[*] Study on an Optimal Design Method for Control Systems based on Bayesian Optimization

Koichi Hirota[†]
Graduate School of Advanced Science and Engineering, Hiroshima University, 1-4-1, Kagamiyama
Higashi-Hiroshima city, Hiroshima, Japan

Shin Wakitani,
Graduate School of Advanced Science and Engineering, Hiroshima University, 1-4-1, Kagamiyama
Higashi-Hiroshima city, Hiroshima, 739-8527 Japan
E-mail: wakitani@hiroshima-u.ac.jp

Toru Yamamoto,
Graduate School of Advanced Science and Engineering,Hiroshima University, 1-4-1, Kagamiyama
Higashi-Hiroshima city, Hiroshima, 739-8527 Japan
E-mail: yama@hiroshima-u.ac.jp

Abstract

When designing a complex product system, must be considered the optimal function distribution between the controller and the plant. This makes it possible to design a plant that is easy to control for fixed PID control, and the performance of the entire system can achieve the desired performance. Here, the evaluation function of the system is calculated from the output error and the input difference value. Then, the optimum design parameters are calculated by using Bayesian optimization.

Keywords: Optimal Allocation of Function, Model-Based Development, Bayesian Optimization, Evaluation Function,.

1. Introduction

Most product systems consist of a combination of multiple subsystems. Appropriate setting and design of operating goals (functional goals) of these subsystems are important to achieve and maintain the target performance of the system. In addition, the performance required by users for products is becoming more diverse and complex year by year. It is necessary to efficiently design a combination of subsystems to satisfy these required performances in a short period of time and at a low cost. In recent years, model-based development (MBD) has been attracting attention in the industrial world, and its introduction is being active. In the MBD process, entire subsystems are modeled as simulation models and are connected to each other on the simulation environment (such as MATLAB/Simulink). Moreover, a user can evaluate the subsystem's performance and easily improve their design without any prototyping. Meanwhile, the actuators of many subsystems are becoming electrified. Therefore, a feedback control system including controllers represented by by-wire technology is used in various situations. However, in many situations of control system design, plant design is completed in advance. In the conditions of this plant characteristic, the structure and parameters of the controller are often adjusted to satisfy the desired control performance. In particular, when the control target includes high-order components, strong non-linearity, and large dead time, it is difficult to achieve high control performance with a low-order controller such as a PID
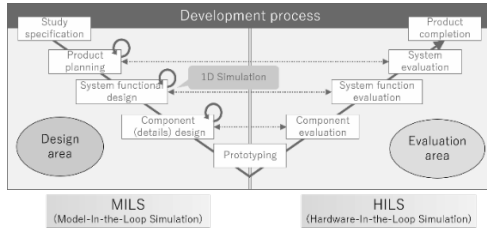
Fig. 1. V-shaped development process.

controller. On the other hand, if a high-order controller is introduced, the adjustment of control parameters becomes complicated, and the uncertainty of the system may impair the stability of the control loop. Therefore, in subsystem design, a plant and a controller are required to be optimized at the same time for the functional goals determined upstream design process. Based on the background, this study considers the simultaneous optimum design of the controller and the plant.

2. Simultaneous Optimal Design of Controller and Plant in MBD

2.1. V-shaped development process

Figure 1 shows the product development process, V-shaped development process, in MBD. The V-shaped development process (V-model) is divided into the model-in-the-loop simulation (MILS) which is a design area using a model and the hardware-in-the-loop simulation (HILS) which is a verification area that uses hardware and a model simultaneously. Since this research is mainly related to the design method, only MILS is be explained. Firstly, the necessary functions for a designed product system and specifications will be examined based on the results of market research. Next, the product system is decomposed into several subsystems, and the functions that these subsystems should satisfy are examined using 1D simulation technology. Furthermore, the parts (detailed design), that compose a subsystem is designed so that it operates in the same way as the functional model determined in the upstream design process. In this process, a detailed design including the shape of parts is performed using 3D simulation technology.

2.2. Optimal allocation of plant and controller functions

When designing a feedback system composed of a plant and a controller, the controller is usually designed from the state in which the plant has already been designed.

However, when designing a control system, there are input restrictions and non-linearities in the designed plant, and these characteristics appear as constraints when designing the controller. Therefore, it is often the case that the desired control performance cannot be obtained with low-order controllers such as PID controllers that are usually introduced in systems. However, if the spread of MBD makes it possible to allocate the optimum functions of the plant and controller in advance, the performance of the entire system will be closer to the desired performance by designing a plant that is easy to control for fixed PID controller. Therefore, this study, considers the optimum allocation method of plant and controller functions for the system function in the design phase. In the subsystem, to reach a given functional goal is the most important. Where, it is assumed that the functional goal can be described as the dynamic input/output relationship, that is the reference model[1]. The error between the output of the subsystem corresponding to a certain input scenario and the output of the reference model is evaluated as follows.

$$I_e = \frac{1}{N} \sum_{k=1}^{N} \left[1 - \exp\left\{ -\frac{(y_r(k) - y(k))^2}{2\sigma_e^2} \right\} \right] . \quad (1)$$

However, N represents the number of data acquired in the simulation evaluation interval. In Eq. (1), the Gaussian kernel is used for the output error. As a result, the value of Eq. (1) is normalized between 0 to 1 and the permissible ratio to the error can be adjusted by σ_e^2. Moreover, in the controller design, it is desirable to impose some restrictions on the controller output from the viewpoint of actuator protection and noise robustness. Here, the following evaluation function is designed for the input difference value $\Delta u(t) = u(t) - u(t-1)$.

$$I_e = \frac{1}{N} \sum_{k=1}^{N} \left[1 - \exp\left\{ -\frac{(\Delta u(k))^2}{2\sigma_{\Delta u}^2} \right\} \right] . \quad (2)$$

By using the Gaussian kernel as in Eq. (2), the evaluation value for the input difference value is normalized, and a certain permissible ratio is given as $\sigma_{\Delta u}$. For Eqs. (1) and (2), the evaluation function for optimizing the entire subsystem is defined by the following equation.

$$I = \lambda I_e + (1 - \lambda) I_{\Delta u} . \quad (3)$$

However, $\lambda (0 \leq \lambda \leq 1)$ is an adjustment parameter.
Based on the above equation, if the design parameter vector composed of control parameters and plant parameters is $\boldsymbol{\theta} = (\theta_1, \theta_2, \cdots, \theta_n,)^T$, the optimization problem is formulated by the following equation.

$$\boldsymbol{\theta}^* = \arg \min_{\boldsymbol{\theta}} \left(I(\boldsymbol{\theta}) \right) . \quad (4)$$

The next chapter explains the design parameter optimization method based on Bayesian optimization.

3. Bayesian Optimization

Bayesian optimization[2,3] can be divided into the regression process based on Gaussian process regression using obtained input/output data and a process for determining the next search point by calculating an acquisition function.

3.1. Gaussian process regression

In the Gaussian process regression, the mean and variance of the system output are estimated. First, this calculation method is explained. Let the input x and the output y are expressed by Eqs. (5) and (6).

$$x = (\theta_1, \theta_2, \cdots, \theta_N)^T . \quad (5)$$
$$y = (I_1, I_2, \cdots, I_N)^T . \quad (6)$$

Where, N is the number of data.

The feature vector of x is given as Eq. (7).

$$\phi = (\phi_0(x), \phi_1(x), \cdots, \phi_H(x))^T . \quad (7)$$

If the weights vector is described as $w = (w_0, w_1, \cdots, w_H)^T$, the linear regression model becomes Eq. (8).

$$\begin{pmatrix} \hat{y}_1 \\ \hat{y}_2 \\ \vdots \\ \hat{y}_N \end{pmatrix} = \begin{pmatrix} \phi_0(x_1) & \phi_1(x_1) & \cdots & \phi_H(x_1) \\ \phi_0(x_2) & \phi_1(x_2) & \cdots & \phi_H(x_2) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_0(x_N) & \phi_1(x_N) & \cdots & \phi_H(x_N) \end{pmatrix} \begin{pmatrix} w_0 \\ w_1 \\ \vdots \\ w_H \end{pmatrix} . \quad (8)$$

$$\hat{y} = \Phi w . \quad (9)$$

However, $\hat{y}$ is the predicted value of the output y and is shown by Eq. (10).

$$\hat{y} = (\hat{y}_1, \hat{y}_2, \cdots, \hat{y}_N)^T . \quad (10)$$

In addition, Φ is called a design matrix and becomes Eq. (11).

$$\Phi = \begin{pmatrix} \phi_0(x_1) & \phi_1(x_1) & \cdots & \phi_H(x_1) \\ \phi_0(x_2) & \phi_1(x_2) & \cdots & \phi_H(x_2) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_0(x_N) & \phi_1(x_N) & \cdots & \phi_H(x_N) \end{pmatrix} . \quad (11)$$

Where, by assuming $y = \hat{y}$, Eq. (9) can be obtained.

$$y = \Phi w . \quad (12)$$

Here, if a weights vector w is generated from the Gaussian distribution $w \sim N(0, \lambda^2 I)$, y also follows a Gaussian distribution, then the mean E $[y]$ and the covariance matrix K becomes as follows.

$$E[y] = E[\Phi w] = \Phi E[w] = 0 . \quad (13)$$

$$\begin{aligned} K &= E[y y^T] - E[y]E[y] \\ &= E[(\Phi w)(\Phi w)^T] \\ &= \Phi E[w w^T]\Phi^T \\ &= \lambda^2 \Phi \Phi^T . \end{aligned} \quad (14)$$

Therefore, the distribution of y is a multivariate Gaussian distribution shown in Eq. (12).

$$y \sim N(0, \lambda^2 \Phi \Phi^T). \quad (15)$$

The function that finds the value of the K element K_{ij} is called a kernel function. Kernel functions represent the similarity of inputs. There are several types, and the Gaussian kernel is widely used. The Gaussian kernel is shown in Eq. (13).

$$k(x_i, x_j) = \theta_1 \exp\left(-\frac{\|x_i - x_j\|^2}{\theta_2} \right) . \quad (16)$$

$\|\cdot\|$ indicates the Euclidean norm of the vector. In addition, θ_1 and θ_2 are parameters that determine the properties of the kernel function. Here, consider finding the prediction point y^* for an arbitrary vector x^*. First, using the output data vector in Eq. (6) and the predicted output y^*, a new vector y^{new} is defined as follows.

$$y^{new} = (y^T, y^*)^T . \quad (17)$$

If the covariance matrix K^{new} is calculated from y^{new}, x, and x^*, then the following relationship is hold.

$$y^{new} \sim N(0, K^{new}). \quad (18)$$

Here, the covariance matrix K^{new} is expressed by the following equation.

$$K^{new} = \begin{pmatrix} K & k_* \\ k_*^T & k_{**} \end{pmatrix} . \quad (19)$$

$$k_* = (k(x^*, x_1), k(x^*, x_2), \cdots, k(x^*, x_N))^T . \quad (20)$$

$$k_{**} = k(x^*, x^*). \quad (21)$$

From Eq. (18), the mean and variance of y^* can be calculated by Eq. (23).

$$y^* \sim N(\mu, \sigma^2). \quad (22)$$

$$\sim N(k_*^T K^{-1} y, k_{**} - k_*^T K k_*). \quad (23)$$

From the above results, the mean μ and variance σ^2 of the unknown y^* corresponding to arbitrary point x^* can be derived.

3.2. Acquisition function

The acquisition function is an equation calculated using the mean μ and standard deviation σ obtained by Gaussian process regression. This function is used to determine the next point to search. There are several types of acquisition functions, in this study, the following Lower Confidence Bound (LCB) is introduced.

$$a_{LCB} = \mu - \sqrt{\frac{\log N}{N}} \sigma . \quad (24)$$

The LCB searches for a place with a large standard deviation, that is, an unknown region, as the number of data decreases. Also, as the number of data increases, the search proceeds based on the data. From the above, it is possible to search for the optimum value of the function obtained by Bayesian optimization globally and efficiently.

4. Numerical Example

The effectiveness of the proposed design method is evaluated by numerical example. Here, consider a control system consisting of a controller, actuator, and plant as shown in Fig. 2.
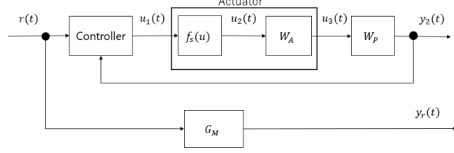

Fig. 2 Feedback control system

Here, $f_s(u)$ and W_A are the characteristics of the actuator, and W_P is the dynamic characteristics of the plants that make up the subsystem. $f_s(u)$ is a function that expresses the input saturation of the actuator and is given by the following equation.

$$u_2(t) = \begin{cases} u_{min}, & (u_1(t) < u_{min}) \\ u_1(t), & (u_{min} \leq u_1(t) < u_{max}) \\ u_{max}, & (u_1(t) \geq u_{max}) \end{cases} \quad (25)$$

The transfer function of plant 1 (W_A) is shown in the following equation.

$$W_A(s) = \frac{K_A}{1 + T_A s}. \quad (26)$$

Here, T_A is the time constant of the actuator, and K_A is the maximum output of the actuator. The transfer function of plant 2 (W_P) is shown in the following equation.

$$W_P = \frac{K\omega^2}{s^2 + 2\zeta\omega s + \omega^2}. \quad (27)$$

Where,

$$K = \frac{1}{k_P}, \zeta = \frac{D_P}{2\sqrt{k_P M_P}}, \omega = \sqrt{\frac{k_P}{M_P}}. \quad (28)$$

However, it is assumed that the system parameters M_P, k_P, and D_p are all positive real numbers. In this simulation, K_A, k_P, and D_P are the design parameters, and the other parameters are the fixed values shown in Table 1. In the controller design, the PID gains are determined based on the pole assignment controller design only using the model of plant 2. Here, the reference model expressing the desired function of the subsystem is given by the following equation.

$$G_m = \frac{1}{\left(1 + \frac{\sigma}{3}s\right)^3}. \quad (29)$$

The parameter $\sigma = 0.5$ for the rise time was set in this case. The evaluation function uses Eq. (3), and Table 2 shows the parameters used in the function design.

Table 1. Plant parameters

parameters	value
u_{min}	0
u_{max}	1
T_A	0.01
M_P	10

Table 2. Evaluation function parameters

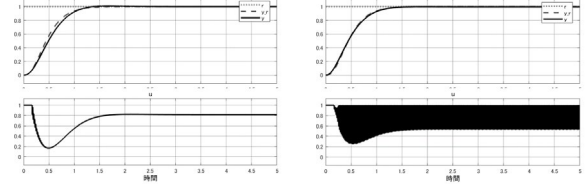| parameters	value
λ	0.5
σ_e	0.1
$\sigma_{\Delta u}$	0.01

Fig. 3. Control result using parameters (left-hand side: $\lambda = 0.5$, right-hand side: $\lambda = 1.0$).

As a result of Bayesian optimization under the condition $\lambda = 0.5$, the optimum values became $K_A = 61$, $k_P = 50$, $D_P = 0.1$. At this time, the control result is shown in Fig. 3 (left-hand side). From the control result, it can be seen that the output y generally follows the reference output y_r. If $\lambda = 0.5$ in the evaluation function, the optimum values are $K_A = 102$, $k_P = 80$, $D_P = 35$ when Bayesian optimization is performed. At this time, the result of the simulation is shown in Fig. 3 (right-hand side). Compared with the result obtained by $\lambda = 0.5$, it seems that the output error is smaller. However, since it can be seen that the input fluctuates drastically, it can conclude that a system in which the fluctuation of the input is suppressed can be designed by adding $I_{\Delta u}$ to the evaluation function.

5. Conclusion

This study proposes an optimal design method of a subsystem composed of a plant and a controller by using Bayesian Optimization. This paper defines a criterion for optimizing the plant parameters so that the plant is designed suitable for a controller. The effectiveness of the proposed design scheme is verified by a simulation example.

6. References

1. Takeshi Shigemasa, Yasuo Takagi, Yoshinori Ichikawa and Toshiyuki Kitamura, A Practical Reference Model for Control System Design, *Transactions of the Society of Instrument and Control Engineers*, Vol. 19, Issue. 7, 1983, pp.592-594.
2. Daichi Mochihashi and Ooba Shigeyuki, Gaussian Process and Machine Learning ~ Machine Learning Professional Series, Kodansya, Tokyo,Japan, 2019, pp. 57-95.
3. C. M. Bishop, Pattern Recognition and Machine Learning, Springer, New York, 2012