Data-Driven Tracking Control
Insensitive to Noise

Yasumasa Fujisaki ∗ and Masao Ikeda ∗∗

∗ Department of Information and Physical Sciences, Osaka University, Suita, Osaka 565–0871, Japan
(Tel: +81-6-6879-7868; e-mail: fujisaki@ist.osaka-u.ac.jp).
∗∗ Administration Bureau, Osaka University, Suita, Osaka 565–0871, Japan
(Tel: +81-6-6879-4819; e-mail: ikeda@mech.eng.osaka-u.ac.jp).

Abstract: This paper presents a data-driven algorithm for tracking control of a linear discrete-time plant. No traditional mathematical model of the plant such as transfer function or state equation is employed. Instead, the plant dynamics is represented by an array whose elements are plant input-output data. Sensitivity of the array to input and output noise is introduced. Then, it is shown that, for an arbitrary reference signal, the tracking control input which minimizes the sensitivity can readily be computed by solving a linear matrix inequality which is composed of the array.

Keywords: Data-driven control; Input-output data array; Tracking; Noise; Linear discrete-time system.

1. INTRODUCTION

If our available information about a plant is its input-output data, we do not have to introduce a traditional mathematical model such as transfer function, state equation (Kalman, 1960), or kernel representation (Willems, 1991). Any input-output data are constrained in a certain subspace which is determined by the plant dynamics. Thus, using a sufficient number of the input-output data, we must be able to represent the plant dynamics and also to compute the input which achieves the desired output specified by the reference signal.

From this point of view, the authors of the present paper have proposed a system representation and a control strategy based on input-output data in several settings (Ikeda, Fujisaki, and Hayashi, 2001; Fujisaki, Duan, and Ikeda, 2004; Fujisaki, Duan, and Ikeda, 2005; Kai and Fujisaki, 2005). Then, the idea has further been formalized in a behavioral setting (Markovsky and Rapisarda, 2008). In these papers, the underlying idea of a data-driven tracking control algorithm is presented under the condition that the obtained data contain no noise. Note that this algorithm requires neither the mathematical model of the plant nor that of the controller, while other works (Kawamura, 1998; Furuta and Wongsaisiwan, 1995; Chan, 1996; Shi and Skelton, 2000; Park and Ikeda, 2009) concerning data-driven control are interested in the design of controllers which are described by state equation or difference equation.

The advantage of the system representation based on the input-output data appears when the plant dynamics has uncertainties. Obviously, the uncertainties influence the data as noise. Then, if our available information about the plant is its input-output data and we can not know what causes the uncertainties, the best information about the plant is the data themselves. In this case, once we identify the plant dynamics as the traditional mathematical model, the information degrades by applying the data to a restrictive model structure. In contrast, if we represent the plant dynamics using its input-output data themselves, we do not have to reduce any information which the obtained data have. That is, the system representation based on the data must be suitable for controlling the plant under the existence of the uncertainties when our available information about the plant is its input-output data.

In this paper, we extend the data-driven tracking control algorithm (Ikeda et al., 2001) to the case that the obtained data contain noise. To this end, we first summarize the control algorithm proposed in Ikeda et al. (2001) and review it from a new viewpoint via a matrix equation on the data array. Then, we introduce a representation of a data array which contains noise, and derive the most robust control input in the sense that the error of the equation is the most insensitive to the noise. Furthermore, we show that, in order to find such control input, it is enough to solve a linear matrix inequality based on the data array. Finally, we present a numerical example and confirm that the proposed algorithm works well when there exists noise in the obtained data.

2. TRACKING CONTROL BASED ON INPUT-OUTPUT DATA

In this section, we summarize the original version of the data-driven tracking control algorithm (Ikeda et al., 2001). The situation considered there is ideal, that is, it is assumed that the obtained data contain no noise.
In the following, we assume that a plant is of minimum phase and its McMillan degree and relative degree are known. To simplify our discussion, throughout this paper, let us consider a single-input single-output plant of the McMillan degree 3 and the relative degree 2 as an example.

We first arrange its input-output data in an array matrix as

\[
\begin{bmatrix}
y_3 & y_2 & y_1 & y_0 & u_1 & u_0 \\
y_4 & y_3 & y_2 & y_1 & u_2 & u_1 \\
y_5 & y_4 & y_3 & y_2 & u_3 & u_2 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
y_9 & y_8 & y_7 & y_6 & u_7 & u_6 \\
y_{10} & y_9 & y_8 & y_7 & u_8 & u_7 \\
y_{11} & y_{10} & y_9 & y_8 & u_9 & u_8 \\
y_{12} & y_{11} & y_{10} & y_9 & u_{10} & u_9 \\
\end{bmatrix}
\]  

(1)

where \( u_k \) and \( y_k \) are the input and the output at time \( k \).

This array has 6 columns. However, if the obtained data contain no noise, its rank is less than or equal to 5 and never becomes 6. The reason is as follows. In general, the plant dynamics is represented as

\[
y_k + a_1 y_{k-1} + a_2 y_{k-2} + a_3 y_{k+3} = b_1 u_{k-2} + b_2 y_{k-3}
\]  

(2)

with constant coefficients \( a_1, a_2, a_3, b_1, b_2 \). Multiplying the array (1) by the column vector

\[
\theta = \begin{bmatrix} 1 & a_1 & a_2 & a_3 & -b_1 & -b_2 \end{bmatrix}^T
\]  

(3)

from the right-hand side, the result is a zero vector. Thus, we see that the rank of the array does not exceed 5. Furthermore, identification theory (Ljung, 1987) says that its rank equals 5 if the input series \( u_0, u_1, u_2, \ldots \) is persistently exciting.

Using this fact that the rank of the data array (1) is 5, we can compute the control input for tracking an arbitrary reference signal as follows (Ikeda et al., 2001).

For example, we consider that time changes from 9 to 10. Then, the data from \( y_0 \) and \( u_0 \) to \( y_9 \) and \( u_9 \) are known. In the row marked \( \dagger \), \( y_{10} \) is not measured at this time. However, we can predict \( y_{10} \) as this row becomes a linear combination of rows above \( \dagger \) because the rank of the array is 5 and 5 elements on the row are known. Similarly, using this estimated \( y_{10} \), we can predict \( y_{11} \) in the row marked \( \dagger \) \( \dagger \). Next, we come up with the row marked \( \dagger \) \( \dagger \) \( \dagger \) which has two unknown data \( y_{12} \) and \( u_{10} \).

Here, let us suppose that our objective is to control the plant such that the output tracks a given reference signal. Thus, we replace \( y_{12} \) with \( \hat{y}_{12} \). Then, \( u_{10} \) becomes the only unknown data in this row. Therefore, we can determine \( u_{10} \) as this row becomes an linear combination of rows above it if the relative degree is 2, that is, \( b_1 \neq 0 \).

Consequently, repeating the above procedure at each time, we can compute the control input for tracking control.

In this algorithm, the plant dynamics is represented as not transfer function nor state equation but the input-output data array (1). In other words, the array (1) is a new system representation which can introduce a new control algorithm without any traditional mathematical models of the plant.

3. ROBUST TRACKING CONTROL BASED ON INPUT-OUTPUT DATA

3.1 Basic Idea

In the previous section, we have summarized the data-driven tracking control algorithm proposed in Ikeda et al. (2001). In this subsection, we investigate this algorithm from another viewpoint in order to extend to the case that there exist input and output noise.

We first define two matrices \( D_p \) and \( D_f \) as

\[
\begin{bmatrix}
y_3 & y_2 & y_1 & y_0 & u_1 & u_0 \\
y_4 & y_3 & y_2 & y_1 & u_2 & u_1 \\
y_5 & y_4 & y_3 & y_2 & u_3 & u_2 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
y_9 & y_8 & y_7 & y_6 & u_7 & u_6 \\
y_{10} & y_9 & y_8 & y_7 & u_8 & u_7 \\
y_{11} & y_{10} & y_9 & y_8 & u_9 & u_8 \\
y_{12} & y_{11} & y_{10} & y_9 & u_{10} & u_9 \\
\end{bmatrix}
\]  

\[
\begin{bmatrix} D_p \\ D_f \end{bmatrix}
\]  

(4)

Here, \( D_p \) contains data all of which are known at time 9, while \( D_f \) contains three unknown variables \( \hat{y}_{11}, \hat{y}_{10}, \hat{u}_{10} \) and the reference signal \( y_{12} \).

The algorithm summarized in the previous section determines three unknown variables in \( D_f \) according to the fact that the matrices in (4) satisfy the rank condition

\[
\text{rank} D_p = \text{rank} \begin{bmatrix} D_p \\ D_f \end{bmatrix} .
\]  

(5)

Obviously, this condition holds if and only if there exists a real matrix \( X \) satisfying the matrix equation

\[
XD_p = D_f .
\]  

(6)

This equation is consistent with the fact that each row of \( D_f \) can be described as a linear combination of the rows of \( D_p \). Namely, \( D_f \) can be obtained by an elementary row operation to \( D_p \) and \( X \) is its transformation matrix. In other words, the algorithm summarized in the previous section is to select the unknown variables in \( D_f \) such that there exists \( X \) satisfying the equation (6).

Note that the solution \( X \) to (6) is not unique if \( D_p \) has many rows more than 5 as (4). However, three unknown variables in \( D_f \) is determined uniquely and independently of the selection of \( X \) satisfying (6) if the rank condition

\[
\text{rank} D_p = 5
\]  

(7)

holds and the plant is of the McMillan degree 3 and of the relative degree 2, that is, the first and the fifth elements of the coefficient vector \( \theta \) are not zero.

If the input-output data of the plant contain noise, then, in general, \( D_p \) is of full column rank, that is,

\[
\text{rank} D_p = 6 .
\]  

(8)

In this case, the unknown variables in \( D_f \) are not determined independently of the selection of \( X \). However, if we fix \( X \), then the variables in \( D_f \) are determined uniquely following the equation (6).

Now, our objective in this paper is to determine a suitable \( \hat{u}_{10} \) in \( D_f \) in the case that the obtained data contain noise. Therefore, in the following, we regard this objective as to select an appropriate \( X \), and devote ourselves to solving this problem.
### 3.2 Representation of Noisy Input-Output Data

In order to clarify the relations between signals and noise, let us assume that the noise are added at the input and the output of the plant as shown in Fig. 1, where $v_k$ is the input noise and $w_k$ is the output noise.

In this case, the input-output data array (4) can be represented as two arrays. One is the array of pure signals constrained by the plant dynamics, and the other is the array of pure noise, that is,

$$
\begin{bmatrix}
D_p \\
D_f
\end{bmatrix} = \begin{bmatrix}
D_{p0} \\
D_{f0}
\end{bmatrix} + \begin{bmatrix}
N_p \\
N_f
\end{bmatrix}
$$

Here, the signal matrices $D_{p0}$ and $D_{f0}$ are defined as

$$
\begin{bmatrix}
D_{p0} \\
D_{f0}
\end{bmatrix} = \begin{bmatrix}
y_3 & y_2 & y_1 & y_0 & u_1 & u_0 \\
y_4 & y_3 & y_2 & y_1 & u_2 \\
y_5 & y_4 & y_3 & y_2 & u_3 & u_2 \\
& \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
y_9 & y_8 & y_7 & y_6 & u_6 \\
y_{10} & y_{10} & y_9 & y_8 & y_7 & u_7 & u_6 \\
y_{11} & y_{11} & y_{10} & y_9 & y_8 & y_7 & u_8 \\
y_{12} & y_{11} & y_{10} & y_9 & y_8 & y_7 & u_8
\end{bmatrix}
$$

and the noise matrices $N_p$ and $N_f$ are defined as

$$
\begin{bmatrix}
N_p \\
N_f
\end{bmatrix} = \begin{bmatrix}
w_3 & w_2 & w_1 & w_0 & v_1 & v_0 \\
w_4 & w_3 & w_2 & w_1 & v_2 \\
w_5 & w_4 & w_3 & w_2 & v_3 & v_2 \\
& \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
w_9 & w_8 & w_7 & w_6 & v_6 \\
w_{10} & w_{10} & w_9 & w_8 & v_7 & v_6 \\
w_{11} & w_{11} & w_{10} & w_9 & v_8 \\
w_{12} & w_{11} & w_{10} & w_9 & v_8
\end{bmatrix}
$$

In the above representation, we introduce the weights $V$, $W_p$, and $W_f$ on the noise matrices to normalize the size of the noise, where these weights are supposed to be nonsingular matrices. The normalized noise matrices $\Delta_p$ and $\Delta_f$ are assumed to

$$
\begin{bmatrix}
\Delta_p \\
\Delta_f
\end{bmatrix} = \begin{bmatrix}
\Delta^T_p & \Delta_f
\end{bmatrix} \leq \delta I
$$

where $\delta$ is a positive number and represents the size of the normalized noise.

Note that we here consider an additive noise to the input-output data array, and discuss robustness against not the perturbation of the coefficients in (2) but the noise in the data array. Note also that the matrix $D_f$ contains the unknown variables as the definition (4) and they appear only in $D_{f0}$.

Here, let us consider how to select the weights $V$, $W_p$, and $W_f$. For simplicity, we assume that all of the weights are diagonal matrices.

Each element of $V$ is the weight on each column of the noise matrices $\Delta_p$ and $\Delta_f$. Noting that the input-output data array has two kinds of the columns corresponding to $y_k$ and $u_k$, the weight $V$ can be used to normalize the noise contained in $y_k$ and $u_k$. For example, we may select each element of $V$ as each variance of the noise contained in $y_k$ and $u_k$.

Each element of $W_p$ and $W_f$ is the weight on each row of the noise matrices $\Delta_p$ and $\Delta_f$ normalized as (12). Noting that each row of $D_p$ contains 6 real data but each row of $D_f$ contains real data equal to or less than 5, the weight $W_f$ can be used to compensate the difference of the number of the noisy data in each row of $D_f$. For example, we may select the weight $W_f$ as the diagonal matrix

$$
\begin{bmatrix}
5/6 & 0 & 0 \\
0 & 4/6 & 0 \\
0 & 0 & 2/6
\end{bmatrix}
$$

The weight $W_p$ is enough to be identity matrix if the size of the noise in the data does not change as passing time. However, we can use the weight to cope with perturbations of the system dynamics if the data matrix $D_p$ has sufficient number of rows. For example, we may set the upper elements of $W_p$ corresponding to the past data to have large values. This means that the past data are supposed to contain large noise, and the weight works as a kind of forgetting factor which is usually introduced in identification theory (Ljung, 1987).

### 3.3 Robust Control Input

Now, if the input-output data array (9) contains no noise, the matrix equation (6) becomes

$$
X_0 D_{p0} = D_{f0}
$$

where $X_0$ is a solution in this ideal situation. In case we can know the signal matrices $D_{f0}$ and $D_{p0}$, we can solve the above matrix equation on the variables $X_0$ and $\hat{y}_{11}$, $\hat{u}_{10}$ in $D_{f0}$. The result $\hat{u}_{10}$ is the most suitable control input because it is consistent with the pure signals contained in the input-output data array.

However, in real situation, we can never know the signal matrices $D_{p0}$ and $D_{f0}$. Our available input-output data array contains noise, thus we can not solve the matrix equation (14) but (6). We can not compute $X_0$ but $X$. The equation (14) with $X$ must have an error, and it must be represented as

$$
XD_{p0} = D_{f0} + W_f E V
$$

where $E$ denotes the normalized error caused by the noise in the data. Here, we introduce the weights $V$ and $W_f$ on the error $E$ because the error of the array has the same structure as that of $D_f$ whose noise matrix is weighted with $V$ and $W_f$ in (9).
Substituting (9) into (6), we see that the error $E$ is represented as

$$E = -W_f^{-1}XW_p\Delta_p + \Delta f$$

$$= \left[ -W_f^{-1}XW_p I \right] \left[ \Delta_p \Delta_f \right].$$  \hfill \text{(16)}

Hence, the size of the error $E$ depends on the selection of $X$. Here, let us quantize the size of the error introducing a positive number $\varepsilon$ as

$$EE^T \leq \varepsilon I.$$ \hfill \text{(17)}

By using the condition (12), this inequality becomes

$$W_f^{-1}XW_p(W_f^{-1}XW_p)^T + I \leq \frac{\varepsilon}{\delta} I.$$ \hfill \text{(18)}

Based on the consideration so far, in this paper, we propose the following control strategy. Let us consider the matrix inequality

$$W_f^{-1}XW_p(W_f^{-1}XW_p)^T + I \leq \gamma I$$ \hfill \text{(19)}

where $\gamma$ is a positive number. Then, the control strategy proposed in this paper is to select the variables $X$ and $\hat{y}_{11}$, $\hat{y}_{10}$, $\hat{u}_{10}$ in $D_f$ to minimize $\gamma$ subject to the inequality (19) and the equation (6).

Note that $\gamma$ is sensitivity of the error $E$ to the noise matrices $\Delta_p$ and $\Delta_f$. Namely, this control strategy produces the most robust input in the sense that the error in (15) is the most insensitive to the noise contained in the input-output data array (4) because we minimize $\gamma$ and determine the unknown variables in $D_f$.

While the weights $W_p$ and $W_f$ appear naturally in the above inequality (19), the weight $V$ does not appear anywhere. This is because when we weight $\Delta_p$ and $\Delta_f$ with the same $V$ form the right-hand side, $V$ is canceled in the error $E$ composed of summation of these noise. Although $V$ does not affect the sensitivity $\gamma$ in this paper, introducing $V$ may be useful if we focus on another control performance instead of $\gamma$.

### 3.4 Calculation of Control Input via LMI

Let us recast the above sensitivity minimization problem as an LMI optimization problem.

The matrix inequality (19) is equivalent to

$$(W_f^{-1}XW_p)^T(W_f^{-1}XW_p) \leq (\gamma - 1)I$$ \hfill \text{(20)}

and the equation (6) can be rewritten as

$$XW_pW_p^{-1}D_p = D_f.$$ \hfill \text{(21)}

Therefore, if there exists $X$ satisfying the condition (19) and (6), that is, (20) and (21), then the matrix inequality

$$(W_f^{-1}D_f)^TW_f^{-1}D_f \leq (\gamma - 1)(W_p^{-1}D_p)^TW_p^{-1}D_p$$ \hfill \text{(22)}

holds.

Conversely, if the matrix inequality (22) holds, then the rank condition (5) holds. This ensures the existence of a solution $X$ to the equation (21). Here, we select a solution $X$ as

$$\hat{X} = D_f(W_p^{-1}D_p)^+W_p^{-1}$$ \hfill \text{(23)}

where $\bullet^+$ denotes Moore-Penrose pseudo inverse. This equation is rewritten as

$$W_f^{-1}\hat{X}W_p = W_f^{-1}D_f(W_p^{-1}D_p)^+.$$ \hfill \text{(24)}

Using the above equation, the matrix inequality (22), and well-known properties of the pseudo inverse (Golub and Van Loan, 1996), we obtain

$$(W_f^{-1}\hat{X}W_p)^TW_f^{-1}\hat{X}W_p$$

$$\leq ((W_p^{-1}D_p)^+)^TW_f^{-1}D_f(W_p^{-1}D_p)^+$$

$$\leq (\gamma - 1)((W_p^{-1}D_p)^+)^TW_f^{-1}D_f(W_p^{-1}D_p)^+$$

$$\leq (\gamma - 1)I.$$ \hfill \text{(25)}

Namely, there exists $X$ satisfying the equations (20) and (21) simultaneously if and only if the matrix inequality (22) holds.

Now, the condition (22) can be represented as a linear matrix inequality

$$\begin{bmatrix}
\gamma - 1)(W_p^{-1}D_p)^TW_p^{-1}D_p & D_f^TW_f^{-1}
\end{bmatrix}
\begin{bmatrix}
I
\end{bmatrix} \geq 0$$ \hfill \text{(26)}

on variables $\gamma$ and $D_f$. Therefore, in order to compute the most robust input, it is enough for us to solve the minimization problem of $\gamma$ subject to the LMI condition (26) and to determine the unknown variables in $D_f$.

Note that we only need the weights $W_p$, $W_f$ and the input-output data arrays $D_p$, $D_f$ to determine the input, and we do not have to compute $X$. Furthermore, the size of the inequality (26) depends on the McMillan degree and the relative degree of the plant, and is independent of the number of the rows of the input-output data array $D_p$, that is, the number of data stored in the array. Note also that the input given as the solution of this minimization problem is equal to the input obtained by the algorithm in Ikeda et al. (2001) in case there is no noise.

### 4. A NUMERICAL EXAMPLE

In this section, we present a numerical example and see that the proposed algorithm works well.

We assume that the plant can be described by the input-output difference equation (2), whose coefficients are originally

$$a_1 = 1.2, \quad a_2 = 0.27, \quad a_3 = -0.04,$$

$$b_1 = 0.6, \quad b_2 = 0.21,$$ \hfill \text{(27)}

and becomes

$$a_1 = 1, \quad a_2 = 0.19, \quad a_3 = -0.03,$$

$$b_1 = 0.4, \quad b_2 = 0.17,$$ \hfill \text{(28)}

at $k = 25$. Furthermore, we assume that noise $u_k$ and $w_k$ having uniform distributions $[-n_k, n_k]$ are added at the input and the output of the plant respectively.

The algorithm collects the input-output data up to $k = 9$ in order to make the array (1), and starts to control the plant at $k = 9$. Here, we used the input data

$$u_0 = 10, \quad u_1 = -2, \quad u_2 = 3, \quad u_3 = 0.5,$$

$$u_4 = 4, \quad u_5 = -1, \quad u_6 = -4, \quad u_7 = 2.5,$$

$$u_8 = 7.5, \quad u_9 = 5,$$
and obtained the output data
\[ y_0 = 2.00, \quad y_1 = -0.70, \quad y_2 = 6.65, \quad y_3 = -6.81, \]
\[ y_4 = 7.73, \quad y_5 = -6.24, \quad y_6 = 7.63, \quad y_7 = -6.93, \]
\[ y_8 = 3.39, \quad y_9 = -1.23 \]
for a certain initial condition.

Fig. 2 and Fig. 3 illustrate the simulation results obtained by the existing algorithm (Ikeda et al., 2001) and the proposed algorithm respectively. In each figure, the solid line is the output \( y_k \), and the dotted line is the reference signal \( y^*_k \). Here, we chose the weight \( W_p \) as a diagonal matrix having diagonal elements \( 1, 1, 1, 1, \ldots \) from the bottom and the weight \( W_f \) as the equation (13).

The algorithm proposed in Ikeda et al. (2001) selects independent rows from the input-output data array according to a numerical machine precision. This selection is sensitive to the given data. Therefore, as shown in Fig. 2, the plant perturbation and the noise affect the controlled output dramatically. On the other hand, by using the algorithm proposed in this paper, the output tracks the reference signal almost exactly as shown in Fig. 3, even if the input and output noise are larger than those of Fig. 2.

5. CONCLUDING REMARKS

In this paper, we have proposed a data-driven tracking control algorithm for an arbitrary reference signal. We have reviewed the existing algorithm in noise free case (Ikeda et al., 2001) through a matrix equation on data arrays. When control and measurement noise exists, it appears on the data array additively and the matrix equation has an error term. We have proposed the most robust control input to minimize the sensitivity of this error term to the noise. This input can be computed via an LMI optimization problem based on the input-output data arrays, which is a significant feature of the proposed algorithm.

REFERENCES


Fig. 2. Output responses via the conventional method

![Fig. 2. Output responses via the conventional method](image1.png)

(b) Noise: \( n_\delta = 0.1 \)

Fig. 3. Output responses via the proposed method

![Fig. 3. Output responses via the proposed method](image2.png)


