsymbol{u}}_{\mathrm{f}}(t)$ that achieves the control objective is determined as a linear weighted average (LWA) of corresponding $\boldsymbol{c}_i$ to selected $\boldsymbol{a}_i$ close to $\boldsymbol{b}$. Yamamoto[26] demonstrated that $\hat{\boldsymbol{u}}_{\mathrm{f}}(t)$ can be expressed as $\hat{\boldsymbol{u}}_{\mathrm{f}}(t) = \boldsymbol{C}\boldsymbol{w}$, using the least-norm solution of $\boldsymbol{A}\boldsymbol{w} = \boldsymbol{b}$ with matrices $\boldsymbol{A}$ and $\boldsymbol{C}$ constructed from $\boldsymbol{a}_i$ and $\boldsymbol{c}_i$ which are close to $\boldsymbol{b}$. Furthermore, by eliminating the need to explicitly choose the nearest neighbors of $\boldsymbol{b}$, Yamamoto[27] proposed the model-free predictive control based on ℓ_1-norm minimization, as outlined in Algorithm 1;

$$\begin{aligned} \hat{u}_{\mathrm{f}}(t) &= \boldsymbol{C}\hat{\boldsymbol{w}} && (2.13) \\ \hat{\boldsymbol{w}} &= \arg\min_{\boldsymbol{w}} \|\boldsymbol{w}\|_1 \text{ subject to } \boldsymbol{A}\boldsymbol{w} = \boldsymbol{b}, && (2.14) \end{aligned}$$

where

$$\begin{aligned} \boldsymbol{A} &= \begin{bmatrix} \boldsymbol{a}_1 & \cdots & \boldsymbol{a}_N \end{bmatrix} \in \mathbb{R}^{(n+m-1+h)\times N}, && (2.15) \\ \boldsymbol{C} &= \begin{bmatrix} \boldsymbol{c}_1 & \cdots & \boldsymbol{c}_N \end{bmatrix} \in \mathbb{R}^{h\times N}, && (2.16) \\ \boldsymbol{w} &= \begin{bmatrix} w_1 \\ \vdots \\ w_N \end{bmatrix} \in \mathbb{R}^{N}. && (2.17) \end{aligned}$$

The ℓ_1-norm is defined as $\|\boldsymbol{w}\|_1 = |w_1| + \cdots + |w_N|$. It is known that ℓ_1-norm minimization frequently results in sparse solutions. In this context, “sparse” denotes a solution vector in which a substantial proportion of its components are either zero or nearly zero. This implies that only a limited number of components possess non-zero values. There have been numerous algorithms proposed for ℓ_1-norm minimization [36], and tools are currently available to solve such problems [37].

To cope with the nonlinearity of the controlled system, a polynomial regression expression is introduced as in [31]. First, we define the pseudo-tensor $\tilde{\otimes}$ as removing duplicated terms in the usual tensor (Kronecker) product $\otimes$. For example,

$$\begin{aligned} \begin{bmatrix} a & b \end{bmatrix}^\top \tilde{\otimes} \begin{bmatrix} a & b \end{bmatrix}^\top &= \begin{bmatrix} a^2 & ab & b^2 \end{bmatrix}^\top && (2.18) \\ \begin{bmatrix} a & b \end{bmatrix}^\top \otimes \begin{bmatrix} a & b \end{bmatrix}^\top &= \begin{bmatrix} a^2 & ab & ab & b^2 \end{bmatrix}^\top && (2.19) \end{aligned}$$

Using the polynomial regression model expression, we can reformulate the model-free predictive control as follows:

$$\begin{aligned} \boldsymbol{a}_i^P &= \boldsymbol{a}_i^{P-1} \tilde{\otimes} \boldsymbol{a}_i \quad , \quad \boldsymbol{a}_1^P = \boldsymbol{a}_1 && (2.20) \\ \boldsymbol{b}^P &= \boldsymbol{b}^{P-1} \tilde{\otimes} \boldsymbol{b}_i \quad , \quad \boldsymbol{b}_1 = \boldsymbol{b} && (2.21) \end{aligned}$$

Algorithm 1 Model-Free predictive control algorithm

Define n, m, h, and N.
Construct $\boldsymbol{A}$ in (2.22) and $\boldsymbol{C}$ in (2.16) from the storage data.
while $t \leq \max(n, m)$ **do**
    Measure output $y(t)$ and apply an appropriate input $u(t)$ in the system and increment the time as $t \leftarrow t + 1$.
end while
repeat
    Measure output $y(t)$ and construct a query vector $\boldsymbol{b}$ in (2.23).
    Estimate $\hat{\boldsymbol{w}}$ by solving ℓ_1-norm minimization (2.14).
    Apply the first element of $\hat{\boldsymbol{u}}_{\mathrm{f}}(t) = \boldsymbol{C}\hat{\boldsymbol{w}}$ to the system input $u(t)$.
    Increment the time as $t \leftarrow t + 1$.
until a terminate condition is met.

where $i = 0, 1, 2, \cdots, N$, $P = 1, 2, \cdots$

$$\boldsymbol{A} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ \boldsymbol{a}_1^1 & \boldsymbol{a}_2^1 & \cdots & \boldsymbol{a}_N^1 \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{a}_1^P & \boldsymbol{a}_2^P & \cdots & \boldsymbol{a}_N^P \end{bmatrix} \in \mathbb{R}^{K \times N}, \tag{2.22}$$

$$\boldsymbol{b} = \begin{bmatrix} 1 \\ \boldsymbol{b}_1 \\ \boldsymbol{b}_2 \\ \vdots \\ \boldsymbol{b}_P \end{bmatrix} \in \mathbb{R}^{K} \tag{2.23}$$

where

$$K = 1 + \sum_{P=1}^{P} \binom{P + n + m + h - 2}{n + m + h - 2} = 1 + \sum_{P=1}^{P} \frac{(P + n + m + h - 2)!}{P!(n + m + h - 2)!} \tag{2.24}$$

is the polynomial regression elements. As the polynomial order increases, $\boldsymbol{K}$ also increases. Moreover, certain data in $\boldsymbol{A}$ and $\boldsymbol{b}$ may not be essential to the system and could potentially lead to divergence. Therefore, It is neccassary to select the relevant data for the polynomial regression elements.

Chapter 3

Order Selection in Model-Free Predictive Control

3.1 Probabilistic Model Selection

Normally, if we increase the model order to fit all the data obtained from a system, this may lead to over-fitting and subsequently to suboptimal system performance. To avoid such issues, there are many model selection criteria, of which the most popular are the Akaike information criterion (AIC) [38], the Bayesian Information Criterion (BIC) [35] and Generalized information criterion (GIC) [42]. All three criteria introduce a penalty for the complexity of the models while rewarding the system efficiency as follows

$$\text{Probabilistic Model Selection} = N \cdot \log(MAE) + K \cdot \eta, \tag{3.1}$$

where N is the number of available data, K is the polynomial regression elements, η is coefficient of penalty term, and MAE is the mean absolute error of the output prediction defined by

$$MAE = \frac{1}{T} \sum_{t=0}^{T} |e(t)|, \tag{3.2}$$

where e is the error between the reference $r(t)$ and the output $y(t)$, and T is the time interval under consideration. The value of MAE is the one that determines system efficiency. The appropriate order is given by the smallest probabilistic model selection criterion.

To provide a clearer understanding, we evaluated the performance of the order selection criteria through three cases of probabilistic model selection

Table 3.1: Probabilistic Model Selection criterion

Criterion	η
AIC	2
BIC	$log(N)$
GIC	$\nu \in [3:6]$

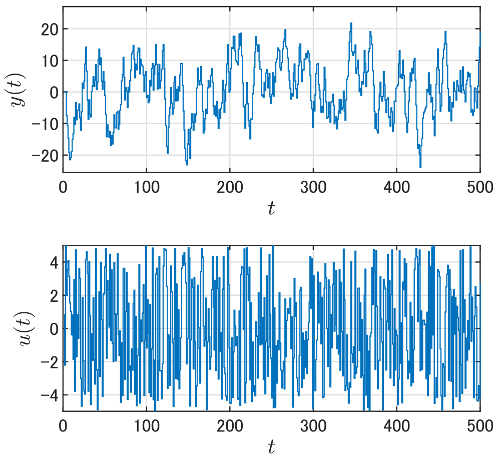


Figure 3.1: Stored data of the linear system (3.3) where the input $u(t)$ is a uniform random sequence: the measured output y (top) and the input u (bottom)

based on (3.1) and Table 3.1 in the model-free predictive control of the linear system, specifying the output order $n = 2$ and the input order $m = 1$,

$$y(t+1) = y(t) - 0.16y(t-1) - 1.5u(t) + \epsilon(t) \tag{3.3}$$

where ϵ is an i.i.d. Gaussian random noise with a zero mean and a variance $\sigma^2 = 0.05$. As shown in Fig. 3.1, we generated input/output data when an input $u(t)$ following a uniform distribution between -5 and 5 was applied to the linear system (3.3). We stored a total of 500 samples of the input/output data specifically for the model-free predictive control. In the model-free predictive control, the square signal

$$r(t) = \begin{cases} 0 & 200i \leq t < 50 + 200i \\ 1 & 50 + 200i \leq t < 100 + 200i \\ 0 & 100 + 200i \leq t < 150 + 200i \\ -1 & 150 + 200i \leq t < 200 + 200i \end{cases} \qquad i = 0, 1, 2, \cdots \tag{3.4}$$

was used as the reference trajectory.

We conducted Algorithm 1 using a grid search method, evaluating all combinations of $n = 1, 2, 3$, $m = 1, 2, 3$, and $h = 1, 2, 3$ with $P = 0$. Table 3.2 illustrates that while using the three cases of probabilistic model selection for determining the correct order, examining all 27 possible order combinations is required. As shown in the table, the optimal order is $(P, n, m, h) = (0, 2, 1, 1)$, which matches with the linear system (3.3). Our results showed that $MAE = 0.0405$, $BIC = -1578.94$, $AIC = -1595.78$, $GIC3 = -1591.78$, $GIC4 = -1587.78$, $GIC5 = -1583.78$, and $GIC6 = -1579.78$ which are graphically depicted in Fig. 3.2. It is clear that the application all of three cases of probabilistic model selection criterion in conjunction with the model-free

Table 3.2: Results of *MAE* and *BIC* when applying model-free predictive control to the linear system (3.3) across 27 combinations of n, m, and h. Bold indicates the optimal value

(n,m,h)	K	MAE	BIC	AIC	$GIC3$	$GIC4$	$GIC5$	$GIC6$
(1, 1, 1)	3	0.0494	−1484.82	−1497.41	−1494.41	−1491.41	−1488.41	−1485.41
(1, 1, 2)	4	0.0470	−1497.23	−1518.21	−1513.21	−1508.21	−1503.21	−1498.21
(1, 1, 3)	5	0.0523	−1431.55	−1461.03	−1454.03	−1447.03	−1440.03	−1433.03
(1, 2, 1)	4	0.0412	−1569.84	−1586.63	−1582.68	−1578.68	−1574.68	−1570.68
(1, 2, 2)	5	0.0411	−1559.01	−1584.28	−1578.28	−1572.28	−1566.28	−1560.28
(1, 2, 3)	6	0.0405	−1553.55	−1587.28	−1579.25	−1571.25	−1563.25	−1555.25
(1, 3, 1)	5	0.0411	−1550.95	−1586.10	−1581.10	−1576.10	−1571.10	−1566.10
(1, 3, 2)	6	0.0412	−1550.95	−1580.60	−1573.60	−1566.60	−1559.60	−1552.60
(1, 3, 3)	7	0.0411	−1539.39	−1577.48	−1568.48	−1559.48	−1550.48	−1541.48
(2, 1, 1)	**4**	**0.0405**	**−1578.94**	**−1595.78**	**−1591.78**	**−1587.78**	**−1583.78**	**−1579.78**
(2, 1, 2)	5	0.0406	−1565.17	−1590.44	−1584.44	−1578.44	−1572.44	−1566.44
(2, 1, 3)	6	0.0405	−1552.92	−1586.61	−1578.61	−1570.61	−1562.61	−1554.61
(2, 2, 1)	5	0.0406	−1570.41	−1591.47	−1586.47	−1581.47	−1576.47	−1571.47
(2, 2, 2)	6	0.0407	−1557.72	−1587.20	−1580.20	−1573.20	−1566.20	−1559.20
(2, 2, 3)	7	0.0405	−1546.87	−1584.78	−1575.78	−1566.78	−1557.78	−1548.78
(2, 3, 1)	6	0.0405	−1566.36	−1591.79	−1585.79	−1579.79	−1573.79	−1567.79
(2, 3, 2)	7	0.0406	−1552.69	−1586.55	−1578.55	−1570.55	−1562.55	−1554.55
(2, 3, 3)	8	0.0406	−1540.33	−1582.63	−1572.63	−1562.63	−1552.63	−1542.63
(3, 1, 1)	5	0.0405	−1572.29	−1593.51	−1588.51	−1583.51	−1578.51	−1573.51
(3, 1, 2)	6	0.0406	−1559.00	−1588.65	−1581.65	−1574.65	−1567.65	−1560.65
(3, 1, 3)	7	0.0405	−1546.75	−1584.83	−1575.83	−1566.83	−1557.83	−1548.83
(3, 2, 1)	6	0.0426	−1540.88	−1566.34	−1560.33	−1554.34	−1548.34	−1542.34
(3, 2, 2)	7	0.0431	−1522.43	−1556.32	−1548.32	−1540.32	−1532.32	−1524.32
(3, 2, 3)	8	0.0435	−1505.03	−1547.35	−1537.35	−1527.35	−1517.35	−1507.35
(3, 3, 1)	7	0.0419	−1542.46	−1572.13	−1565.13	−1558.13	−1551.13	−1544.13
(3, 3, 2)	8	0.0441	−1504.48	−1542.58	−1533.58	−1524.58	−1515.58	−1506.58
(3, 3, 3)	9	0.0437	−1497.06	−1543.59	−1532.59	−1521.59	−1510.59	−1499.59

predictive control is an effective way to determine an appropriate order. Although the three cases of probabilistic model selection criteria can determine the optimal order of the system, the various coefficients of the penalty term are influenced by the number of data N and the level of noise σ^2. To identify the optimal criterion for application with model-free predictive control, we conducted additional tests by exploring all combinations as previously mentioned, considering $N = [250, 500, 750, 1000, 1500, 2000]$ and σ^2 ranging from 0 to 0.1 in increments of 0.01. As depicted in Fig. 3.3, it is evident that with an increase noise, the accuracy of all criteria tends to decrease. Similarly, when utilizing a larger dataset N, the performance of the criteria also decrease correspondingly. Conversely, BIC-criterion distinguishes itself from the others. Even with increasing noise, BIC-criterion consistently identifies the optimal order, as evidenced clearly at $N = 1500$. AIC-criterion failed to identify the optimal order, particularly when $\sigma^2 = 0.01$ and under other criterion. In contrast, BIC-criterion maintains its ability to determine the optimal order. The difference between these criteria is that the AIC is sensitive to highly complex models [39], whereas the BIC has the ability to overcome such sensitivity issues by introducing a simpler structure than the AIC [40]. Nevertheless, as

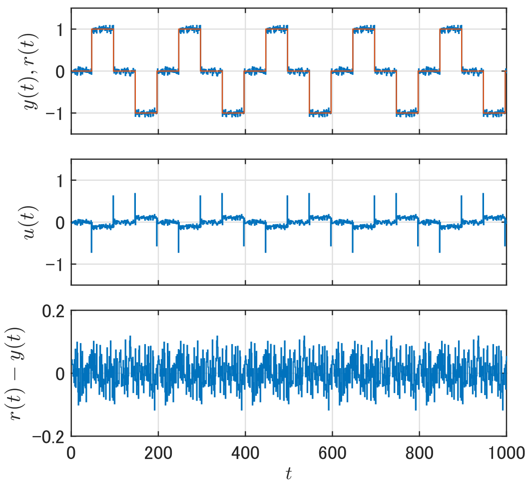


Figure 3.2: Simulation result of model-free predictive control using the appropriate order $(P, n, m, h) = (0, 2, 1, 1)$ for the linear system (3.3). Red represents the desired reference trajectory.

illustrated in Fig. 3.3, employing a larger dataset N leads to a less efficient utilization of the criterion. Based on all the reasons mentioned above, in this thesis, we choose the BIC-criterion, and the number of data N utilized does not exceed 1000 samples.

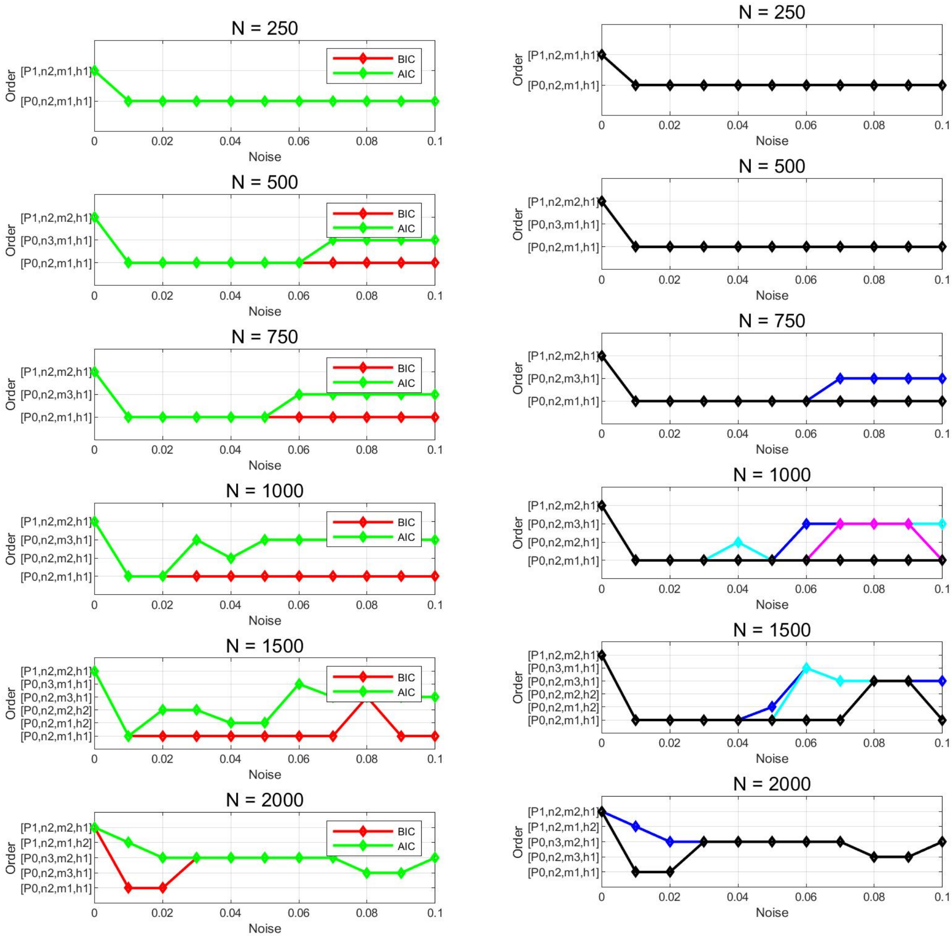
However, if we were to use probabilistic model selection criteria, such as BIC, AIC or GIC, it would be beneficial to determine the optimal model order, to avoid phenomena of overfitting and/or underfitting. A notable drawback of using this technique is that we need to estimate the residual error for each possible combination across the range of orders to identify the lowest value and ascertain the most suitable order. Another widely adopted approach to determine the appropriate order of a linear system is to use the rank of matrices. Among the techniques for evaluating the rank of matrices, the singular value decomposition (SVD) technique stands out. In the next section, we explore the relationship between model-free predictive control and a linear equation, paving the way for the application of the SVD method in this context.

3.2 Singular Value Decomposition Approach

Singular Value Decomposition (SVD) of matrix $\boldsymbol{A}$ can be defined as the mathematical factorization of A by the product of three matrices as depicted below:

$$\boldsymbol{A} = \boldsymbol{U}\boldsymbol{\Sigma}V^{\top} \tag{3.5}$$

where, $A \in \mathbb{R}^{m\times n}$, $\boldsymbol{U} \in \mathbb{R}^{m\times m}$, and $V^{\top} \in \mathbb{R}^{n\times n}$ are set of orthogonal matrices and normalized vectors m-dimensional column vector $\boldsymbol{u}$ and n-dimensional columns vector $\boldsymbol{v}$, respectively. $\Sigma \in \mathbb{R}^{m\times n}$ is diagonal matrix with non-negative singular values (σ_i), arranged in deceasing order $(\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \geq 0)$. The matrix $\boldsymbol{A}$ can be expressed as:



(a) AIC (green) and BIC (red)

(b) GIC3 (navy blue), GIC4 (light blue), GIC5 (pink) , GIC6 (black)

Figure 3.3: Comparative performance of probabilistic model selection: AIC, BIC, and GIC

$$A = \begin{bmatrix} \boldsymbol{u}_1 & \boldsymbol{u}_2 & \cdots & \boldsymbol{u}_m \end{bmatrix} \begin{bmatrix} \sigma_1 & 0 & 0 & \cdots & 0 & \cdots & 0 \\ 0 & \sigma_2 & 0 & \cdots & 0 & \cdots & 0 \\ 0 & 0 & \sigma_3 & \cdots & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 & \cdots & 0 \\ 0 & 0 & 0 & \ddots & \sigma_r & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{V}_1^\top \\ \boldsymbol{V}_2^\top \\ \vdots \\ \boldsymbol{V}_n^\top \end{bmatrix} \tag{3.6}$$

or

$$A = \begin{bmatrix} \boldsymbol{u}_1 & \boldsymbol{u}_2 & \cdots & \boldsymbol{u}_m \end{bmatrix} \begin{bmatrix} \sigma_1 \boldsymbol{V}_1^\top \\ \sigma_2 \boldsymbol{V}_2^\top \\ \vdots \\ \sigma_r \boldsymbol{V}_r^\top \\ \vdots \\ 0 \end{bmatrix} = \sigma_1 \boldsymbol{u}_1 \boldsymbol{V}_1^\top + \sigma_2 \boldsymbol{u}_2 \boldsymbol{V}_2^\top + \cdots + \sigma_r \boldsymbol{u}_r \boldsymbol{V}_r^\top \tag{3.7}$$

One method to evaulate the rank of matrix is through on SVD of the matrix is by considering that the number of non-zero singular values of Σ is r ($rank(A) = rank(\Sigma) = r \leq min(m, n)$). For Ideal case or Noise-free data, the rank of matrix can be inferred from non-zero singular values, as singular values beyond the rank are zero. However, in the presence of noisy data, the singular values may represent non-zero values but tends to be very small. Hence, the utilization of singular values for matrix rank estimation is susceptible to perturbation, particularly in the presence of noise. Consequently, it is important to determine an optimal threshold for selecting the smallest singular value detection. However, the determination of appropriate threshold depends on many infulencing factors, such as the size of dataset and the intensity of the noise which is difficult to determine the appropriate threshold [43].

3.2.1 Singular Value Decomposition appoarch

We will consider the case where the discrete time system (2.1) can be described as a linear system as follow:

$$y(t) = \boldsymbol{\theta}^\top \begin{bmatrix} y(t-n) \\ \vdots \\ y(t-1) \\ u(t-m) \\ \vdots \\ u(t-1) \end{bmatrix}, \tag{3.8}$$

where the exact order n, m, and the model parameter $\boldsymbol{\theta} \in \mathbb{R}^{n+m}$ are assumed to be unknown.

Hereafter, in the context of mathematical expressions, we will emphasize cases where $m \geq 2$. While expressions for $m = 1$ are analogous, they will be omitted for the sake of simplicity.

From (3.8), we obtain

$$\begin{bmatrix} y(t+1) & \cdots & y(t+h) \end{bmatrix} = \boldsymbol{\theta}^\top \begin{bmatrix} y(t-n+1) & \cdots & y(t-n+h) \\ \vdots & \cdots & \vdots \\ y(t) & \cdots & y(t+h-1) \\ u(t-m+1) & \cdots & u(t-m+h) \\ \vdots & \cdots & \vdots \\ u(t) & \cdots & u(t+h-1) \end{bmatrix} \tag{3.9}$$

From these, since $y(t+1), \ldots, y(t+h)$ are linear combination of $y(t-n+1), \ldots, y(t)$ and $u(t-m+1), \ldots, y(t+h-1)$, there exists a matrix $\boldsymbol{\Theta}$ such that

$$\boldsymbol{y}_\mathrm{f}(t) = \begin{bmatrix} y(t+1) \\ \vdots \\ y(t+h) \end{bmatrix} = \boldsymbol{\Theta}^\top \begin{bmatrix} y(t-n+1) \\ \vdots \\ y(t) \\ u(t-m+1) \\ \vdots \\ u(t-1) \\ u(t) \\ \vdots \\ u(t+h-1) \end{bmatrix} = \boldsymbol{\Theta}^\top \begin{bmatrix} \boldsymbol{y}_\mathrm{p}(t) \\ \boldsymbol{u}_\mathrm{p}(t) \\ \boldsymbol{u}_\mathrm{f}(t) \end{bmatrix} \tag{3.10}$$

Here, we note that for $h = 1$, the matrix $\boldsymbol{\Theta}^\top$ becomes a vector that is identical to $\boldsymbol{\theta}^\top$ in (3.9) and (3.10), while for $h \geq 2$, the first $n+m-1$ elements of the first row of $\boldsymbol{\Theta}^\top$ are equivalent to $\boldsymbol{\theta}^\top$ in (3.9) and (3.10). Using the collected data

$$\boldsymbol{Y}_\mathrm{f} = \begin{bmatrix} \boldsymbol{y}_\mathrm{f}(t_1) & \cdots & \boldsymbol{y}_\mathrm{f}(t_N) \end{bmatrix} \in \mathbb{R}^{h \times N} \tag{3.11}$$

$$\boldsymbol{Y}_\mathrm{p} = \begin{bmatrix} \boldsymbol{y}_\mathrm{p}(t_1) & \cdots & \boldsymbol{y}_\mathrm{p}(t_N) \end{bmatrix} \in \mathbb{R}^{n \times N} \tag{3.12}$$

$$\boldsymbol{U}_\mathrm{p} = \begin{bmatrix} \boldsymbol{u}_\mathrm{p}(t_1) & \cdots & \boldsymbol{u}_\mathrm{p}(t_N) \end{bmatrix} \in \mathbb{R}^{(m-1) \times N} \tag{3.13}$$

$$\boldsymbol{U}_\mathrm{f} = \begin{bmatrix} \boldsymbol{u}_\mathrm{f}(t_1) & \cdots & \boldsymbol{u}_\mathrm{f}(t_N) \end{bmatrix} \in \mathbb{R}^{h \times N} \tag{3.14}$$

we can solve

$$\boldsymbol{Y}_\mathrm{f} = \boldsymbol{\Theta}^\top \begin{bmatrix} \boldsymbol{Y}_\mathrm{p} \\ \boldsymbol{U}_\mathrm{p} \\ \boldsymbol{U}_\mathrm{f} \end{bmatrix}. \tag{3.15}$$

Note that from (2.15) and (2.16) in the model-free predictive control, we can obtain

$$\begin{bmatrix} \boldsymbol{A} \\ \boldsymbol{C} \end{bmatrix} = \begin{bmatrix} \boldsymbol{a}_1 & \cdots & \boldsymbol{a}_N \\ \boldsymbol{c}_1 & \cdots & \boldsymbol{c}_N \end{bmatrix} = \begin{bmatrix} \boldsymbol{y}_\mathrm{p}(t_1) & \cdots & \boldsymbol{y}_\mathrm{p}(t_N) \\ \boldsymbol{y}_\mathrm{f}(t_1) & \cdots & \boldsymbol{y}_\mathrm{f}(t_N) \\ \boldsymbol{u}_\mathrm{p}(t_1) & \cdots & \boldsymbol{u}_\mathrm{p}(t_N) \\ \boldsymbol{u}_\mathrm{f}(t_1) & \cdots & \boldsymbol{u}_\mathrm{f}(t_N) \end{bmatrix} = \begin{bmatrix} \boldsymbol{Y}_\mathrm{p} \\ \boldsymbol{Y}_\mathrm{f} \\ \boldsymbol{U}_\mathrm{p} \\ \boldsymbol{U}_\mathrm{f} \end{bmatrix}. \tag{3.16}$$

After rearranging (3.16) in the model-free predictive control, this is equivalent with the linear equation matrices in (3.15). It is widely known that the order of a linear system can be determined by evaluating the rank of a matrix. Singular Value Decomposition (SVD) is a commonly used computational technique for determining matrix ranks. Hence, by considering the linear system relationships in the model-free predictive control, the rank condition can be effectively applied to evaluate the appropriate order of the model-free predictive control. This evaluation can be accomplished solely by using input and output data via matrices $\boldsymbol{A}$ and $\boldsymbol{C}$ as in (3.16).

By partitioning $\boldsymbol{\Theta}$ into three matrices as

$$\boldsymbol{\Theta}^\top = \begin{bmatrix} \boldsymbol{\Theta}_1^\top & \boldsymbol{\Theta}_2^\top & \boldsymbol{\Theta}_3^\top \end{bmatrix} \in \mathbb{R}^{h\times(n+m-1+h)}, \tag{3.17}$$

the equation (3.15) is equivalent to

$$\begin{bmatrix} \boldsymbol{\Theta}_1^\top & -\boldsymbol{I}_h & \boldsymbol{\Theta}_2^\top & \boldsymbol{\Theta}_3^\top \end{bmatrix} \begin{bmatrix} \boldsymbol{Y}_\mathrm{p} \\ \boldsymbol{Y}_\mathrm{f} \\ \boldsymbol{U}_\mathrm{p} \\ \boldsymbol{U}_\mathrm{f} \end{bmatrix} = \boldsymbol{0}, \tag{3.18}$$

where $\boldsymbol{I}_h$ is the identity matrix of size h. In addition, (3.18) is equivalent to

$$\begin{bmatrix} \boldsymbol{\Theta}_1^\top & -\boldsymbol{I}_h & \boldsymbol{\Theta}_2^\top & \boldsymbol{\Theta}_3^\top \end{bmatrix} \begin{bmatrix} \boldsymbol{A} \\ \boldsymbol{C} \end{bmatrix} = \boldsymbol{0}. \tag{3.19}$$

Here, we consider the following SVD

$$\begin{bmatrix} \boldsymbol{A} \\ \boldsymbol{C} \end{bmatrix} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^\top, \tag{3.20}$$

where $\boldsymbol{U} \in \mathbb{R}^{(n+m-1+2h)\times(n+m-1+2h)}$ and $\boldsymbol{V} \in \mathbb{R}^{N\times N}$ are square orthogonal matrices and $\boldsymbol{\Sigma}_{r^*} = \mathrm{diag}(\sigma_1, \ldots, \sigma_{r^*})$ with nonzero singular values $\sigma_1 \geq \cdots \geq \sigma_{r^*} > 0$ and

$$\boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{\Sigma}_{r^*} & \boldsymbol{0}_{r^*\times(N-r^*)} \\ \boldsymbol{0}_{r^0\times r^*} & \boldsymbol{0}_{r^0\times(N-r^*)} \end{bmatrix} \in \mathbb{R}^{(n+m-1+2h)\times N}, \tag{3.21}$$

where r^0 is the number of zero singular values and $r^* + r^0 = n + m - 1 + 2h$.

Since $\boldsymbol{V}$ is the orthogonal matrix, multiplying from the right side of (3.19) with $\boldsymbol{V}$, we can obtain

$$\begin{bmatrix} \boldsymbol{\Theta}_1^\top & -\boldsymbol{I}_h & \boldsymbol{\Theta}_2^\top & \boldsymbol{\Theta}_3^\top \end{bmatrix} \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^\top\boldsymbol{V} = \begin{bmatrix} \boldsymbol{\Theta}_1^\top & -\boldsymbol{I}_h & \boldsymbol{\Theta}_2^\top & \boldsymbol{\Theta}_3^\top \end{bmatrix} \boldsymbol{U}\boldsymbol{\Sigma} = \boldsymbol{0}. \tag{3.22}$$

Note that $\boldsymbol{U}$ is partitioned into two matrices in accordance with $\boldsymbol{\Sigma}$ as

$$\boldsymbol{U} = \begin{bmatrix} \boldsymbol{U}_{r^*} & \boldsymbol{U}_{r^*}^\perp \end{bmatrix}, \tag{3.23}$$

(3.22) is equivalent to

$$\begin{bmatrix} \boldsymbol{\Theta}_1^\top & -\boldsymbol{I}_h & \boldsymbol{\Theta}_2^\top & \boldsymbol{\Theta}_3^\top \end{bmatrix} \boldsymbol{U}_{r^*}\boldsymbol{\Sigma}_{r^*} = \boldsymbol{0}. \tag{3.24}$$

Since $\boldsymbol{\Sigma}_{r^*}$ is nonsingular, (3.24) is equivalent to

$$\begin{bmatrix} \boldsymbol{\Theta}_1^\top & -\boldsymbol{I}_h & \boldsymbol{\Theta}_2^\top & \boldsymbol{\Theta}_3^\top \end{bmatrix} \boldsymbol{U}_{r^*} = \boldsymbol{0}. \tag{3.25}$$

Here, by partitioning $\boldsymbol{U}_{r^*}$ into four matrices according to the matrix in (3.25) as

$$\boldsymbol{U}_{r^*} = \begin{bmatrix} \boldsymbol{U}_1 \\ \boldsymbol{U}_2 \\ \boldsymbol{U}_3 \\ \boldsymbol{U}_4 \end{bmatrix} \tag{3.26}$$

and by defining

$$\boldsymbol{U}_{134} = \begin{bmatrix} \boldsymbol{U}_1 \\ \boldsymbol{U}_3 \\ \boldsymbol{U}_4 \end{bmatrix} \in \mathbb{R}^{(n+m-1+h)\times r^*} \tag{3.27}$$

we obtain a linear equation for $\boldsymbol{\Theta}$ as

$$\boldsymbol{\Theta}^\top \boldsymbol{U}_{134} = \boldsymbol{U}_2. \tag{3.28}$$

When $n + m - 1 + h \leq r^*$, $\boldsymbol{\Theta}$ is given as the least-squares solution

$$\boldsymbol{\Theta}^\top = \boldsymbol{U}_2 \boldsymbol{U}_{134}^\top \left(\boldsymbol{U}_{134} \boldsymbol{U}_{134}^\top\right)^{-1}. \tag{3.29}$$

When $n + m - 1 + h > r^*$, (3.28) becomes an underdetermined system. In this case, the matrix division operation in MATLAB can identify a basic solution, which has at most nonzero component.

When $h = 1$, the matrix $\boldsymbol{\Theta}^\top$ in (3.29) becomes a vector that is identical to $\boldsymbol{\theta}^\top$ in (3.9) and (3.10), while for $h \geq 2$, the first $n + m - 1$ elements of the first row of $\boldsymbol{\Theta}^\top$ are equivalent to $\boldsymbol{\theta}^\top$ in (3.9) and (3.10). While overestimating the degrees of m and n, it is expected that the elements of $\boldsymbol{\theta}$ obtained by (3.29) corresponding to the extra degrees will becomes zero. In addition, the column size r^* of $\boldsymbol{U}_{r^*}$ and the number of zero singular values r^0 play a significant role in evaluating $\boldsymbol{\theta}$, as shown below. In general, the presence of noise signifies that the singular values that should ideally be zero becomes nonzero, making it impossible to select an appropriate r^*. This problem can be circumvented by computing of the ratio of singular values σ_i/σ_{i+1} for $i = 1, 2, \cdots$. Specifically, let r^0 be the smallest integer i where σ_i/σ_{i+1} is maximum. As will be shown in numerical examples later, even when the values of (P, n, m, h) make it difficult to determine whether σ_i can be considered zero, the appropriate r^0 (equivalently r^*) can be determined at the location where σ_i/σ_{i+1} changes abruptly. It is also important that the method described here can be executed solely using storage data.

Application of SVD approach in order selection mentioned above is a linear regression case. As from polynomail case, we extend (2.22) and (3.19) are equivalent to

$$\begin{bmatrix} \boldsymbol{\Theta}_1^\top & -\boldsymbol{I}_h & \boldsymbol{\Theta}_2^\top & \boldsymbol{\Theta}_3^\top \end{bmatrix} \begin{bmatrix} 1 \\ \boldsymbol{A}^1 \\ \vdots \\ \boldsymbol{A}^P \\ \boldsymbol{C} \end{bmatrix} = \boldsymbol{0}. \tag{3.30}$$

From (2.8), (3.30) can be rewritten as

$$\begin{bmatrix} \boldsymbol{\Theta}_1{}^\top & \boldsymbol{\Theta}_2^\top & -\boldsymbol{I}_h & \boldsymbol{\Theta}_3^\top & \boldsymbol{\Theta}_4^\top \end{bmatrix} \begin{bmatrix} 1 \\ \boldsymbol{Y}_\mathrm{p}^1 \\ \boldsymbol{Y}_\mathrm{f}^1 \\ \boldsymbol{U}_\mathrm{p}^1 \\ \boldsymbol{A}^2 \\ \vdots \\ \boldsymbol{A}^P \\ \boldsymbol{C} \end{bmatrix} = \boldsymbol{0}, \tag{3.31}$$

In order to clarify the relationship between $\boldsymbol{\Theta}$ and polynomial dataset of storage data, (3.31) can be rewritten as

$$\boldsymbol{\Theta}_1{}^\top + \boldsymbol{\Theta}_2^\top \boldsymbol{Y}_\mathrm{p}^1 - \boldsymbol{Y}_\mathrm{f}^1 + \boldsymbol{\Theta}_3^\top \begin{bmatrix} \boldsymbol{U}_\mathrm{p}^1 \\ \boldsymbol{A}^2 \\ \vdots \\ \boldsymbol{A}^P \end{bmatrix} + \boldsymbol{\Theta}_4^\top \boldsymbol{C} = \boldsymbol{0}, \tag{3.32}$$

Here, we consider the following SVD as mention before, from (3.31) is equivalent to

$$\begin{bmatrix} \boldsymbol{\Theta}_1{}^\top & \boldsymbol{\Theta}_2^\top & -\boldsymbol{I}_h & \boldsymbol{\Theta}_3^\top & \boldsymbol{\Theta}_4^\top \end{bmatrix} \boldsymbol{U}_{r^*} = \boldsymbol{0}, \tag{3.33}$$

Here, by partitioning $\boldsymbol{U}_{r^*}$ into five matrices according to the matrix in (3.31) as

$$\boldsymbol{U}_{r^*} = \begin{bmatrix} \boldsymbol{U}_1 \\ \boldsymbol{U}_2 \\ \boldsymbol{U}_3 \\ \boldsymbol{U}_4 \\ \boldsymbol{U}_5 \end{bmatrix} \tag{3.34}$$

and by defining

$$\boldsymbol{U}_{1245} = \begin{bmatrix} \boldsymbol{U}_1 \\ \boldsymbol{U}_2 \\ \boldsymbol{U}_4 \\ \boldsymbol{U}_5 \end{bmatrix} \in \mathbb{R}^{(K-h)\times r^*} \tag{3.35}$$

Table 3.3: Estimations of $\boldsymbol{\theta}$ and singular values σ_i for the linear system (3.3), obtained for $P = 0$, and $h = 1$ and various m and n. r^* represents the number of singular values considered nonzero based on the ratio of singular values. Elements of $\boldsymbol{\theta}$ considered to be redundant are indicated in bold, whereas singular values σ_i in bold represent values which are considered to be zero, and bold ratios of singular values reflect values used to determine r^*. In the case of $(P, n, m, h) = (0, 2, 1, 1)$, $\boldsymbol{\theta}$ has no redundant elements and an accurate value is therefore obtained.

n	1	1	1	2	2	2	3	3	3
m	1	2	3	1	2	3	1	2	3
h	1	1	1	1	1	1	1	1	1
K	3	4	5	4	5	6	5	6	7
$\boldsymbol{\theta}$	$\begin{bmatrix} -1.02 \\ 0 \end{bmatrix}$	$\begin{bmatrix} -1.02 \\ -0.28 \\ -1.50 \end{bmatrix}$	$\begin{bmatrix} -1.01 \\ -0.06 \\ -0.30 \\ -1.5 \end{bmatrix}$	$\begin{bmatrix} -0.16 \\ 1 \\ -1.5 \end{bmatrix}$	$\begin{bmatrix} -0.16 \\ 1 \\ \mathbf{0} \\ -1.5 \end{bmatrix}$	$\begin{bmatrix} -0.16 \\ 1 \\ \mathbf{0} \\ \mathbf{0} \\ -1.5 \end{bmatrix}$	$\begin{bmatrix} \mathbf{0} \\ -0.16 \\ 1 \\ -1.5 \end{bmatrix}$	$\begin{bmatrix} \mathbf{0} \\ -0.16 \\ 1 \\ \mathbf{0} \\ -1.5 \end{bmatrix}$	$\begin{bmatrix} \mathbf{0} \\ -0.16 \\ 1 \\ \mathbf{0} \\ \mathbf{0} \\ -1.5 \end{bmatrix}$
σ_i	270.59 97.80 **8.11**	271.39 97.98 56.56 **1.60**	271.76 98.45 65.57 46.11 **0.62**	318.81 121.74 70.78 **0.54**	319.71 127.40 85.60 8.40 **0.54**	319.71 127.56 85.93 56.38 **1.66** **0.54**	354.36 151.68 88.56 57.19 **0.54**	354.78 157.76 90.99 70.92 **0.62** **0.48**	355.83 159.51 96.74 83.42 8.41 **0.62** **0.62**
σ_i/σ_{i+1}	2.77 **12.06**	2.77 1.73 **35.28**	2.76 1.50 1.42 **74.02**	2.62 1.72 **132.17**	2.51 1.49 10.19 **15.69**	2.51 1.48 1.52 **33.87** 3.11	2.34 1.71 1.55 **106.71**	2.25 1.73 1.28 **113.49** 1.31	2.23 1.65 1.16 9.92 **13.46** 1.31
r^*	2	3	4	3	4	4	4	4	5

we obtain a linear equation for $\boldsymbol{\Theta}$ as

$$\boldsymbol{\Theta}^\top \boldsymbol{U}_{1245} = \boldsymbol{U}_3. \tag{3.36}$$

When $K - h \le r^*$, $\boldsymbol{\Theta}$ is given as the least-squares solution

$$\boldsymbol{\Theta}^\top = \boldsymbol{U}_3 \boldsymbol{U}_{1245}^\top \left(\boldsymbol{U}_{1245} \boldsymbol{U}_{1245}^\top\right)^{-1}. \tag{3.37}$$

Consequently, we demonstrated that the optimal order can be obtained using the method presented for the linear system (3.3) discussed earlier with $n = 2$, $m = 1$, and $\boldsymbol{\theta} = \begin{bmatrix} -0.16 & 1 & -1.5 \end{bmatrix}^\top$ using a grid search approach considering all combinations of $P = 0, 1, 2$, $n = 1, 2, 3$, $m = 1, 2, 3$, and $h = 1$. In Table 3.3 presents the estimates of $\boldsymbol{\theta}$ obtained from (3.29). We evaluated the presence of redundant terms based on the consideration of the maximum value observed in σ_i/σ_{i+1}. To clarify this, the maximum value of the ratio for $(P, n, m, h) = (0, 2, 1, 1)$, is $\sigma_3/\sigma_4 = 132.17$. This implies that $\sigma_4 = 0.54$ should be regarded as zero with respect to other singular values. The estimated value of $\boldsymbol{\theta}$ obtained in (3.29) matches that of (3.3), thereby justifying the earlier discussion. As depicted in the polynomial case, Table 3.4 and 3.5 present the estimates of $\boldsymbol{\theta}$ obtained from (3.37) with $P = 1$, and $P = 2$, respectively. Similarly, in the polynomial case, it is evident that the appropriate

Table 3.4: Similar to Table 3.3, obtained for polynomial case of $P = 1$, $(P, n, m, h) = (0, 2, 1, 1)$ remains appropriate order as shown in Table 3.3.

n	1	1	1	2	2	2	3	3	3
m	1	2	3	1	2	3	1	2	3
h	1	1	1	1	1	1	1	1	1
K	4	5	6	5	6	7	6	7	8
$\boldsymbol{\theta}$	$\begin{bmatrix} \mathbf{0} \\ 0.86 \\ -1.53 \end{bmatrix}$	$\begin{bmatrix} \mathbf{0} \\ 0.86 \\ -0.28 \\ -1.50 \end{bmatrix}$	$\begin{bmatrix} \mathbf{0} \\ 0.80 \\ -0.05 \\ -0.30 \\ -1.5 \end{bmatrix}$	$\begin{bmatrix} \mathbf{0} \\ -0.16 \\ 1 \\ -1.5 \end{bmatrix}$	$\begin{bmatrix} \mathbf{0} \\ -0.16 \\ 1 \\ \mathbf{0} \\ -1.5 \end{bmatrix}$	$\begin{bmatrix} \mathbf{0} \\ -0.16 \\ 1 \\ \mathbf{0} \\ \mathbf{0} \\ -1.5 \end{bmatrix}$	$\begin{bmatrix} \mathbf{0} \\ \mathbf{0} \\ -0.16 \\ 1 \\ -1.5 \end{bmatrix}$	$\begin{bmatrix} \mathbf{0} \\ \mathbf{0} \\ -0.16 \\ 1 \\ \mathbf{0} \\ -1.5 \end{bmatrix}$	$\begin{bmatrix} \mathbf{0} \\ \mathbf{0} \\ -0.16 \\ 1 \\ \mathbf{0} \\ \mathbf{0} \\ -1.5 \end{bmatrix}$
σ_i	54.85 22.27 19.52 **1.74**	55 22.27 19.55 11.33 **0.64**	55.06 22.26 19.65 13.11 9.25 **0.58**	64.68 24.34 22.25 14.12 **0.54**	64.86 25.46 22.27 17.08 1.80 **0.54**	64.84 25.5 22.25 17.14 11.30 **0.71** **0.51**	71.95 30.31 22.26 17.71 11.38 **0.54**	72.04 31.52 22.26 18.18 14.14 **0.64** **0.48**	72.24 31.87 22.27 19.33 16.63 1.81 **0.64** **0.48**
σ_i/σ_{i+1}	2.46 1.14 **11.21**	2.47 1.14 1.73 **17.62**	2.47 1.13 1.50 1.42 **16.01**	2.66 1.09 1.58 **25.94**	2.55 1.14 1.30 **9.47** 3.32	2.54 1.15 1.30 1.52 **15.99** 1.37	2.37 1.36 1.26 1.56 **20.89**	2.29 1.42 1.22 1.29 **21.93** 1.35	2.27 1.43 1.15 1.16 **9.21** 2.80 1.36
r^*	3	4	5	4	4	5	5	5	5

Table 3.5: Similar to Table 3.3, obtained for polynomial case of $P = 2$, $(P, n, m, h) = (0, 2, 1, 1)$ remains appropriate order as shown in Table 3.3.

(P,n,m,h) = (2,2,1,1)				(P,n,m,h) = (2,3,1,1)			
$\boldsymbol{\theta}$	σ_i	σ_i/σ_{i+1}	r^*	$\boldsymbol{\theta}$	σ_i	σ_i/σ_{i+1}	r^*
$\begin{bmatrix} \mathbf{0} \\ -0.16 \\ 1 \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ -1.5 \end{bmatrix}$	246.8 75.49 64.61 40.69 28.8 23.64 14.09 13.35 8.06 4.44 **0.54**	3.27 1.17 1.58 1.41 1.22 1.68 1.06 1.66 1.81 **8.17**	10	$\begin{bmatrix} \mathbf{0} \\ \mathbf{0} \\ -0.16 \\ 1 \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ -1.5 \end{bmatrix}$	297.4 103.6 71.8 57.8 48.29 38.57 29.09 18.09 17.48 16.13 11.4 10.27 7.50 5.77 3.60 **0.53**	2.87 1.44 1.24 1.20 1.25 1.33 1.61 1.04 1.08 1.42 1.11 1.37 1.30 1.60 **6.74**	15

order of (3.3) is $(P, n, m, h) = (0, 2, 1, 1)$. Note that in Table 3.5, a grid search was conducted considering all combinations, as before, with a total of 9 orders. However, due to the higher polynomial resulting in a larger amount of data, we are presenting only 2 orders here that are clear from our proposed method, which can evaluate the appropriate order as $(P, n, m, h) = (0, 2, 1, 1)$. To clarify this, in $(P, n, m, h) = (2, 3, 1, 1)$, the estimated value of $\boldsymbol{\theta}$ obtained from (3.37) exhibits numerous zero elements, indicating that the appropriate

Table 3.6: Similar to Table 3.3, even when changing h to $h = 2, 3$, $(P, n, m, h) = (0, 2, 1, 1)$ in Table 3.3 remains appropriate.

n	2	2	2	2	2	2
m	1	1	2	2	3	3
h	2	3	2	3	2	3
K	5	6	6	7	7	8
$\boldsymbol{\theta}$	$\begin{bmatrix} -0.160 \\ 1 \\ -1.50 \\ \mathbf{0} \end{bmatrix}$	$\begin{bmatrix} -0.160 \\ 1 \\ -1.50 \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix}$	$\begin{bmatrix} -0.161 \\ 1 \\ \mathbf{0} \\ -1.50 \\ \mathbf{0} \end{bmatrix}$	$\begin{bmatrix} -0.161 \\ 1 \\ \mathbf{0} \\ -1.50 \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix}$	$\begin{bmatrix} -0.161 \\ 1 \\ \mathbf{0} \\ \mathbf{0} \\ -1.50 \\ \mathbf{0} \end{bmatrix}$	$\begin{bmatrix} -0.161 \\ 1 \\ \mathbf{0} \\ \mathbf{0} \\ -1.50 \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix}$
σ_i	354.78 157.76 90.99 70.91 **0.62** **0.48**	382.29 190.18 108.52 84.89 70.84 **0.66** **0.56** **0.45**	355.83 159.51 96.74 83.42 8.41 **0.62** **0.48**	383.33 190.56 112.93 87.96 82.93 8.40 **0.66** **0.56** **0.45**	355.19 159.95 96.80 83.73 56.38 **1.67** **0.62** **0.48**	382.35 191.11 113.02 87.98 83.27 56.37 **1.67** **0.66** **0.56** **0.45**
σ_i/σ_{i+1}	2.25 1.73 1.28 **113.49** 1.31	2.01 1.75 1.28 1.20 **107.79** 1.17 1.25	2.23 1.65 1.16 9.92 **13.46** 1.31	2.01 1.69 1.28 1.06 9.87 **12.80** 1.17 1.25	2.22 1.65 1.16 1.48 **33.85** 2.67 1.31	2.00 1.69 1.29 1.06 1.48 **33.83** 2.54 1.17 1.26
r^*	4	5	5	6	5	6

order is $(P, n, m, h) = (0, 2, 1, 1)$. Similarly, for all combinations of $n = 1, 2$ and $m = 1, 2, 3$ in Table 3.3, 3.4 and 3.5, respectively, the values of $\boldsymbol{\theta}$ elements at locations corresponding to the extra degrees are determined to be zero, allowing us to find the appropriate values for n and m. This result was obtained by treating small singular values as zero. It should be noted that although the proposed method can determine the appropriate output order n and input order m, it fails to do so for h.

The reason why the appropriate control horizon h cannot be determined lies in the fact that a redundancy is introduced when the horizon is increased. Nonetheless, Table 3.6 shows the results of varying the horizon parameter h from 1 to 3 for combinations of $n = 2$ and $m = 1, 2, 3$. These results demonstrate that, regardless of any changes in the horizon h, similar outcomes can be obtained as in Table 3.3, indicating that increasing the horizon h leads to redundancy. Therefore, this method is not well-suited for effectively identifying the appropriate control horizon h.

Algorithm 2 Two-stage order selection based on SVD and BIC

1st stage
Fix $h = 1$ and N. Select several combinations of (P, n, m).
for all (P, n, m) do
    Construct $\boldsymbol{A}$ in (2.15) and $\boldsymbol{C}$ in (2.16) from storage data.
    Evaluate SVD of $\begin{bmatrix} \boldsymbol{A} \\ \boldsymbol{C} \end{bmatrix} = \boldsymbol{U\Sigma V}^\top$
    Evaluate $\boldsymbol{U}_{r^*}$ from the maximum ratio σ_i/σ_{i+1} of nonzero singular values.
    Determine $\boldsymbol{\theta}$ from $\boldsymbol{\Theta}$ obtained by (3.29).
end for
Determine the best order (P, n, m) via nonzero elements of all $\boldsymbol{\theta}$ obtained above.
2nd stage
Select several h (e.g. $h = 1, 2, 3$).
for all h do
    Perform Algorithm 1
    Calculate $\boldsymbol{MAE}$ (3.2) and $\boldsymbol{BIC}$ (3.1).
end for
Determine the best h as the one that gives the smallest $\boldsymbol{BIC}$.

3.2.2 Two-stage order selection

To determine the appropriate order m, n, as well as the control horizon h in the model-free predictive control, we present a two-stage approach in Algorithm 2. In the first stage of Algorithm 2, suitable values of m and n are determined by SVD as discussed in chapter. 3.2.1, while the second stage determines the best control horizon h by the BIC-criterion as discussed in chapter. 3.1.

All simulations in this paper were performed in MATLAB on a 64-bit Microsoft Windows PC with 3.00GHz Intel Core i7-9700 processor and 8 GB of memory. The execution time of ℓ_1-minimization did not exceed 0.01 s of each sample.

3.2.3 Simulation in Linear system

In chapter. 3.2.1, the first stage of Algorithm 2 was executed for the linear system (3.3). Subsequently, the appropriate order ($n = 2$ and $m = 1$) was discerned. While retaining the order ($n = 2$ and $m = 1$), the model-free predictive control was applied with $h = 1, 2$, and 3, as represented in Table 3.7, marking the second stage of Algorithm 2. The data from Table 3.7 clearly indicates that the BIC identifies $h = 1$ as the optimal horizon. A prominent feature of the proposed two-stage approach is its computational efficiency; following the assessment of six order combinations in Table 3.6, it requires the evaluation of $\boldsymbol{BIC}$ for only three control horizon combinations as shown in Table 3.7. This approach ensures a marked reduction in the number of experiments for model-free predictive control (numerical simulations in the context of this

Table 3.7: Second stage of Algorithm 2. Based on Tables 3.3 and 3.6 and assuming $(P, n, m) = (0, 2, 1)$ is optimal, we obtained MAE and BIC by applying model-free predictive control for $h = 1, 2, 3$. Bold indicates the optimal value.

h	MAE	BIC
1	**0.04045**	**−1578.94**
2	0.04056	−1565.17
3	0.04054	−1552.92

paper), especially when compared to the exhaustive approach of assessing BIC for 27 combinations of orders and horizons without implementing the two-stage approach.

Furthermore, there is a significant reduction in computation time in first stage to determine the appropriate (n, m) using SVD approach requires only 0.1 seconds among all possible combinations and only 30 seconds for second stage to evaluated the appropriate horizon h. Using the two-stage method, the computation time to investigate the 27 combinations shown in Table 3.2, which would normally take 270 seconds, is reduced to just 30.1 seconds, achieving a notable reduction.

3.2.4 Simulation in Nonlinear system 1

We evaluated Algorithm 2 by applying it to a nonlinear system

$$y(t+1) = \frac{y(t)}{1 + y^2(t)} + u^3(t) + \epsilon(t), \tag{3.38}$$

where ϵ is an i.i.d. Gaussian random noise with a zero mean and a variance $\sigma^2 = 0.05$.

Initially, as shown in Fig. 3.4, to obtain input/output data $\mathcal{D}_{open}\{(u, y)\}$, an input $u(t)$ following a uniform distribution between −2 and 2 was applied to the nonlinear system (3.38). The results of the first stage of Algorithm 2 are presented in Table 3.8, where all nine combinations of orders are shown. As depicted in the polynomial case, Table 3.9 and 3.10 present a polynomial order case of $P = 1$, and $P = 2$, respectively. Similarly, it is evident that the appropriate order of (3.38) is $(P, n, m, h) = (0, 1, 1, 1)$. Notably, the order $(P, n, m) = (0, 1, 1)$ is identified as the most suitable.

In the second stage of Algorithm 2 to collect 500 samples of input/output data $\mathcal{D}_{PI}\{(u, y)\}$ around the reference $r(t)$ as specified in (3.4), which are illustrated in Fig. 3.5, we employed PI control

$$u(t) = K_{\mathrm{P}}\left(r(t) - y(t)\right) + K_{\mathrm{I}} \sum_{i=0}^{t} \left(r(i) - y(i)\right) + v(t) \tag{3.39}$$

where $K_{\mathrm{P}} = 0.6$ and $K_{\mathrm{I}} = 0.4$, and $v(t)$ was chosen from a uniform distribution between −0.1 and 0.1 as the excitation signal.

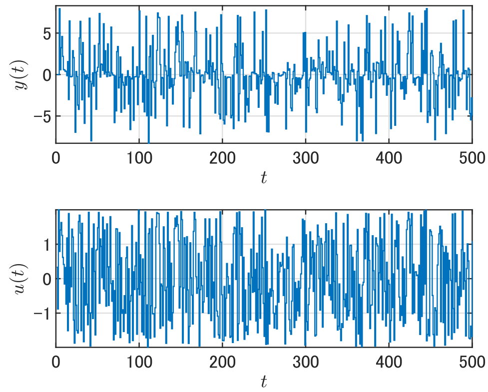


Figure 3.4: Input/output data $\mathcal{D}_{open}\{(u, y)\}$ used in the first stage of Algorithm 2 of the nonlinear system (3.38): measured output y (top) and input u (bottom)

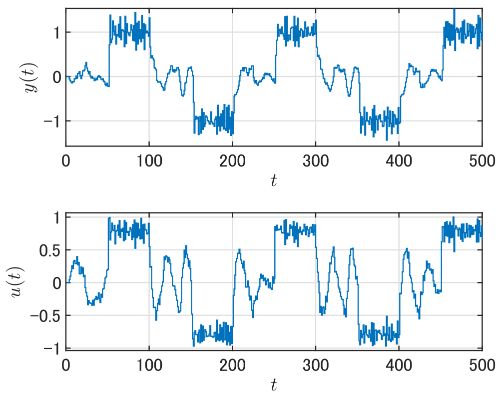


Figure 3.5: Stored data $\mathcal{D}_{PI}\{(u, y)\}$ used in model-free predictive control in second stage of Algorithm 2 of the nonlinear system (3.38) obtained using PI control: measured output y (top) and input u (bottom)

As discussed in Refs. [29, 31], it is evident that PI control can effectively enhance the datasets $\mathcal{D}_{PI}\{(u, y)\}$ used in model-free predictive control because the storage data $\{u(t_j), y(t_j)\}$ $(j = 1, 2, \ldots, N)$ used in model-free predictive control must be sufficiently rich within the range where $y(t)$ can follow the reference signal $r(t)$. While retaining the order ($n = 1$ and $m = 1$), we conducted model-free predictive control to the nonlinear system (3.38) as the second stage of Algorithm 2. As shown in Table 3.11, the optimal horizon with the lowest BIC is $h = 3$. The results of the model-free predictive control using the obtained order and horizon are illustrated in Fig. 3.6.

We note that using $\mathcal{D}_{PI}\{(u, y)\}$ in the first stage does not yields a satisfactory result. Determination of (n, m) by SVD in the first stage presupposes the linearity of the system, whereas $\mathcal{D}_{PI}\{(u, y)\}$ contains input-output data that can follow a wide range of the reference signal r, thereby representing input-output data that can capture the nonlinearity of the system.

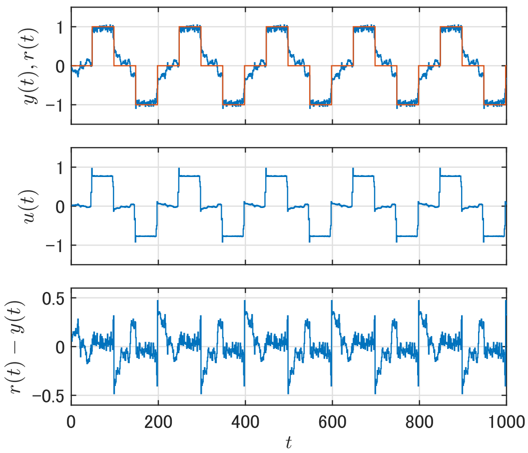


Figure 3.6: Simulation result of model-free predictive control with the appropriate order $(n, m, h) = (1, 1, 3)$ for the nonlinear system (3.38). Red represents the desired reference trajectory.

Table 3.8: Estimations of $\boldsymbol{\theta}$ and singular values σ_i for the nonlinear system (3.38), obtained for $P = 0$, $h = 1$ and various m and n. r^* represents the number of singular values considered nonzero based on the ratio of singular values. Elements of $\boldsymbol{\theta}$ considered to be redundant are indicated in bold. Singular values σ_i in bold are values considered to be zero, and bold ratios of singular values are the values used to determine r^*.

n	1	1	1	2	2	2	3	3	3
m	1	2	3	1	2	3	1	2	3
h	1	1	1	1	1	1	1	1	1
K	3	4	5	4	5	6	5	6	7
$\boldsymbol{\theta}$	$\begin{bmatrix} 0.05 \\ 2.88 \end{bmatrix}$	$\begin{bmatrix} 0.05 \\ \mathbf{0} \\ 2.87 \end{bmatrix}$	$\begin{bmatrix} 0.06 \\ \mathbf{0} \\ \mathbf{0} \\ 2.87 \end{bmatrix}$	$\begin{bmatrix} \mathbf{0} \\ 0.05 \\ 2.87 \end{bmatrix}$	$\begin{bmatrix} \mathbf{0} \\ 0.05 \\ \mathbf{0} \\ 2.87 \end{bmatrix}$	$\begin{bmatrix} \mathbf{0} \\ 0.05 \\ \mathbf{0} \\ \mathbf{0} \\ 2.87 \end{bmatrix}$	$\begin{bmatrix} \mathbf{0} \\ \mathbf{0} \\ 0.05 \\ 2.87 \end{bmatrix}$	$\begin{bmatrix} \mathbf{0} \\ \mathbf{0} \\ 0.05 \\ \mathbf{0} \\ 2.87 \end{bmatrix}$	$\begin{bmatrix} \mathbf{0} \\ \mathbf{0} \\ 0.05 \\ \mathbf{0} \\ \mathbf{0} \\ 2.87 \end{bmatrix}$
σ_i	76.98 69.90 **9.97**	78.44 72.37 **10.89** **9.02**	78.44 72.42 **26.39** **10.85** **8.87**	77.86 72.17 68.09 **9.98**	79.46 72.88 69.86 **10.87** **8.97**	80.28 74.15 71.72 **11.29** **9.92** **8.56**	78.03 74.69 69.00 67.64 **9.98**	79.63 74.73 71.12 67.97 **10.90** **8.93**	80.71 75.54 71.72 69.57 **11.27** **9.91** **8.51**
σ_i/σ_{i+1}	1.10 **7.01**	1.08 **6.65** 1.21	1.08 **2.74** 2.43 1.22	1.08 1.06 **6.83**	1.09 1.04 **6.43** 1.21	1.08 1.03 **6.35** 1.14 1.16	1.05 1.08 1.02 **6.78**	1.07 1.05 1.05 **6.23** 1.22	1.07 1.05 1.03 **6.17** 1.14 1.16
r^*	2	2	2	3	3	3	4	4	4

3.2.5 Simulation in Nonlinear system 2

We also evaluated Algorithm 2 by applying it to a higher nonlinear system [41]

$$\begin{aligned} y(t) &= \frac{z(t)(y(t-3)-1)u(t-2)+u(t-1)}{1+y(t-3)^2+y(t-2)^2} + \varepsilon(t), \\ z(t) &= y(t-1)y(t-2)y(t-3), \end{aligned} \tag{3.40}$$

Table 3.9: Similar to Table 3.8, obtained for polynomial case of $P = 1$, $(P, n, m, h) = (0, 1, 1, 1)$ remains appropriate order as shown in Table 3.8.

n	1	1	1	2	2	2	3	3	3
m	1	2	3	1	2	3	1	2	3
h	1	1	1	1	1	1	1	1	1
K	4	5	6	5	6	7	6	7	8
$\boldsymbol{\theta}$	$\begin{bmatrix}\mathbf{0}\\0.05\\2.88\end{bmatrix}$	$\begin{bmatrix}\mathbf{0}\\0.05\\\mathbf{0}\\2.87\end{bmatrix}$	$\begin{bmatrix}\mathbf{0}\\0.06\\\mathbf{0}\\\mathbf{0}\\2.87\end{bmatrix}$	$\begin{bmatrix}\mathbf{0}\\\mathbf{0}\\0.05\\2.87\end{bmatrix}$	$\begin{bmatrix}\mathbf{0}\\\mathbf{0}\\0.05\\\mathbf{0}\\2.87\end{bmatrix}$	$\begin{bmatrix}\mathbf{0}\\\mathbf{0}\\0.05\\\mathbf{0}\\\mathbf{0}\\2.87\end{bmatrix}$	$\begin{bmatrix}\mathbf{0}\\\mathbf{0}\\\mathbf{0}\\0.05\\2.87\end{bmatrix}$	$\begin{bmatrix}\mathbf{0}\\\mathbf{0}\\\mathbf{0}\\0.05\\\mathbf{0}\\2.87\end{bmatrix}$	$\begin{bmatrix}\mathbf{0}\\\mathbf{0}\\\mathbf{0}\\0.05\\\mathbf{0}\\\mathbf{0}\\2.87\end{bmatrix}$
σ_i	76.99 69.90 **22.36** **9.97**	78.44 72.37 **22.36** **10.89** **9.00**	78.45 72.42 **26.42** **22.34** **10.85** **8.84**	77.86 72.17 68.09 **22.36** **9.96**	79.47 72.88 69.86 **22.36** **10.87** **8.97**	80.28 74.15 71.72 **22.37** **11.29** **9.92** **8.53**	78.03 74.69 69.00 67.64 **22.36** **9.96**	79.63 74.73 71.12 67.97 **22.36** **10.90** **8.90**	80.71 75.54 71.72 69.57 **22.37** **11.27** **9.91** **8.47**
σ_i/σ_{i+1}	1.10 **3.13** 2.24	1.08 **3.24** 2.05 1.21	1.08 **2.74** 1.18 2.06 1.23	1.08 1.06 **3.05** 2.25	1.09 1.04 **3.12** 2.06 1.22	1.08 1.03 **3.21** 1.98 1.14 1.16	1.05 1.08 1.02 **3.03** 2.24	1.07 1.05 1.05 **3.04** 2.05 1.23	1.07 1.05 1.03 **3.11** 1.98 1.14 1.17
r^*	2	2	2	3	3	3	4	4	4

Table 3.10: Similar to Table 3.8, obtained for polynomial case of $P = 2$, $(P, n, m, h) = (0, 1, 1, 1)$ remains appropriate order as shown in Table 3.8.

(P,n,m,h) = (2,1,1,1)				(P,n,m,h) = (2,2,1,1)				(P,n,m,h) = (2,3,1,1)			
$\boldsymbol{\theta}$	σ_i	σ_i/σ_{i+1}	r^*	$\boldsymbol{\theta}$	σ_i	σ_i/σ_{i+1}	r^*	$\boldsymbol{\theta}$	σ_i	σ_i/σ_{i+1}	r^*
$\begin{bmatrix}\mathbf{0}\\0.05\\\mathbf{0}\\\mathbf{0}\\\mathbf{0}\\2.87\end{bmatrix}$	480.1 365.7 217.1 76.61 69.84 **16.43** **9.95**	1.31 1.68 2.83 1.09 **4.25** 1.65	5	$\begin{bmatrix}\mathbf{0}\\\mathbf{0}\\0.05\\\mathbf{0}\\\mathbf{0}\\\mathbf{0}\\\mathbf{0}\\\mathbf{0}\\\mathbf{0}\\2.86\end{bmatrix}$	535.1 369.8 353.3 256.8 207.5 203 76.9 71.5 67.85 **14.88** **9.92**	1.45 1.05 1.37 1.24 1.02 2.64 1.08 1.05 **4.56** 1.5	9	$\begin{bmatrix}\mathbf{0}\\\mathbf{0}\\\mathbf{0}\\0.05\\\mathbf{0}\\\mathbf{0}\\\mathbf{0}\\\mathbf{0}\\\mathbf{0}\\\mathbf{0}\\\mathbf{0}\\\mathbf{0}\\\mathbf{0}\\\mathbf{0}\\2.85\end{bmatrix}$	584.3 375.7 359.2 348.4 269.7 244.9 237.3 205.4 199.7 188.6 76.33 73.3 68.45 66.2 **13.71** **9.85**	1.56 1.05 1.03 1.29 1.10 1.03 1.16 1.03 1.06 2.47 1.04 1.07 1.03 **4.83** 1.39	14

where ϵ is an i.i.d. Gaussian random noise with a zero mean and a variance $\sigma^2 = 0.01$.

As shown in Fig. 3.7, we generated input/output data $\mathcal{D}_{open}\{(u, y)\}$ with an input $u(t)$ following a uniform distribution between -2 and 2 for the higher nonlinear system (3.40). The results of the first stage of Algorithm 2 are outlined in Table 3.12, which presents all nine combinations of orders. Remarkably, the order $(n, m) = (1, 1)$ has been identified as the most suitable.

Table 3.11: Second stage of Algorithm 2. Based on Table 3.8, assuming $(n, m) = (1, 1)$ is optimal, we obtained MAE and BIC by applying model-free predictive control for $h = 1, 2, 3$. Bold indicates the optimal value.

h	MAE	BIC
1	0.1660	−879.27
2	0.1483	−923.16
3	**0.1157**	**−1034.78**

Table 3.12: Estimations of $\boldsymbol{\theta}$ and singular values σ_i for the higher nonlinear system (3.40), obtained for $h = 1$ and various m and n. r^* represents the number of singular values considered nonzero based on the ratio of singular values.

n	1	1	1	2	2	2	3	3	3
m	1	2	3	1	2	3	1	2	3
h	1	1	1	1	1	1	1	1	1
K	3	4	5	4	5	6	5	6	7
$\boldsymbol{\theta}$	$\begin{bmatrix}0.05\\2.88\end{bmatrix}$	$\begin{bmatrix}0.05\\\mathbf{0}\\2.87\end{bmatrix}$	$\begin{bmatrix}0.06\\\mathbf{0}\\\mathbf{0}\\2.87\end{bmatrix}$	$\begin{bmatrix}\mathbf{0}\\0.05\\2.87\end{bmatrix}$	$\begin{bmatrix}\mathbf{0}\\0.05\\\mathbf{0}\\2.87\end{bmatrix}$	$\begin{bmatrix}\mathbf{0}\\0.05\\\mathbf{0}\\\mathbf{0}\\2.87\end{bmatrix}$	$\begin{bmatrix}\mathbf{0}\\\mathbf{0}\\0.05\\2.87\end{bmatrix}$	$\begin{bmatrix}\mathbf{0}\\\mathbf{0}\\0.05\\\mathbf{0}\\2.87\end{bmatrix}$	$\begin{bmatrix}\mathbf{0}\\\mathbf{0}\\0.05\\\mathbf{0}\\\mathbf{0}\\2.87\end{bmatrix}$
σ_i	76.96 69.89 **9.95**	78.41 72.37 **10.87** **8.99**	78.42 72.42 **26.39** **10.83** **8.84**	77.82 72.18 68.07 **9.95**	79.43 72.89 69.85 **10.85** **8.94**	80.24 74.17 71.71 **11.28** **9.89** **8.54**	77.98 74.69 69.00 67.62 **9.95**	79.59 74.73 71.12 67.96 **10.89** **8.90**	80.66 75.55 71.71 69.57 **11.26** **9.87** **8.48**
σ_i/σ_{i+1}	1.10 **7.02**	1.08 **6.66** 1.21	1.08 **2.74** 2.43 1.23	1.08 1.06 **6.84**	1.09 1.04 **6.44** 1.21	1.08 1.03 **6.36** 1.14 1.16	1.04 1.08 1.02 **6.79**	1.06 1.05 1.05 **6.24** 1.22	1.07 1.05 1.03 **6.19** 1.14 1.16
r^*	2	2	2	3	3	3	4	4	4

In the second stage of Algorithm 2, we collect **500** samples of input/output data $\mathcal{D}_{PI}\{(u, y)\}$ around the reference $\boldsymbol{r}(t)$ in (3.4) using PI control as discussed in previous subsections and illustrated in Fig. 3.8. As shown in Table 3.13, the results indicate that the optimal order is $(n, m, h) = (1, 1, 2)$, supporting the conclusions drawn in the previous subsections. We note here that the control results in Fig. 3.9 have deteriorated compared to those in Fig. 3.6, indicating that controlling the higher nonlinear system (3.40) is more difficult that than system (3.38).

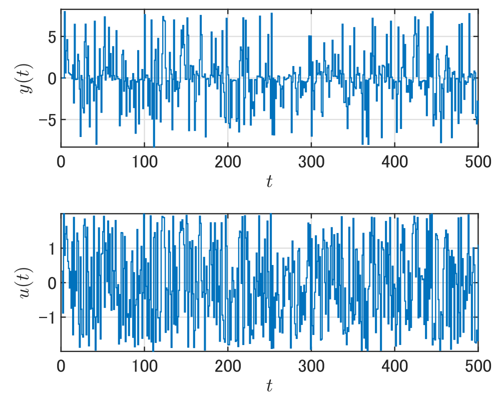


Figure 3.7: Input/output data $\mathcal{D}_{open}\{(u, y)\}$ used in the first stage of Algorithm 2 of the nonlinear system (3.40): measured output y (top) and input u (bottom)

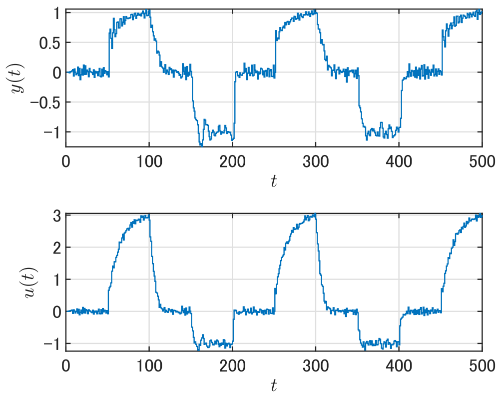


Figure 3.8: Stored data $\mathcal{D}_{PI}\{(u, y)\}$ used in model-free predictive control in second stage of Algorithm 2 of the nonlinear system (3.40) obtained using PI control: measured output y (top) and input u (bottom)

Table 3.13: Second stage of Algorithm 2. Based on Table 3.12, assuming $(n, m) = (1, 1)$ is optimal, we obtained *MAE* and *BIC* by applying model-free predictive control for $h = 1, 2, 3$. Bold indicates the optimal value.

h	*MAE*	*BIC*
1	0.2126	−755.47
2	**0.1563**	**−896.79**
3	0.1572	−881.60

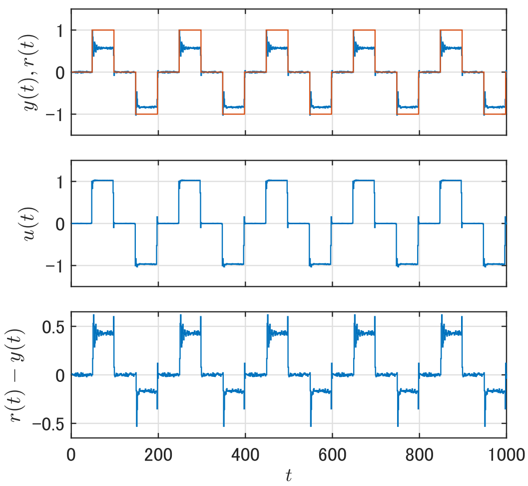


Figure 3.9: Simulation result of model-free predictive control with the appropriate order $(n, m, h) = (1, 1, 2)$ for the higher nonlinear system (3.40). Red represents the desired reference trajectory.

Chapter 4

A Singular Value Decomposition Approach in Model-Free Predictive control

In this section, we discuss methods as alternatives to the ℓ_1-norm minimization approach in the model-free predictive control.

As in [44], by combining (2.13) and (2.14), we obtain

$$\begin{bmatrix} \boldsymbol{A} \\ \boldsymbol{C} \end{bmatrix} \boldsymbol{w} = \begin{bmatrix} \boldsymbol{b} \\ \hat{\boldsymbol{u}}_{\mathrm{f}} \end{bmatrix}. \tag{4.1}$$

Here, we consider the Singular Value Decomposition (SVD)

$$\begin{bmatrix} \boldsymbol{A} \\ \boldsymbol{C} \end{bmatrix} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^{\top}, \tag{4.2}$$

where $\boldsymbol{U} \in \mathbb{R}^{(n+m-1+2h)\times(n+m-1+2h)}$ and $\boldsymbol{V} \in \mathbb{R}^{N\times N}$ are orthogonal matrices, and

$$\boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{\Sigma}_q & \boldsymbol{0}_{q\times(N-q)} \\ \boldsymbol{0}_{z\times q} & \boldsymbol{0}_{z\times(N-q)} \end{bmatrix} \in \mathbb{R}^{(q+z)\times N}, \tag{4.3}$$

is a diagonal matrix, where $\boldsymbol{\Sigma}_q = \mathrm{diag}(\sigma_1, \dots, \sigma_q) \in \mathbb{R}^{q\times q}$ with nonzero singular values $\sigma_1 \geq \cdots \geq \sigma_q > 0$ and where z is the number of zero singular values and

$$q + z = n + m - 1 + 2h. \tag{4.4}$$

Here, $\boldsymbol{U}$ and $\boldsymbol{V}$ are partitioned into matrices in accordance with $\boldsymbol{\Sigma}$ as

$$\boldsymbol{U} = \begin{array}{cc} & \begin{array}{cc} q & z \end{array} \\ \begin{array}{c} n+m-1+h \\ h \end{array} & \begin{pmatrix} \boldsymbol{U}_{11} & \boldsymbol{U}_{12} \\ \boldsymbol{U}_{21} & \boldsymbol{U}_{22} \end{pmatrix} \end{array}, \tag{4.5}$$

$$\boldsymbol{V} = \begin{array}{cc} & \begin{array}{cc} q & N-q \end{array} \\ N & \begin{pmatrix} \boldsymbol{V}_1 & \boldsymbol{V}_2 \end{pmatrix} \end{array}. \tag{4.6}$$

By using SVD (4.2), (4.1) can be rewritten as

$$\boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^{\top}\boldsymbol{w} = \begin{bmatrix} \boldsymbol{b} \\ \hat{\boldsymbol{u}}_{\mathrm{f}} \end{bmatrix}. \tag{4.7}$$

By multiplying $\boldsymbol{U}^\top$ from the left of (4.7), we obtain

$$\begin{bmatrix} \boldsymbol{\Sigma}_q & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{V}_1^\top \\ \boldsymbol{V}_2^\top \end{bmatrix} \boldsymbol{w} = \begin{bmatrix} \boldsymbol{U}_{11}^\top & \boldsymbol{U}_{21}^\top \\ \boldsymbol{U}_{12}^\top & \boldsymbol{U}_{22}^\top \end{bmatrix} \begin{bmatrix} \boldsymbol{b} \\ \hat{\boldsymbol{u}}_\mathrm{f} \end{bmatrix}. \tag{4.8}$$

From this,

$$\boldsymbol{U}_{12}^\top \boldsymbol{b} + \boldsymbol{U}_{22}^\top \hat{\boldsymbol{u}}_\mathrm{f} = \boldsymbol{0}. \tag{4.9}$$

If $h = z$ and $\boldsymbol{U}_{22}$ is nonsingular, the future input $\hat{\boldsymbol{u}}_\mathrm{f}$ can be obtained by

$$\hat{\boldsymbol{u}}_\mathrm{f} = -\boldsymbol{U}_{22}^{-T} \boldsymbol{U}_{12}^\top \boldsymbol{b}. \tag{4.10}$$

The condition $h = z$ is equivalent to

$$q = n + m - 1 + h. \tag{4.11}$$

From (4.7), we also obtain

$$\begin{bmatrix} \boldsymbol{U}_{11} \\ \boldsymbol{U}_{21} \end{bmatrix} \boldsymbol{\Sigma}_q \boldsymbol{V}_1^\top \boldsymbol{w} = \begin{bmatrix} \boldsymbol{b} \\ \hat{\boldsymbol{u}}_\mathrm{f} \end{bmatrix}. \tag{4.12}$$

If (4.11), and $\boldsymbol{U}_{11}$ is nonsingular, we have

$$\hat{\boldsymbol{u}}_\mathrm{f} = \boldsymbol{U}_{21} \boldsymbol{U}_{11}^{-1} \boldsymbol{b}. \tag{4.13}$$

Since $\boldsymbol{U}$ is orthogonal, under the condition (4.11) and that $\boldsymbol{U}_{11}$ and $\boldsymbol{U}_{22}$ are nonsingular, (4.10) and (4.13) are identical as $\boldsymbol{U}_{21} \boldsymbol{U}_{11}^{-1} = -\boldsymbol{U}_{22}^{-T} \boldsymbol{U}_{12}^\top$.

4.1 Examples

In this section, we compare the performance of model-free predictive control using SVD and ℓ_1-norm minimization approaches. All numerical simulations were executed under the following conditions, with certain exceptions:

- The system noise $\boldsymbol{\varepsilon}$ is assumed to be an i.i.d. Gaussian random noise with a zero mean and a variance $\sigma^2 = 0.01$.

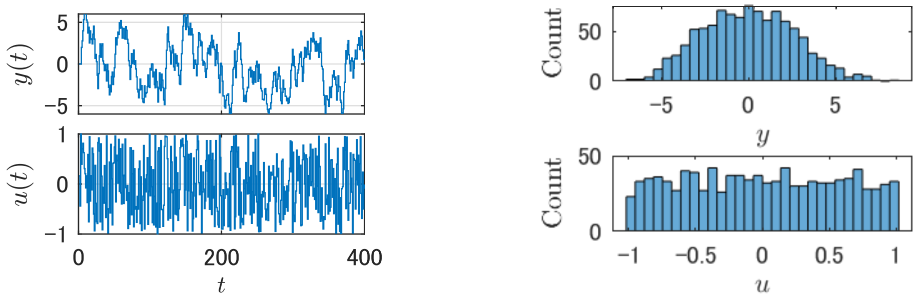
- The input/output dataset is generated by applying $\boldsymbol{u}(t)$ following a uniform distribution $\mathcal{U}(a, b)$ between a and b to the system in an open-loop fashion. A total of $N = 1000$ samples are used for the model-free predictive control.

- Utilizing the dataset and the model order selection method described in Sec. 4, the order of the system was identified for the model-free predictive control.

- In the model-free predictive control, the control horizon is set to $h = 1$, and the sinusoidal signal

  $$r(t) = \sin 0.01\pi t \tag{4.14}$$

  is used as the reference trajectory.



(a) $(u(t)), y(t))$ until $t = 400$ (b) ℓ_1-norm min. based

Figure 4.1: The dataset of the linear system (4.16).

- The control performance is evaluated by the mean absolute error (MAE) defined by

$$MAE = \frac{1}{T_e - T_s} \sum_{t=T_s}^{T_e} |r(t) - y(t)|, \tag{4.15}$$

where $T_s = 0$ and $T_e = 800$.

4.1.1 Simulation in Linear system

We consider the linear system

$$y(t) = \boldsymbol{\theta}^\top \boldsymbol{x}(t) + \varepsilon(t), \quad \boldsymbol{\theta}^\top = \begin{bmatrix} -0.1 & 1 & 2 \end{bmatrix} \tag{4.16}$$

where the order of the system is $(n, m) = (2, 1)$.

The dataset, depicted in Fig. 4.1 by only 400 samples and represented by the histograms of N samples, was generated by a uniform distribution $u(t) \sim \mathcal{U}(-1, 1)$. By using the dataset, the order of the system was identified as $(n, m) = (2, 1)$ by the model order selection in Sec. 3.2.

Figure 4.2 illustrates the simulation results of model-free predictive control by SVD and ℓ_1-norm minimization approaches. The performance using MAE by both approaches are nearly equivalent as shown in Table 4.1.

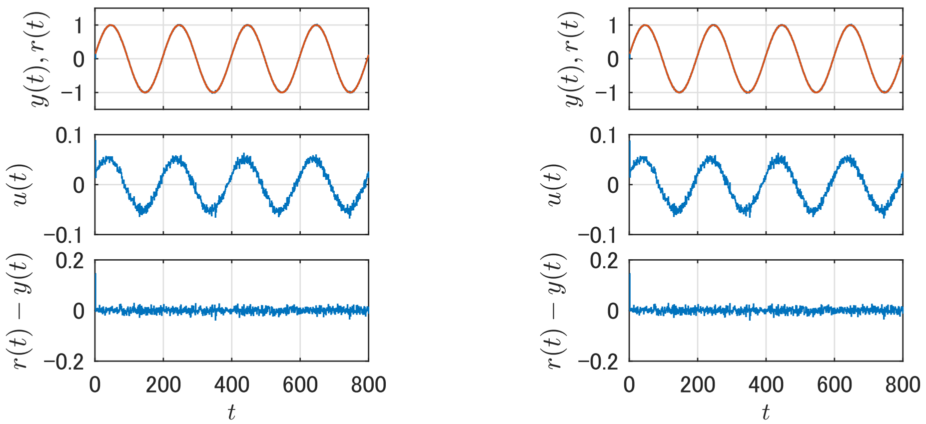
4.1.2 Simulation in Switched linear system

We consider the switched linear system

$$\begin{aligned} y(t) &= \boldsymbol{\theta}_i^\top \boldsymbol{x}(t) + \varepsilon(t), \quad i = 1, 2, \\ \boldsymbol{\theta}_1 &= \begin{bmatrix} -0.1 & 1.2 & 2.2 \end{bmatrix}^\top \text{ when } |y(t-1)| \geq 0.5, \\ \boldsymbol{\theta}_2 &= \begin{bmatrix} 0.1 & 0.8 & 1.9 \end{bmatrix}^\top \text{ when } |y(t-1)| < 0.5, \end{aligned} \tag{4.17}$$

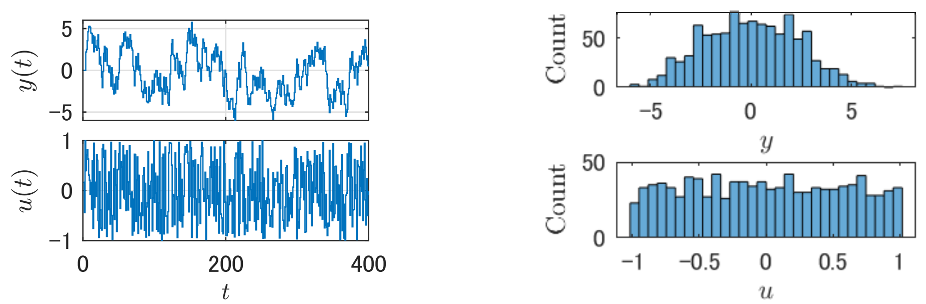
Table 4.1: Comparison of MAE between SVD and ℓ_1-norm minimization approach for the linear system (4.16)

SVD based	ℓ_1-norm min. based
0.0081	0.0081



(a) SVD based (b) histogram of $(u(t), y(t))$

Figure 4.2: Simulation result of model-free predictive control for the linear system (4.16). Red represents the reference trajectory $r(t)$.



(a) $(u(t)), y(t))$ until $t = 400$ (b) ℓ_1-norm min. based

Figure 4.3: The dataset of the switched linear system (4.17).

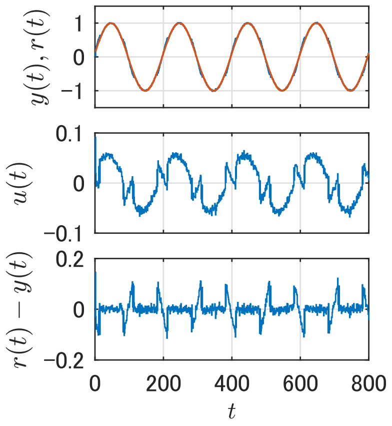
where the order of the system is $(n, m) = (2, 1)$.

The dataset, which is depicted in Fig. 4.3 with only 400 samples and whose distribution is represented by the histograms of N samples, was generated by a uniform distribution $u(t) \sim \mathcal{U}(-1, 1)$. It was found that upon analysis of the dataset that the application of the stricter criteria for zero element determination of the identified $\boldsymbol{\theta}$, as described in Sec. 3.2, did not allow for a unique determination of the system order. However, by relaxing the criteria, the system order was established as $(n, m) = (2, 1)$.

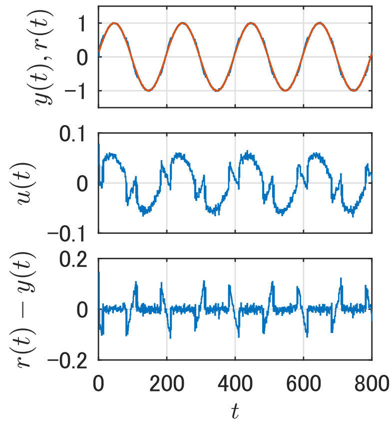
Figure 4.4 illustrates the simulation results of model-free predictive control by two approaches. The performance using MAE by both approaches are nearly equivalent as shown in Table 4.2.

Table 4.2: Comparison of MAE between SVD approach and ℓ_1-norm minimization approach in switched linear system (4.17)

SVD based	ℓ_1-norm min. based
0.0220	0.0216

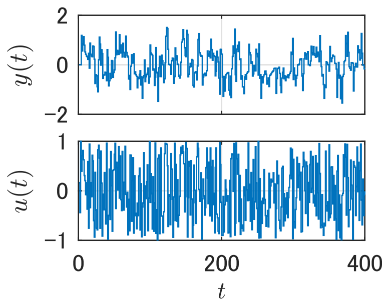


(a) SVD based

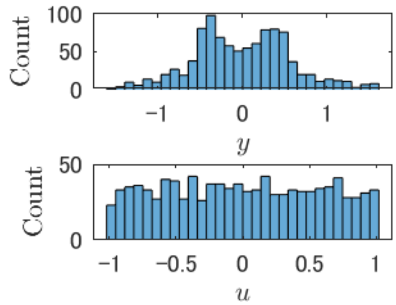


(b) histogram of $(u(t),y(t))$

Figure 4.4: Simulation result of model-free predictive control for the switched linear system (4.17). Red represents $r(t)$.



(a) $(u(t)), y(t))$ until $t = 400$



(b) ℓ_1-norm min. based

Figure 4.5: The first dataset of the nonlinear system (4.18).

4.1.3 Simulation in nonlinear system

We consider the nonlinear system

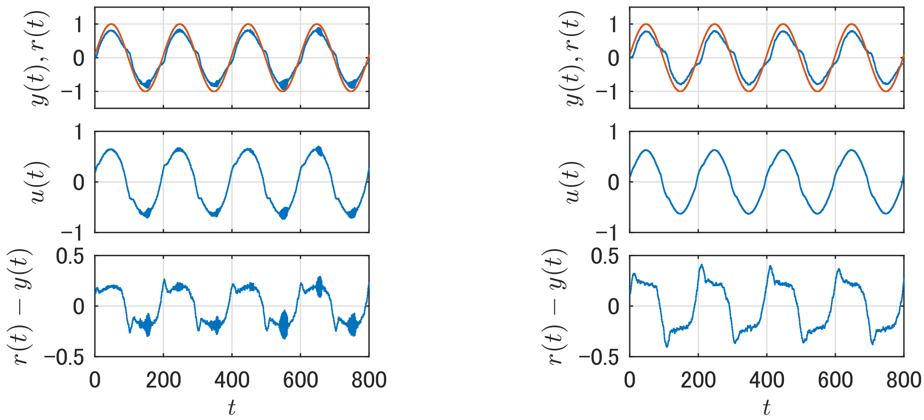
$$y(t) = \frac{y(t-1)}{1 + y^2(t-1)} + 1.2u^3(t-1) + \varepsilon(t), \tag{4.18}$$

where the order of the system is $(n, m) = (1, 1)$.

The dataset, which is depicted in Fig. 4.5 with only 400 samples and whose distribution is represented by the histograms of N samples, was generated by a uniform distribution $u(t) \sim \mathcal{U}(-1, 1)$. Despite altering the criteria for zero element determination when utilizing the dataset, the system order could not be uniquely determined. Therefore, it was set to $(n, m) = (1, 1)$.

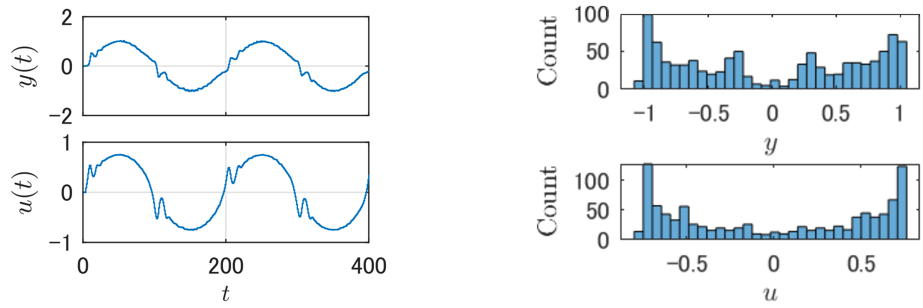
Figure 4.6 illustrates the simulation results of model-free predictive control using two approaches. The control performance has noticeably deteriorated compared to previous results for the linear and switched linear systems, as clearly indicated by MAE in Table 4.3.

As discussed in [30], one way to improve control performance is to collect a dataset around the reference trajectory $r(t)$ for model-free predictive control.



(a) SVD based (b) histogram of $(u(t), y(t))$

Figure 4.6: Simulation result of model-free predictive control using dataset in Fig. 4.5 for the nonlinear system (4.18). Red represents $r(t)$.



(a) $(u(t)), y(t))$ until $t = 400$ (b) ℓ_1-norm min. based

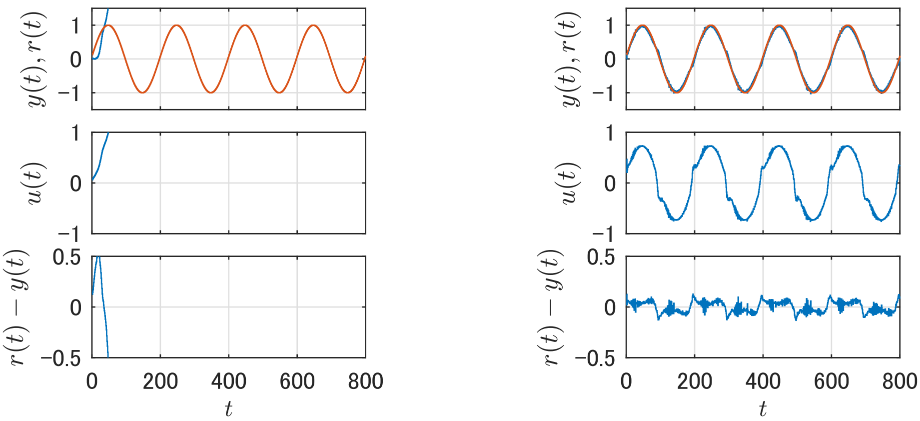
Figure 4.7: The second dataset of the nonlinear system (4.18). The input $u(t)$ is generated by PI control (4.19).

To collect data around the reference trajectory, PI control was employed:

$$u(t) = K_{\mathrm{P}} (r(t) - y(t)) + K_{\mathrm{I}} \sum_{i=0}^{t} (r(i) - y(i)), \tag{4.19}$$

where $K_{\mathrm{P}} = 0.3$ and $K_{\mathrm{I}} = 0.5$. The obtained $u(t)$ by this PI control with the excitation signal $v(t) \sim \mathcal{U}(-0.0004, 0.004)$ was applied to the nonlinear system (4.18) in an open-loop fashion to obtain a dataset. Figure 4.7 shows only 400 samples and the histograms of N samples. In this case also, despite modifying the criteria for zero element determination when utilizing the dataset, the system order could not be uniquely determined. Consequently, it was set to $(n, m) = (1, 1)$.

The simulation results of the model-free predictive control utilizing the dataset shown in Fig. 4.7 are illustrated in Fig. 4.8. The model-free predictive control employing the SVD approach exhibits instability. However, as Table 4.3 clearly demonstrates, MAE by the ℓ_1-norm minimization method is improved. These findings demonstrate the sensitivity of model-free predictive control to variations in the dataset.



(a) SVD based (b) histogram of $(u(t), y(t))$

Figure 4.8: Simulation result of model-free predictive control using dataset in Fig. 4.7 for the nonlinear system (4.18). Red represents $r(t)$.

Table 4.3: Comparison of MAE between SVD approach and ℓ_1-norm minimization approach in nonlinear system (4.18)

	SVD based	ℓ_1-norm min. based
Fig. 4.6	0.1576	0.2132
Fig. 4.8	NaN	0.0494

Chapter 5

Conclusion

In this thesis, we introduced a two-stage technique for determining the appropriate order in model-free predictive control using Singular Value Decomposition (SVD) and Bayesian Information Criterion (BIC). It is important to emphasize that although the traditional order selection using the BIC-criterion can yield the appropriate order, it is typically necessary to consider all possible orders, resulting to the repetition of control experiments for evaluation. However, using the utility of the singular value through SVD can be obtained an appropriate selection only input and output orders, however, it does not provide an optimal value of the horizon in the predictive control. Moreover, this method eliminates the need to simulate every possible order, relying solely on storage data. Consequently, by integrating the strengths of these two methods offers potential in determining the appropriate order. The proposed technique also has the ability to identify orders containing redundant terms that need to be eliminated. Additionally, by examining the decline ratio of singular values, it becomes feasible to detect redundant values even in the presence of noisy signal data. However, when Singular Value Decomposition (SVD) was compared with ℓ_1-norm minimization approach with model-free predictive control. For linear systems, there was no difference in control performance between the two methods, whereas for nonlinear systems, the approach based on ℓ_1-norm minimization was confirmed to be superior through numerical simulations. In the future, we intend to further explore the applicability of our proposed method in the context of nonlinear systems for which a PI-controller and further investigation will not be required and developing methods to acquire high-quality datasets for model-free predictive control, particularly for nonlinear systems.

Publications

Journal paper

1. P. Jiravit and S. Yamamoto: A Two-Stage Method for Order Selection in Model-Free Predictive Control, IEEJ Transactions on Electronics and Electronic Engineering, Vol. 19, No. 6, 2024 DOI: 10.1002/tee.24030.

Conference papers

1. P. Jiravit and S. Yamamoto: A Singular Value Decomposition Approach in Model-Free Predictive Control, SICE International Symposium on Control Systems 2024, March 18th - 20th, 2024, Hiroshima, Japan.

Bibliography

[1] M. Morari and J. H. Lee: Model predictive control: Past, present and future, Computers and Chemical Engineering, Vol. 23, No. 4–5, pp. 667–682, 1999.

[2] C. E. Garcia, D. M. Prett, and M. Morari, Model predictive control: Theory and practice – a survey, Automatica, Vol. 25, No. 3, pp. 335–348, 1989.

[3] L. Ljung: System Identification: Theory for the User, 2nd ed. Upper Saddle River, New Jersey: Prentice Hall, 1999.

[4] A. Stenman: Model-free predictive control, Proc. of the 38th IEEE Conference on Decision and Control, Phoenix, AZ, USA, Vol. 4, pp. 3712–3717, 1999.

[5] G. Cybenko: Just-in-time learning and estimation, NATO ASI series of Computer and Systems Sciences, Vol. 153, pp. 423–434, 1996.

[6] A. Stenman, F. Gustafsson, and L. Ljung: Just in time models for dynamical systems, Proc. of the 35th IEEE Conference on Decision and Control, Vol. 1, pp. 1115–1120, 1996.

[7] A. Stenman, A.V. Nazin, and F. Gustafsson: Asymptotic properties of just-in-time models, Proc. of the 11th IFAC Symposium on System Identification, pp. 1249–1254, 1997.

[8] A. Stenman: Model on Demand: algorithms, analysis and applications, PhD thesis, Department of Electrical Engineering, Linkoping University, 1999.

[9] M.W. Braun, D.E. Rivera, and A. Stenman: A model-on-demand identification methodology for nonlinear process systems, International Journal of Control, Vol. 74, No. 18, pp. 1708–1717, 2001

[10] G. Bontempi, M. Birattari, and H. Bersini: Lazy learning for local modelling and control design, International Journal of Control, Vol. 72, No. 7-8, pp. 643–658, 1999.

[11] D.W. Aha, D. Kibler, and M.K. Albert: Instance-based learning algorithms, Machine Learning, Vol. 6, No. 1, pp. 37–66, 1991.

[12] Q. Zheng and H. Kimura: A new just-in-time modeling method and its application to rolling set-up modeling, Trans. of the Society of Instrument and Control Engineers, Vol. 37, No. 2, pp. 640–646, 2001 (in Japanese).

[13] Q. Zheng and H. Kimura: Just-in-time modeling for function prediction and its application, Asian Journal of Control, Vol. 3, No. 1, pp. 35–44, 2001.

[14] M. Kishi, K. Kimura, J. Ohta, and S. Yamamoto: Shrinkage prediction of a steel production via model-on-demand, Proc. of the 11th IFAC Symposium on Automation in Mining, Mineral and Metal Processing pp. 447–450, 2004.

[15] H. Shigemori, M. Kano, and S. Hasebe: Optimum quality design system for steel products through locally weighted regression model, Journal of Process Control, Vol. 21, No. 2, pp. 293–301, 2011.

[16] J. Ohta and S. Yamamoto: Database-driven tuning of PID controller, Trans. of the Society of Instrument and Control Engineers, Vol. 40, No. 6, pp. 664–669, 2004 (in Japanese).

[17] T. Yamamoto, K. Takao, and T. Yamada: Design of a data-driven PID controller, IEEE Transactions on Control Systems Technology, Vol. 17, No. 1, pp. 29–39, 2009.

[18] K. Fujiwara, M. Kano, S. Hasebe, and A. Takinami: Soft-sensor development using correlation-based just-in-time modeling, AICHE Journal, Vol. 55, No. 7, pp. 1754–1765, 2009.

[19] D. Inoue and S. Yamamoto: An operation support system based on database-driven on-demand predictive control, Proc. of 2004 SICE Annual Conference, pp. 2024–2027, 2004.

[20] D. Inoue and S. Yamamoto: A memory - based predictive control approach to a braking assist problem, Proc. of 2005 SICE Annual Conference, pp. 3186–3189, 2005.

[21] D. Inoue and S. Yamamoto: Support for drivers via just-in-time predictive control and fault detection based on a nearest neighbor method during braking to stop trains, Transactions of the Japan Society of Mechanical Engineers C, Vol. 72, pp. 2756–2761, 2006.

[22] E. Konaka: Design of discrete predictive controller using approximate nearest neighbor method, Proc. the 18th IFAC World Congress, pp. 10213–10218, 2011.

[23] K. Fukuda, S. Ushida, and K. Deguchi: Just-in-time control of image-based inverted pendulum systems with a time-delay, Proc. SICE-ICASE 2006, pp. 4016–4021, 2006.

[24] N. Nakpong and S. Yamamoto: Just-in-time predictive control for a two-wheeled robot, Proc. 2012 Tenth Int. Conf. on ICT and Knowledge Engineering, No. 3, pp. 95–98, 2012.

[25] T. Kosaki and M. Sano: Networked just-in-time control of a parallel mechanism with pneumatic linear drives, Communication in Control Science and Engineering (CCSE), Vol. 2, pp. 19–25, 2014.

[26] S. Yamamoto: A new model-free predictive control method using input and output data, Materials Science and Intelligent Technologies Applications, Advanced Materials Research, Vol. 1042, pp. 182–187, 2014.

[27] S. Yamamoto: A model-free predictive control method by ℓ_1-minimization, Proc. of the 10th Asian Control Conference 2015, 1570064495, 2015.

[28] H. Saputra and S. Yamamoto: Comparative study of model-free predictive control and its database maintenance for unstable systems, SICE Journal of Control, Measurement, and System Integration, Vol. 8, No. 6, pp. 390–395, 2015.

[29] H. Li and S. Yamamoto: A model-free predictive control method based on polynomial regression, Proc SICE International Symposium on Control Systems 2016, 3A2-4, 2016.

[30] H. Li and S. Yamamoto: Polynomial regression based model-free predictive control for nonlinear systems, Proc. SICE 2016, pp. 578–582, 2016.

[31] H. Li and S. Yamamoto: Model-free predictive control using polynomial regressors, SICE Journal of Control, Measurement, and System Integration, Vol. 10, No. 2, pp. 93–99, 2017.

[32] H. Li and S. Yamamoto: Multi-input multi-output model-free predictive control and its application to wastewater treatment, IEEJ Transactions on Electrical and Electronic Engineering, Vol. 12, No. 5, pp. 753–758, 2017.

[33] Z. Zhang, H. Li, H. Zhang, J. Zhang, Z. Zhong, and W. Xu: Model-free predictive control of nonlinear systems under false data injection attacks, Computers and Electrical Engineering, Vol. 100, No. 107977, May, 2022.

[34] H. Li, W. Xu, H. Zhang, J. Zhong, and Y. Liu: Polynomial regressors based data-driven control for autonomous underwater vehicles, Peer-to-Peer Networking and Applications, Vol. 13, pp. 1767–1775, 2020.

[35] G. Schwarz: Estimating the dimension of a model, The Annals of Statistics, Vol. 6, No. 2, pp. 461–464, 1978.

[36] A.Y. Yang, Z. Zhou, A. Ganesh, S.S. Sastry, and Y. Ma: Fast ℓ_1-Minimization Algorithms for Robust Face Recognition IEEE Transactions on Image Processing, Vol. 22, No. 8, pp. 3234–3246, 2013.

[37] M. Grant, S. Boyd, and Y. Ye: CVX: Matlab Software for Disciplined Convex Programming, cvxr.com, 2008.

[38] H. Akaike: A new look at the statistical model identification, IEEE Transactions on Automatic Control, Vol. 19, No. 6, pp. 716–723, 1974.

[39] A. Chakrabarti and J. K. Ghosh: AIC, BIC and recent advances in model selection, Philosophy of Statistics, Vol. 7, pp. 583–605, 2011.

[40] T. Inglot and T. Ledwina: Towards data driven selection of a penalty function for data driven Neyman tests, Linear Algebra and its Applications, Vol. 417, No. 1, pp. 124–133, 2006.

[41] K. S. Narendra and K. Parthasarathy: Identification and Control of Dynamical Systems Using Neural Networks, IEEE Transactions on Neural Networks, Vol. 1, No. 1, pp. 4–27, 1990.

[42] R. J. Bhansali and D. Y. Downham: Some properties of the order of an autoregressive model selected by a generalization of Akaike's FPE criterion, Biometrika, Vol. 64, pp. 547–551, 1977.

[43] M. M. Nouri, N. Miki, and N. Nagai: ARMA model order estimation based on the SVD of the data matrix, Journal of the Acoustical Society of Japann (E), Vol. 15, Issue. 6, pp. 383–392, 1994.

[44] M. Chaovalit and S. Yamamoto: Model-Free Predictive Control and Its Relation to Parameter-Estimation-Based Predictive Control, 2019 SICE International Symposium on Control Systems (SICE ISCS), pp. 84–89 , 2019.