Proceedings of the American Control Conference
Philadelphia, Pennsylvania • June 1998

Control-Relevant Experiment Design: A Plant-Friendly, LMI-based Approach

Brian L. Cooley
Jay H. Lee *

Department of Chemical Engineering
Auburn University
Auburn, AL 36849-5127

Stephen P. Boyd

Department of Electrical Engineering
Stanford University
Stanford, CA 94305-4055

Abstract

In this paper, the non-convex control-relevant experiment design developed in earlier works is transformed to a semidefinite program with linear matrix inequality constraint. As a result, the computational burden is greatly reduced. An example is presented to demonstrate that the reduced computation time may be obtained without the cost of degraded performance.

1 Introduction

In recent years, both researchers and practitioners have increasingly pointed to the data generation (i.e., plant test) stage as the most crucial in the system identification process. Indeed, it is the only stage that requires interactions with the process, and as a result it tends to be the dominant factor influencing the time and cost associated with the system identification process.

Clearly, minimizing the impact of the data generation stage on normal plant operations is highly desirable. This translates into two objectives: 1) gathering the most useful data in the allotted experiment time and 2) insuring that the product manufactured during the plant test meets quality specifications. Meeting these objectives using only experience and intuition is often difficult.

Developing a rigorous framework for meeting the above-mentioned objectives requires a mathematical definition of "most useful data" and a mathematical expression of the acceptable operating window. Undoubtedly, the usefulness of a set of data depends on the intended application of the identified model. In this paper, we will restrict our attention to generating data for models to be used in model-based controller design.

Cooley and Lee [2] recently developed a rigorous framework for generating data containing the most useful information for designing a model-based controller. The key idea is to choose the test signal such that the expected value of a closed-loop performance index is minimized. While the work in [2] rigorously addresses designing test signals for controller performance and incorporates operating window constraints, the resulting design procedure is computationally burdensome. In this paper, we seek to address these computational aspects by approximating the problem using a first-order Taylor series expansion. The resulting optimization problem is a linear matrix inequality (LMI) which is more easily solved.

This paper is arranged as follows. In Section 2, the experiment design problem is developed. In Section 3, we present the main result, a computationally efficient solution to the experiment design problem. In Section 4, we demonstrate the usefulness of the approach through a numerical example. Finally, in Section 5 we offer our conclusions.

2 Problem Formulation

In this section, we will briefly develop the control-relevant experiment design problem for a specific choice of controller performance measure and model structure. For a more complete development of the experiment design problem, the interested reader is referred to [2].

We assume that the underlying system may be described as the following multivariable linear time invariant system:

$$y(t) = G(q)\,u(t) + d(t) \tag{1}$$

In the above, $y \in \Re^{n_y}$ is the output, $u \in \Re^{n_u}$ is the

*To whom all the correspondence should be addressed, phone (205)844-2060, fax (205)844-2063, e-mail jhl@eng.auburn.edu

0-7803-4530-4/98 $10.00 © 1998 AACC

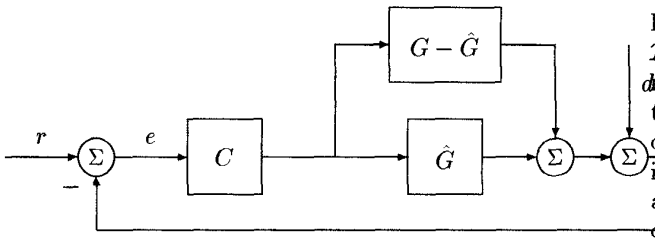

Figure 1: The closed-loop system.

input, and G is the system transfer function. q is the forward shift operator (i.e., $q^{-1}u(t) = u(t-1)$). We take $d \in \Re^{n_y}$ to be a random sequence representing noise / disturbance effects.

In this paper, we are concerned with designing experiments for identifying models to be used in model-based control schemes. That is, we will assume a model structure and identify a set of parameters ($\hat{\Theta}$) to obtain a model of the plant ($\hat{G}(q,\hat{\Theta})$). Based on that model, a controller will be designed. Note that in general the control depends on the model and the disturbance / reference signal characteristics through a mapping c:

$$C(q) = c(\hat{G}(q,\hat{\Theta})) \tag{2}$$

The resulting closed-loop system is depicted in Figure 1.

It is only natural that the data used to generate the model will influence the model quality, and thus, the resulting controller performance. What we desire is a method to design the experiment such that the best performance is achieved.

At the time the experiment is designed, the plant, the model (and therefore the controller), and the disturbance characteristics are unknown. The experimenter may have some prior knowledge about the disturbance characteristics and the plant. For practical and computational reasons, it is convenient to formulate the input design problem in a statistical framework. Here, we treat the plant, model, and controller as random variables and minimize the expectation of the performance index conditional to any prior knowledge available. The minimization is performed with respect to any input constraints that may exist. Although the plant is unknown, we can also add output constraints by satisfying them at a specified probability μ. Thus, the input design problem becomes:

$$\min_{u(1),\cdots,u(N)} \; E\{J|\mathcal{I}\} \tag{3}$$

$$u(t) \in \mathcal{U} \quad t = 1,\cdots,N \tag{4}$$

$$P_1[y(t) \in \mathcal{Y}] \geq \mu \quad t = 1,\cdots,N \tag{5}$$

In the above, J is a measure of closed-loop performance, $\mathcal{I}$ is the *information vector* consisting of all a priori information such as the prior distribution of the plant and the disturbance characteristics, and $Ea|b$ denotes the conditional expectation of a with respect to b. $P_1[a]$ is the probability that event a will occur. The sets $\mathcal{U}$ and $\mathcal{Y}$ are mathematical descriptions of the input and output constraints and are typically characterized by magnitude and rate bounds. N is the number of data collected.

The above formulation is intuitively appealing for iterative designs in which the model is updated between successive experiments and the next experiment is designed based on the updated model. In such a situation, (3)-(5) become:

$$\min_{u_k(1),\cdots,u_k(N)} \; E\{J_k|\mathcal{I}_k\} \tag{6}$$

$$u_k(t) \in \mathcal{U}_k \quad t = 1,\cdots,N \tag{7}$$

$$P[y_k(t) \in \mathcal{Y}_k] \geq \mu_k \quad t = 1,\cdots,N \tag{8}$$

where the index k has been introduced to denote the k^{th} experiment (e.g., $\mathcal{I}_k$ is the a priori information available for designing the k^{th} experiment). This idea is easily extendible to an adaptive design framework, in which the model is updated at every (or every few) sampling instants, and the experiment is re-computed in receding horizon fashion [2].

While there are many ways to measure controller performance, in this paper we will restrict our attention to mean squared error as the performance measure:

$$J_k = E\left\{e^T e \middle| G, C_k(\hat{G}_k), \Phi_\eta\right\} = \text{trace}\left(E\left\{ee^T \middle| G, C_k(\hat{G}_k), \Phi_\eta\right.\right. \tag{9}$$

In the above, e is the difference between the output and the reference signal, as depicted in Figure 1. Note that the mean squared error is computed for a *given* plant, controller, and disturbance / reference signal characteristic. This is why the conditional expectation is used in (9). It is straightforward to show the following.

$$J_k = \frac{1}{2\pi}\int_{-\pi}^{\pi} \text{trace}\left[\hat{S}_k\left(I + (G-\hat{G}_k)\hat{G}_k^{-1}\hat{H}_k\right)^{-1}\Phi_\eta\left(I + (G- \right. \tag{10}$$

In the above, $\hat{S}_k = (I + \hat{G}_k C_k)^{-1}$ is the so-called nominal sensitivity function, expressing the closed-loop relationship (based on the model $\hat{G}_k$) between the output error e and the disturbance d / the reference signal r, and $\hat{H}_k = \hat{G}_k C_k (I + \hat{G}_k C_k)^{-1}$ is the complementary sensitivity function (i.e., $\hat{S}_k + \hat{H}_k = I$). Φ_η is the spectrum of the combined reference and disturbance signal ($\eta = r - d$) and $\{\cdot\}^*$ denotes the complex conjugate

transpose (A^{-*} is the inverse of A^*). The above expression is an explicit relationship between model error and performance.

Notice that J_k depends on the unknown plant G, the unknown model $\hat{G}_k$, and the unknown controller C_k contained in the sensitivity function $\hat{S}_k$ and in the complementary sensitivity function $\hat{H}_k$. We treat the unknowns as stochastic quantities and take the conditional expectation of the performance index with respect to the information available prior to the experiment:

$$E\{J_k|\mathcal{I}_k\} = \frac{1}{2\pi}\int_{-\pi}^{\pi} \text{trace}\left[E\left\{\hat{S}_k\left(I+(G-\hat{G}_k)\hat{G}_k^{-1}\hat{H}_k\right)^{-1}\Phi_\eta \times \left(I+(G-\hat{G}_k)\hat{G}_k^{-1}\hat{H}_k\right)^{-*}\hat{S}_k^*\middle|\mathcal{I}_k\right\}\right]d\omega \quad (11)$$

In the above, $\mathcal{I}_k$ is the information vector available at the start of the k^{th} experiment

It is difficult to solve the optimization in (6) with J expressed as in (11) because 1) $\hat{S}_k$ and $\hat{H}_k$ may have complex dependence upon $\hat{G}_k$ and 2) the expectation is difficult to evaluate with respect to $\hat{G}_k$. However, by restricting the controller design to be direct synthesis (i.e., $\hat{S}_k$ is determined a priori) and linearizing e with respect to $\hat{G}_k$ about G_{k-1}, we obtain the following.

$$E\{J_k|\mathcal{I}_k\} \approx \frac{1}{2\pi}\int_{-\pi}^{\pi} \text{trace}\left[\hat{S}_k\Phi_\eta\hat{S}_k^* + \hat{S}_k E\left\{\left(\hat{G}_k-G\right)\left[\hat{G}_{k-1}^{-1}\hat{H}_k\Phi_\eta\hat{H}_k^*\hat{G}_{k-1}^{-*}\right]\left(\hat{G}_k-G\right)^*\middle|\mathcal{I}_k\right\}\hat{S}_k^*\right]d\omega \quad (12)$$

Equation (12) is much easier to use in (6) than (10) because the only unknown is $\hat{G}_k - G$ (all the other matrices are contained in $\mathcal{I}_k$), which is treated as a stochastic variable whose distribution depends on the test signal (and the identification procedure). Given a relationship between the test signal and the distribution of $\hat{G}_k - G$, it will be possible to formulate the optimization problem in (6)-(8) using the performance measure approximation in (12).

For the purposes of this paper, we assume that the underlying system may be described by a finite expansion in a set of orthonormal basis functions $f_k(q)$ [5]:

$$y(t) = \sum_{k=1}^{n}\theta_k f_k(q)\{u(t)\} + d(t) = \Theta^T\phi(t) + d(t) \quad (13)$$

$$\Theta^T = \begin{bmatrix}\theta_1 & \cdots & \theta_n\end{bmatrix} \quad (14)$$

$$\phi(t) = \begin{bmatrix} f_1(q)\{u(t)\} \\ \vdots \\ f_n(q)\{u(t)\}\end{bmatrix} \quad (15)$$

with θ_i representing the (real-valued) expansion coefficient matrices. For ease of exposition, we take $\{d(t)\}$ to be a zero-mean i.i.d. random sequence with the identity matrix as its covariance.

Given the above assumptions and an a priori estimate $\hat{\Theta}_{k-1}$ and covariance Σ_{k-1}, we can obtain unbiased estimates of the parameter matrix Θ via least squares.

$$\hat{\Theta}_k = \Sigma_k\left(\sum_{t=1}^{N}\phi(t)y^T(t) + \Sigma_{k-1}^{-1}\hat{\Theta}_{k-1}\right) \quad (16)$$

$$\Sigma_k = \left(\sum_{t=1}^{N}\phi(t)\phi^T(t) + \Sigma_{k-1}^{-1}\right)^{-1} \quad (17)$$

In the above $\hat{\Theta}_k$ is the updated estimate. Similarly, Σ_k is the updated covariance matrix. We may compute the following expectation:

$$E\left\{\left(\hat{\Theta}_k-\Theta\right)\left(\hat{\Theta}_k-\Theta\right)^T\middle|\mathcal{I}_k\right\} = \left(\sum_{i=1}^{N}\phi(t)\phi^T(t) + \Sigma_{k-1}\right)^{-1} \quad (18)$$

The above expectation may be used to compute the following frequency domain expectation:

$$E\left\{\left(\hat{G}_k(e^{j\omega})-G(e^{j\omega})\right)^*\left(\hat{G}_k(e^{j\omega})-G(e^{j\omega})\right)\middle|\mathcal{I}_k\right\} = \Lambda(e^{j\omega})^* I \quad (19)$$

$$\Lambda(e^{j\omega}) = \begin{bmatrix}\Lambda_1(e^{j\omega}) \\ \vdots \\ \Lambda_n(e^{j\omega})\end{bmatrix} \quad (20)$$

where the Λ_k, $k = 1, \cdots, n$ are the frequency response of the corresponding basis function f_k. For example, if $f_k(q) = q^{-k}$, then $\Lambda_k(e^{j\omega}) \triangleq e^{-j\omega k}I_{n_u}$. The above expression holds because $G = \Theta^T\Lambda$, and it is an explicit relationship between the test signal (which appears in the matrix Σ_k) and the expectation on the left hand side of (19) (which we will loosely refer to as the "frequency domain error covariance").

Based on (12), we present the following result.

Theorem 1 *The approximate control-relevant experiment design problem defined by (3)-(5) with $E\{J_k|\mathcal{I}_k\}$ given in (12) with prior estimate $\hat{\Theta}_{k-1}$, covariance Σ_{k-1}, and disturbance / reference signal spectrum Φ_η corresponds to the following "weighted trace optimal" experiment design problem:*

$$\min_{u(1),\cdots,u(N)} \text{trace}(W_k\Sigma_k) \quad (21)$$

subject to the constraints of (4)-(5). Σ_k is given by (17), and the weighting matrix W_k is given by

$$W_k = \frac{1}{2\pi}\int_{-\pi}^{\pi}\text{trace}\left(\hat{S}_k^*\hat{S}_k\right)\Lambda\hat{G}_{k-1}^{-1}\hat{H}_k\Phi_\eta\hat{H}_k^*\hat{G}_{k-1}^{-*}\Lambda^* d\omega \quad (22)$$

Proof: A proof is presented in [2].

3 The Main Result

3.1 Reduction of Computational Burden

While the proposed experiment design technique has been rather simply formulated as the optimization in (3)-(5) using the linearized performance index in (12), there are several issues associated with solving the optimization problem. Perhaps foremost among the difficulties is the fact that the nonconvexity of the objective function leads to local solutions and a large computational burden. However, we notice the following:

$$\min_{U_k} [\text{trace}(W\Sigma_k)] \tag{23}$$

$$\Downarrow$$

$$\min_{U_k, Z} [\text{trace}(Z)] \tag{24}$$

$$\text{s.t.} \quad \begin{bmatrix} Z & I \\ I & P_k W^{-1} \end{bmatrix} \geq 0 \tag{25}$$

where $P_k = \Sigma_k^{-1}$. The optimization defined by (24) and (25) would be a semidefinite program with linear matrix inequality (LMI) constraint if P_k were an affine function of the test signal. Unfortunately, P_k is a quadratic function of the test signal. If we are willing to accept another approximation, P_k may be transformed into an affine function of U_k by performing a first order Taylor's expansion about a given test signal:

$$P_k = P_k|_{\bar{U}_k} + \sum_{t=1}^{N} \sum_{j=1}^{n_u} \left. \frac{dP_k}{du_j(t)} \right|_{U_k} (u_j(t) - \bar{u}_j(t)) \tag{26}$$

where $\overline{U}_k$ is a nominal test signal and $u_j(t)$ is the value of the j^{th} input at time t. The optimization problem defined by (24)-(26) is a LMI, for which there are very efficient methods of solution (for more on LMI's, see [1]). Thus, the computational burden is greatly reduced at the expense of providing a good nominal test signal. For purposes of designing experiments offline, this implies an iterative approach in which an initial condition is provided, the LMI solved, and the solution is used as the initial condition for re-solving the experiment design problem. Such a procedure would proceed until the objective function converged to within a certain tolerance. An on-line approach is easily amenable to a receding horizon implementation, in which only the first input move is implemented and the remainder are discarded. The discarded input moves can be used as the nominal test signal for the next optimization.

3.2 Output Constraints

In order to insure that the plant remains safe and profitable during experimentation, it is desirable to include output constraints in the experiment design framework. However, since the model is a stochastic quantity, hard constraints cannot be placed on the outputs. Instead, the constraints are satisfied with a given probability specified by the user. Once the probability level is specified, (ellipsoidal) confidence intervals may be drawn based on the probability distribution of the model parameters [3]:

$$\text{vec}(\Theta - \Theta_k)^T \Sigma_k^{-1} \text{vec}(\Theta - \Theta_k) \leq \chi_\gamma^2(p) \tag{27}$$

where γ is the confidence level, $p = n_y \cdot n_u \cdot n$ is the number of parameters, $\chi_\gamma^2(p)$ is the chi-square percentile for confidence γ and p degrees of freedom. The operator $\text{vec}(\cdot)$ stacks the columns of a vector into a matrix. The output constraints must hold for all parameters in the set (27). One ends up with extra linear inequality constraints that must be satisfied for parameters within an ellipsoid. With (24)-(25), this can be shown to be a convex problem (second order cone programming [4]).

In the next section, we will demonstrate how this particular formulation improves the performance of model-based controllers derived from identified models.

4 Numerical Example

4.1 Preliminaries

In this section, we will consider the following system:

$$y(t) = \frac{1}{s+1} \begin{bmatrix} 0.2031 & -0.1452 \\ 0.1952 & -0.2404 \end{bmatrix} u(t) \tag{28}$$

The above system was sampled with a sampling time of $T_s = 1$ and modeled via a finite impulse response model with 5 coefficient matrices ($n = 5$). An initial model was developed using 100 data points corrupted with noise of covariance $(0.3)^2 I_2$. Based on this initial model, experiments consisting of 100 input moves and output data points corrupted with noise of covariance $(0.3)^2 I_2$ were designed using both the nonconvex optimization in (6)-(7) and the LMI-based approach in (24)-(25) along with the constraint (7) and the approximation (26). In each case, the input moves were constrained to have magnitude less than 0.1. The experiments were designed based on Φ_η equal to the spectrum of an integrating system, and $\hat{S}$ as follows:

$$\hat{S} = \frac{1}{s+1} I_2 \tag{29}$$

Once the models were identified and the controllers designed, the performance was tested by subjecting the closed-loop systems to a setpoint change from $r = [0\ 0]^T$

Table 1: Simulation results

	Computation time	Mean-squared error
LMI-based approach	608	11.17
Non-convex optimization	10171	10.68

to $r = [1 \ -1]^T$. The performance was measured by the sum of the squared errors.

Table 1 presents the computation time and performance results of the two experiments. Clearly, the LMI-based solution has greatly reduced the time required for computing the test signal with only a slight increase in mean-squared error of the resulting controller.

5 Conclusions

This paper provides a computationally efficient (Linear Matrix Inequality) procedure for designing experiments. The procedure is based on an earlier developed nonconvex optimization. An example is presented to demonstrate the computational advantages of the approach and to show that the approximation made to arrive at the LMI do not cause dramatic reductions in the effectiveness of the experiment design procedure.

References

[1] S. P. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan. *Linear Matrix Inequalities in System and Control Theory.* SIAM, Philadelphia, 1994.

[2] B. L. Cooley and J. H. Lee. Control-relevant experiment design for multivariable systems. *Automatica*, 1997. submitted.

[3] B. L. Cooley and J. H. Lee. Integrated identification and robust control. *Journal of Process Control*, 1998. accepted for publication.

[4] M. S. Lobo, L. Vandenberghe, S. P. Boyd, and H. Lebret. Applications of second-order cone programming. *Linear Algebra and Applications*, 1998.

[5] P. M. J. Van den Hof, P. S. C. Heuberger, and J. Bokor. System identification with generalized orthonormal basis functions. *Automatica*, 31(12):1821–1834, 1995.