Model Invalidation:
A Connection between Robust Control and Identification

Roy S. Smith and John C. Doyle

Electrical Engineering, 116-81
California Institute of Technology
Pasadena, CA 91125, U.S.A.

Abstract

This paper begins to address the gap between the models used in robust control theory and those obtained from identification experiments by considering the connection between uncertain models and data. The model invalidation problem considered here is: given experimental data and a model with both additive noise and norm-bounded perturbations, is it possible that the model could produce the input/output data?

1 Introduction

Robust control theory now gives the engineer a set of analysis tools for linear models which include two types of uncertainty: additive noise and block structured, norm bounded perturbations entering the model in a linear fractional manner. Once a system is modeled, and the engineer is confident of the applicability of the model, the theory gives techniques for designing systems which are theoretically robust with respect to the these uncertainties. No theory makes statements about the performance or the stability of the actual physical system. Therein lies the problem for the engineer. Before the robust control methods can be applied the uncertainty must, in some sense, be identified. Current identification methods are well developed in the case where all of the residuals, or uncertainty, are attributed to additive noise. For models with both additive noise and norm bounded perturbations no such identification methods exist.

Once a model has been determined, perhaps by ad hoc methods, there must be some method of evaluating its applicability to the actual physical system. The engineer must be confident that the model will describe all input output behaviors of the system. This condition can never be guaranteed but it is possible to test a necessary condition: that the model be able to describe all observed input output behaviors of the system. This is simply the model invalidation question to be discussed here.

In this paper, the model invalidation question is translated into an optimization problem: find the minimum norm noise input meeting the constraints imposed by the assumptions of the model. A Lagrange multiplier approach to this optimization problem is pursued. For a large class of models in the robust control framework, this leads to conditions which are computable as the solution of a quadratic optimization problem. The results in this paper are, not surprisingly, highly reminiscent of the type of results one gets in the μ analysis framework. In particular, lower and upper bounds are derived and conditions under which the bounds are equal are obtained. Familiar issues of global convergence and dependence of the results on the number of blocks arise. This theory also appears to have potential connections with the work of Willems[1] and Krause[2], and we hope to explore this in the future.

The testing of a model against experimental data is often referred to as model validation. This is a misleading term as it is never possible to validate models — only to invalidate them. The title and subsequent preamble reflects this philosophy; however in keeping with the accepted usage we will return to the term model validation.

2 Model Validation

There is no identification theory for robust control models. "Black box" identification is in fact a poorly posed problem; The physical system can only be observed by input output measurements and, modulo considerations of observability from a particular output, the residuals can be attributed either entirely to additive noise or entirely to norm bounded perturbations. In practice an engineer will run many experiments to attempt to isolate the effects of noise from those of perturbations.

An assumption, inherent in the use of any model, is that it can describe any input output behavior that the physical system can produce. It is not possible to test this; however it is possible to find a necessary condition. This is just the model validation question: can the model account for all of the previously observed input output behavior? This will be formulated more rigorously in the context of robust control models in Section 2.3.

The model validation theory described here has additional applications. For example, of significant interest to operating engineers is the problem of fault detection. Given a design model and a controller in operation, the model validation theory gives a means of continuously assessing whether or not the physical system is still described by the design model. It will be seen that the techniques presented here produce the perturbation and noise that come closest to satisfying the conditions of the model. Gradual deterioration in a system may manifest itself as increasing perturbations and/or noise required for accountability of the data. A sudden failure may be identified by a sudden jump in the size of the required perturbations and/or noise.

2.1 A Generic Identification/Validation Model

For the purposes of discussing identification and model validation, a generic (P,Δ) structure is introduced in Figure 1. In identification experiments certain inputs to the system are known. These are denoted by u and partitioned from the other unknown inputs: w. In a physical system u might be actuator inputs and w might represent noise and disturbances. w is assumed to belong to a norm bounded set: $\|w\| \leq 1$. Measured outputs are represented by e and are also assumed to be known.

The Δ is also norm bounded but can in general be block diagonal. If there are f blocks the block structure is a tuple of f integers $(k_1,\ldots,k_f)$, giving the dimension of each block. It is without loss of generality that the blocks are assumed to be square. Define $\mathbf{\Delta}$ as

$$\mathbf{\Delta} = \{\mathrm{diag}\,(\Delta_1 \ldots \Delta_f) |\ \dim(\Delta_i) = k_i.\}$$

and the unit ball of $\mathbf{\Delta}$ by $\mathbf{B\Delta}$. The uncertain model representation will be abbreviated to

$$e = \mathrm{F}_u(P,\Delta)\begin{bmatrix} w \\ u \end{bmatrix}, \quad \Delta \in \mathbf{B\Delta}, \quad \|w\| \leq 1.$$

where $\mathrm{F}_u(\ ,\ )$ signifies that the upper block of the fractional transformation is closed.

2.2 A Brief Review of μ

There is a strong connection between the structured singular value μ ([3], [4]) and the approach to the model validation problem taken here. A brief review gives the results needed for this paper.

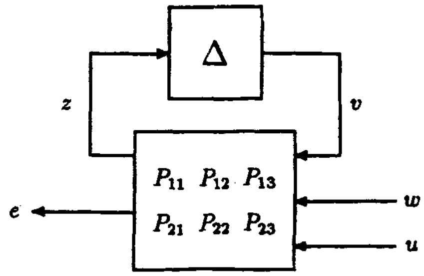

Figure 1: The Generic Structure for Identification and Model Validation Problems.

The positive real valued function μ can be defined on a matrix M by

$$\det(I - M\Delta) \neq 0 \quad \forall \Delta \in \mathbf{\Delta}, \quad \sigma_{\max}(\Delta) \leq \gamma \quad \text{iff } \gamma\mu(M) < 1.$$

γ is used here to illustrate that μ scales, i.e., for all $\alpha \in \mathcal{R}$, $\mu(\alpha M) = |\alpha|\mu(M)$.

The interconnection structure for μ analysis is that shown in Figure 1 with $u = 0$, $P_{13} = 0$, and $P_{23} = 0$. μ is essentially defined as the condition for robust stability: $\mathbf{F}_u(P, \Delta)$ is stable for all $\Delta \in \mathbf{B\Delta}$ iff $\mu(P_{11}) < 1$.

In general μ is difficult to calculate. A singular value test gives a more easily calculated upper bound. Define

$$\mathcal{D} = \{\operatorname{diag}(d_1 I_1, \ldots, d_f I_f) \mid d_i \in \mathcal{R}+, \ \dim(I_i) = k_i\}. \qquad (1)$$

All $D \in \mathcal{D}$ commute with Δ and a sufficient condition for robust stability is therefore

$$\inf_{D \in \mathcal{D}} \sigma_{\max}(D P_{11} D^{-1}) < 1. \qquad (2)$$

2.3 The Model Validation Problem Formulation

Referring to the model structure of Figure 1, the model validation problem can be formulated as follows. Given an interconnection structure P and an input/output experiment data pair (u,e), the model validation problem is:

Problem 1 (Model Validation)
Does there exist (w,Δ), $\|w\| \leq 1$, $\Delta \in \mathbf{B\Delta}$, such that

$$e = \mathbf{F}_u(P, \Delta) \begin{bmatrix} w \\ u \end{bmatrix}.$$

This simply states that there is an element of the model set and an element of the unknown input signal set such that the observed datum is produced exactly.

This paper will present and discuss a method for finding (w,Δ) meeting the constraints of the model: $\|w\| \leq 1$ and $\Delta \in \mathbf{B\Delta}$. Note that no statement is made relating the particular element of the model set or the particular element of the input signal set to any physical system or signal. Such a relationship does not exist. If no (w, Δ) pair meeting the above requirements exists then the model cannot account for the observed behavior. Such a tool is of use in selecting inappropriate models from a group of candidate models.

The model validation test is therefore a necessary condition for any model to describe a physical system. The fact that every experiment can be accounted for in this manner provides little information about the model and the system. There may be experiments, as yet unperformed, which will invalidate the model. The particular w and Δ do not necessarily bear any relationship to physical signals. If a consistent property is observed in the w or Δ from many experiments then it may be possible to reformulate the model with greater fidelity. There is no guarantee of this but any such model could of course be tested against the experimental data with the model validation procedure.

The properties of an admissible (w,Δ) are

$$\|\Delta\| \leq 1, \quad \Delta \in \mathbf{\Delta}, \quad \|w\| \leq 1. \qquad (3)$$

and, using the notation introduced for generic identification models in Figure 1

$$e \;=\; \mathbf{F}_u(P, \Delta_j) \begin{bmatrix} w \\ u \end{bmatrix} \;=\; P_{21} v + P_{22} w + P_{23} u.$$

If v and z are partitioned conformally with the Δ blocks, $\Delta \in \mathbf{\Delta}$ is equivalent to

$$\|v_i\| \leq \|z_i\|, \quad i = 1, \ldots, f.$$

where the subscript i represents the components corresponding to the i^{th} uncertainty block. For convenience define x as the vector $x = [v \ w]^T$ and define Q_i by

$$Q_i = \text{block row}(0_1, \ldots, 0_{i-1}, I_i, 0_{i+1}, \ldots, 0_f)$$

where 0_j is a block of zeros of dimension $k_i \times k_j$ and I_i is an identity of dimension $k_i \times k_i$. Q_i will be used to select the components corresponding to the i^{th} uncertainty block. The following two column vector notation will be used to partition the vector x into v and w and partition Px into z and e.

$$[Q_i \ 0]x = v_i, \quad [0 \ I]x = w, \quad [Q_i \ 0]Px = z_i, \quad \text{and} \quad [0 \ I]Px = e.$$

The existence of (w,Δ) meeting the constraints has now been reduced to $f+1$ norm conditions and an equality condition. Using the above definitions and the square of the norms for the test, Theorem 2 immediately follows.

Theorem 2 (Model Validation) *There exists (w,Δ) solving the model validation problem;*

$$e = \mathbf{F}_u(P, \Delta) \begin{bmatrix} w \\ u \end{bmatrix}, \quad \|w\| \leq 1, \quad \Delta \in \mathbf{B\Delta}$$

iff there exists x such that:

i) $\|[Q_i \ 0]x\|^2 \;\leq\; \left\| [Q_i \ 0] P \begin{bmatrix} x \\ u \end{bmatrix} \right\|^2, \quad i = 1, \ldots, f.$

ii) $\|[0 \ I]x\|^2 \leq 1.$

iii) $e = [P_{21} \ P_{22}]x + P_{23}u.$

3 An Optimization Formulation

Theorem 2 gives $f+2$ conditions on a candidate vector x such that x meets the conditions if and only if there exists w and Δ accounting for the observed data. This can be posed as an optimization problem in a number of ways. The choice of $\|w\|$ as an objective function will be considered here.

Furthermore P and Δ will be treated as constant matrices. This is the simplest case and can be used as the basis for the solution to a number of more complicated problems. For example, if P is a rational transfer function and Δ is an unknown complex constant at each frequency, robust stability and performance can then be treated on a frequency by frequency basis. The model validation problem can also be cast into this framework with these same assumptions, which then reduce the problem to the constant matrix case.

Define the matrices, T_i, $i = 1, \ldots, f$ by $T_i = Q_i^T Q_i$. The optimization problem associated with Theorem 2 is as follows.

Problem 3 (Minimum $\|w\|$, Constant Matrix)

$$\min_x f(x) \quad \text{subject to } g_i(x) \leq 0, \ \ i = 1, \ldots, f, \text{and } g_e(x) = 0.$$

where $f(x)$, $g_i(x)$, $g_e(x)$ are defined by

$$f(x) = x^* \begin{bmatrix} 0 & 0 \\ 0 & I \end{bmatrix} x. \qquad (4)$$

$$g_i(x) = x^* \begin{bmatrix} T_i & 0 \\ 0 & 0 \end{bmatrix} x - [\, x^* \ \ u^* \,] P^* \begin{bmatrix} T_i & 0 \\ 0 & 0 \end{bmatrix} P \begin{bmatrix} x \\ u \end{bmatrix}. \qquad (5)$$

$$g_e(x) = e - P_{23}u - [P_{21} \ P_{22}]x. \qquad (6)$$

In this case, if $\hat{x}$ achieves the above minimum subject to the constraints, then the remaining condition to test is $\| [0\; I]\, \hat{x} \| \leq 1$. If this condition is not satisfied then no other (w,Δ) pair exists meeting the constraints of the model and the model cannot therefore account for the datum.

Solution of the optimization problem above (Problem 3) will yield the minimum $\|w\|$ meeting the constraints. In the single uncertainty block case ($f = 1$) an alternative, equally physically motivated optimization problem arises from the choice of Equation 5 as the objective and Equations 4 and 6 as constraints. The solution of this problem gives the minimum norm Δ meeting the constraints. It would then remain to check if $\|\Delta\| \leq 1$. This formulation could be of interest in attempting to refine the weights on the uncertainty blocks in an identification procedure.

3.1 Removal of the Equality Constraint

In this section x is reparametrized to remove the equality constraint. This is not necessarily the most efficient way of performing the optimization but illustrates some theoretical points of interest. Consider the solutions to

$$e - P_{23}u = [P_{21}\; P_{22}]\, x \quad \text{as} \quad x = x_0 + Vy \tag{7}$$

with x_0 being the least squares solution:

$$e - P_{23}u = [P_{21}\; P_{22}]\, x_0 \quad \text{and} \quad Vy \in \operatorname{Ker}[P_{21}\; P_{22}],$$

with $\dim(y) = \dim(\operatorname{Ker}[P_{21} P_{22}])$. Note that Vy is simply the span of the right singular vectors corresponding to the zero singular values of $[P_{21}\; P_{22}]$.

Several cases are immediately possible:

i) $e - P_{23}u \notin \operatorname{Range}[P_{21}\; P_{22}]$. In this case there is no w or Δ that can account for the datum.

ii) $e - P_{23}u \in \operatorname{Range}[P_{21}\; P_{22}]$ but $\operatorname{Ker}([P_{21}\; P_{22}])$ is empty. There is a unique w and Δ specified by u and e. It now remains to check the norms of w and Δ to determine if the model can account for the datum.

iii) $e - P_{23}u \in \operatorname{Range}[P_{21}\; P_{22}]$ and $\dim(\operatorname{Ker}[P_{21}\; P_{22}]) > 0$. This is the generic case where the reparametrization of x has removed the equality constraint and reduced the dimension of the search for x.

For notational convenience define the set $\mathcal{X}$ by

$$\mathcal{X} = \{x \mid x = x_0 + Vy\}.$$

4 A Lagrange Multiplier Approach

4.1 Preliminaries

The following lemmas relate the properties of the Hessian of a function $f(x)$ (the matrix of second partial derivatives), denoted by $H[f(x)]$, to the convexity properties of the function.

Lemma 4 *A twice differentiable functional $f(x)$, defined on an open set $\Gamma \subset C^n$ is convex iff $H[f(x)]$ is positive semidefinite.*

Lemma 5 *$f(x)$ is a twice differentiable functional, defined on an open set $\Gamma \subset C^n$. If $H[f(x)]$ is positive definite then $f(x)$ is strictly convex on Γ.*

Consider the optimization problem

$$\min_x f(x) \quad \text{subject to } g_i(x) \leq 0, \quad i = 1, \ldots, f,$$

which is referred to as the primal problem. The Lagrangian is defined as

$$L(x,\lambda) = f(x) + \sum_{i=1}^{f} \lambda_i g_i(x), \quad \text{where } \lambda = [\lambda_1, \ldots, \lambda_f]^T.$$

The dual function $h(\lambda)$ and the dual problem are respectively defined as

$$h(\lambda) = \min_x L(x,\lambda), \quad \text{and} \quad \max_\lambda h(\lambda).$$

Under certain assumptions a solution of the dual problem can yield a solution of the primal problem. The following theorems express the well known relationship between the solutions to the dual and primal problems. Refer to Wismer and Chattergy[5] for details.

Theorem 6 *Define E as the region of λ space upon which $h(\lambda)$ is finite.*
$h(\lambda)$ is concave on convex regions of E.

Theorem 7 (Kuhn Tucker Saddlepoint) *A point $(\hat{x},\hat{\lambda})$ with $\hat{\lambda} \geq 0$ is a saddlepoint of the Lagrangian $L(x,\lambda)$ iff*

i) $\hat{x}$ minimizes $L(x,\hat{\lambda})$ over all x.

ii) $g_i(\hat{x}) \leq 0$, for all $i = 1,\ldots,f$.

iii) $\hat{\lambda}_i g_i(\hat{x}) = 0$, for all $i = 1,\ldots,f$.

Notice that condition *i)* requires a global optimization of the Lagrangian.

Theorem 8 *If the point $(\hat{x},\hat{\lambda})$ is a saddlepoint of the Lagrangian $L(x,\lambda)$ then $\hat{x}$ solves the primal problem*

4.2 Application to the Model Validation Problem

Lagrange multiplier techniques are now applied to Problem 3. Introducing the notation

$$A_i = \begin{bmatrix} T_i - P_{11}^* T_i P_{11} & -P_{11}^* T_i P_{12} \\ -P_{12}^* T_i P_{11} & -P_{12}^* T_i P_{12} \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 0 & 0 \\ 0 & I \end{bmatrix}$$

allows the Equations 4 and 5 of the optimization problem (Problem 3) to be expressed as

$$f(x) = x^* B x$$

and

$$g_i(x) = x^* A_i x - 2\,\mathrm{Re}\left\{ x^* \begin{bmatrix} P_{11}^* \\ P_{12}^* \end{bmatrix} T_i P_{13} u \right\} - u^* P_{13}^* T_i P_{13} u.$$

This makes the structure of the Hessian of the Lagrangian clear; the constraints are indefinite quadratic inequalities. The objective functional is however a positive semidefinite quadratic.

Now consider the inclusion of the constraint $x \in \mathcal{X}$. Using Equation 7 is it possible to formulate a Lagrangian in terms of y rather than x. To distinguish the fact that the equality constraint has been included, the objective function is denoted by $f_e(y)$ and the constraints are $g_{ei}(y)$. This gives

$$f_e(y) = y^* V^* B V y + 2\,\mathrm{Re}\{x_0^* B V y\} + x_0^* B x_0$$

and

$$\begin{aligned} g_{ei}(y) \;=\; & y^* V^* A_i V y \\ & + 2\,\mathrm{Re}\left\{ y^* V^* A_i x_0 - y^* V^* \begin{bmatrix} P_{11}^* \\ P_{12}^* \end{bmatrix} T_i P_{13} u \right\} \\ & - 2\,\mathrm{Re}\left\{ x_0^* \begin{bmatrix} P_{11}^* \\ P_{12}^* \end{bmatrix} T_i P_{13} u \right\} \\ & + x_0^* A_i x_0 - u^* P_{13}^* T_i P_{13} u. \end{aligned}$$

Note that the objective functional is still positive semidefinite and the constraints are still indefinite quadratic inequalities.

The Lagrangian, denoted by $L_e(y,\lambda)$ to emphasize the inclusion of the equality constraint, associated with the above is

$$L_e(y,\lambda) = y^* V^* (B + \sum_{i=1}^{f} \lambda_i A_i) V y + 2\,\mathrm{Re}\{y^* C_e(\lambda)\} + d_e(\lambda)$$

where

$$C_e(\lambda) = V^* B x_0 + \sum_{i=1}^{f} \lambda_i \left[V^* A_i x_0 - V^* \begin{bmatrix} P_{11}^* \\ P_{12}^* \end{bmatrix} T_i P_{13} u \right] \tag{8}$$

and

$$d_e(\lambda) = x_0^* B x_0 + \sum_{i=1}^{f} \lambda_i \left[x_0 A_i x_0 - 2\,\mathrm{Re}\left\{ x_0^* \begin{bmatrix} P_{11}^* \\ P_{12}^* \end{bmatrix} T_i P_{13} u \right\} - u^* P_{13}^* T_i P_{13} u \right]$$

Note that in this formulation y is unconstrained. The dual function associated with $L_e(y,\lambda)$ is denoted by $h_e(\lambda)$. The Lagrangian $L_e(y,\lambda)$ is the one of interest in solving Problem 3. However a study of the properties of $L(x,\lambda)$ will aid in understanding the equality constrained case.

4.3 The Properties of $L(x,\lambda)$

Define Λ as the region of λ space upon which the Hessian of $L(x,\lambda)$ is positive definite. Lagrange multipliers are generally defined with components $\lambda_i \geq 0$. Here they will be taken to be $\lambda_i > 0$ with the $\lambda_i = 0$ treated as a special case.

Now the property that the Hessian of the Lagrangian is strictly positive definite is shown to be equivalent to a maximum singular value test. Consider the Lagrangian for $\lambda \in \Lambda$.

The Hessian of the Lagrangian, denoted by $H[L(x,\lambda)]$ is

$$H[L(x,\lambda)] = B + \sum_{i=1}^{f} \begin{bmatrix} \lambda_i T_i - P_{11}^* \lambda_i T_i P_{11} & -P_{11}^* \lambda_i T_i P_{12} \\ -P_{12}^* \lambda_i T_i P_{11} & -P_{12}^* \lambda_i T_i P_{12} \end{bmatrix}$$

Define D by

$$\sum_{i=1}^{f} \lambda_i T_i = D^T D = D^* D = D^2,$$

where $d_i > 0$. Only $\lambda_i > 0$ are being considered so D^{-1} is well defined. Note also that D defined above is an element of $\mathcal{D}$ defined by Equation 1. Now

$$H[L(x,\lambda)] = \begin{bmatrix} D^2 - P_{11}^* D^2 P_{11} & -P_{11}^* D^2 P_{12} \\ -P_{12}^* D^2 P_{11} & I - P_{12}^* D^2 P_{12} \end{bmatrix}$$

The problem is to find D, and consequently λ_i, such that the above is positive definite. Equivalently, using $\gamma_{\min}$ and $\gamma_{\max}$ to denote the minimum and maximum eigenvalue respectively,

$$\gamma_{\min} \begin{bmatrix} D^2 - P_{11}^* D^2 P_{11} & -P_{11}^* D^2 P_{12} \\ -P_{12}^* D^2 P_{11} & I - P_{12}^* D^2 P_{12} \end{bmatrix} > 0$$

$$\iff \quad \gamma_{\max} \begin{bmatrix} P_{11}^* D^2 P_{11} - D^2 & P_{11}^* D^2 P_{12} \\ P_{12}^* D^2 P_{11} & P_{12}^* D^2 P_{12} - I \end{bmatrix} < 0.$$

Premultiplying and postmultiplying by a symmetric matrix does not change the inertia of the above matrix. This leads to the following equivalent condition.

$$\gamma_{\max} \begin{bmatrix} D^{-1} P_{11}^* D^2 P_{11} D^{-1} & D^{-1} P_{11}^* D^2 P_{12} \\ P_{12}^* D^2 P_{11} D^{-1} & P_{12}^* D^2 P_{12} \end{bmatrix} < 1.$$

$$\iff \quad \sigma_{\max}[D P_{11} D^{-1} \; D P_{12}] < 1 \qquad (9)$$

This condition is reminiscent of that for the upper bound to μ: Equation2. This is formalized in the following theorem.

Theorem 9 *There exist $D \in \mathcal{D}$ such that $\sigma_{\max}(D P_{11} D^{-1}) < 1$ iff there exist $\lambda_i > 0$, such that $H[L(x,\lambda)]$ is positive definite.*

Proof of Theorem 9: This follows by noting that there is a degree of freedom in the D required to make $\sigma_{\max}(D P_{11} D^{-1}) < 1$ which can be used to make the contribution of $D P_{12}$ arbitrarily small. ∎

Packard[4] points out that there is no loss of generality in restricting D to be positive and real. This translates exactly into the same requirements on λ.

4.4 The Properties of $L_e(y,\lambda)$

Now define Λ_e as the region of λ space in which the Hessian of the Lagrangian: $L_e(y,\lambda)$ is positive definite and $\lambda_i > 0$. It will be shown for $\lambda \in \Lambda_e$, $h_e(\lambda)$ is finite and concave.

Lemma 10 $\Lambda \subseteq \Lambda_e$

Proof of Lemma 10 Follows by noting that every for every $\lambda \in \Lambda$, Vy can be substituted for x. The Lagrangian remains positive definite and so $\lambda \in \Lambda_e$. ∎

It can now be shown that the existence of $D \in \mathcal{D}$ giving a sufficient condition for robust stability is sufficient to guarantee the existence of λ such that $L_e(y,\lambda)$ is positive definite. This is expressed in the following theorem.

Theorem 11 *If there exist $D \in \mathcal{D}$ such that $\sigma_{\max}(D P_{11} D^{-1}) < 1$ then there exist $\lambda_i > 0$ such that $H[L_e(y,\lambda)]$ is positive definite.*

Proof of Theorem 11: This follows immediately from Theorem 9 and Lemma 10. ∎

4.5 Properties of the Sets Λ and Λ_e

The previous sections showed that for models which meet the sufficient condition for robust stability, the sets Λ and Λ_e are not empty. This section gives some results regarding these sets.

Packard[4] points out that the region of $\mathcal{D}$ such that the $D \in \mathcal{D}$, such that $\sigma_{\max}(DMD^{-1}) < 1$ is a convex region. It it not surprising then that Λ and Λ_e are also convex regions.

Theorem 12 *Λ_e is a convex.*

Proof of Theorem 12 This follows from the fact that the Lagrange multipliers enter the Lagrangian linearly and the Hessian of any convex combination will then be equal to the convex combination of the Hessians. ∎

Corollary 13 *Λ is convex.*

By Lemma 5, $L_e(y,\lambda)$ is a strictly convex function of y for $\lambda \in \Lambda_e$. It is in fact a positive definite quadratic. Now for every $\lambda \in \Lambda_e$, $h_e(\lambda)$ is readily calculated as $L_e(y,\lambda)$ has a unique global minimum.

The above shows that the Lagrangian is well behaved for λ contained within Λ_e. It now remains to examine the properties on the boundary. Two types of boundary are possible for the set Λ (and Λ_e). Denote by $\partial\Lambda$ (respectively $\partial\Lambda_e$), the boundary of Λ such that the minimum eigenvalue of $H[L(x,\lambda)]$ is zero. As the multipliers are constrained to be greater than zero, the hyperplanes defined by each component of λ being equal to zero also bound Λ (and Λ_e). Denote this boundaries by $\partial_0\Lambda$ (and $\partial_0\Lambda_e$). Any λ such that $\lambda_i = 0$ and $H[L(x,\lambda)]$ has a zero eigenvalue is defined to be on $\partial\Lambda$ and not on $\partial_0\Lambda$.

For $\lambda_i > 0$ and $\lambda \notin \Lambda_e \cup \partial\Lambda_e \cup \partial_0\Lambda_e$, $H[L_e(y,\lambda)]$ has a negative eigenvalue and $h_e(\lambda) = -\infty$.

4.6 Properties of the Dual Function

Theorem 14 *For all $\lambda \in \Lambda_e$, the dual function $h_e(\lambda)$ is concave.*

Proof of Theorem 14: This follows from Theorem 6 and Theorem 12. ∎

5 Solving the Validation Problem

The properties of the space Λ_e and the dual function $h_e(\lambda)$ can give a solution to the model validation problem by finding a Kuhn Tucker saddlepoint. It will be shown here that if the solution to the dual problem does not occur on the boundary $\partial\Lambda_e$ then it also solves the primal problem and hence the model validation problem. In the case where the maximization leads to the boundary it may not be possible to find a solution to the primal problem. In these cases it is possible to bound the solution. This is examined in Section 5.2.

5.1 Solution via the Dual Problem

Theorem 15

$$\text{If} \quad \max_{\lambda} h_e(\lambda) = L_e(\hat{y}, \hat{\lambda}) \quad \text{with } \hat{\lambda} \notin \partial\Lambda_e,$$

then $\hat{x} = x_0 + V\hat{y}$ *solves Problem 3.*

Proof of Theorem 15: Note that the constraints are simply the gradient of the Lagrangian with respect to λ. If $\hat{\lambda} \in \Lambda_e$ the constraints are all zero. If $\hat{\lambda} \in \partial_0\Lambda_e$ then the nonzero constraints are negative and correspond to zero components in $\hat{\lambda}$. ∎

5.2 Bounds on a Solution

A detailed examination of the dual function on the boundary is postponed until Section 6. The result of interest here is that it is always possible to solve the single uncertainty block case (Section 6.4). This allows the calculation of an lower bound on the minimum $\|w\|$ required to account for the observed datum.

If the structure of the uncertainty is ignored a simpler problem can be posed. This is equivalent to simply setting $f = 1$. Denote the solution to this as Δ_l and w_l.

Theorem 16 *There exist no* w *with* $\|w\| \leq \|w_l\|$ *solving the model validation problem (Problem 3)*

Proof of Theorem 16: This follows from the fact that every Δ structure with $f > 1$ is contained within the set of unstructured $(f = 1)$ Δ. ∎

Solution via the dual problem involves finding x such that the constraints $g_{ei}(x)$ are met exactly $(g_{ei}(x) = 0)$ or the constraint is inactive at the solution $(\lambda_i = 0)$. In the case where the maximization of the dual function leads to the boundary it will be possible to find x, denoted by x_u such that $g_{ei}(x) < 0$ and $\lambda_i > 0$ for some i. In this case x_u meets the constraints but it is not possible to say that it achieves the minimum of the primal problem. x_u can then be considered as an upper bound on x solving the model validation problem.

The Lagrange multiplier approach leads to the possibility of a gap in the solution of the model validation problem. In the case where a solution is found on the interior of Λ_e or on $\partial_0\Lambda_e$ the minimum $\|w\|$ meeting the constraints is found. In the case where $f = 1$ and $\lambda \in \partial\Lambda_e$, the minimum $\|w\|$ is still found. This allows the model validation question to be answered in a yes/no manner.

If $f \geq 3$ and $\lambda \in \partial\Lambda_e$ maximizes the dual function, it is possible to have a gap in the answer to the model validation question. If $\|w_l\| > 1$ then the answer is no (i.e. there is no vector x meeting all of the constraints of the model). Alternatively if $\|w_u\| < 1$ the answer is yes; in fact x_u is a vector meeting all of the constraints of the model. However if $\|w_l\| < 1$ and $\|w_u\| > 1$, there is no conclusive statement that can be made with regard to the model validation question without further computation.

This gap is analogous to that between μ and its upper bound $\sigma_{\max}(DMD^{-1})$. This gap exists for four or more uncertainty blocks. For three uncertainty blocks the gap also exists in the model validation problem. To consider any equivalence between the number of uncertainty blocks in each problem the equality constraint must be counted as an uncertainty block. Section 6.5 provides an example to show that the gap does indeed exist for $f = 3$. This example is the same one constructed by Doyle (refer to Packard[4]) and used to demonstrate the gap in the calculation of μ. The $f = 2$ case for model validation is still under investigation.

6 The Dual Function on $\partial\Lambda_e$

For λ close to the boundary $\partial\Lambda_e$, the Hessian of the Lagrangian has an eigenvalue close to zero. In general the y achieving the minimum of $L_e(y, \lambda)$ will become larger in norm as λ approaches the boundary. This is not always the case and the next section will give necessary and sufficient conditions under which $h_e(\lambda)$ is finite on the boundary.

6.1 When is $h_e(\lambda)$ finite on $\partial\Lambda_e$

For $\lambda \in \partial\Lambda_e$ the Hessian of the Lagrangian has a zero eigenvalue. Whether or not the minimum of $L_e(y, \lambda)$ for $\lambda \in \partial\Lambda_e$ is finite depends on the associated eigenvector and the linear term of the Lagrangian: $C_e(\lambda)$. For $\hat{\lambda} \in \partial\Lambda_e$, define

$$\hat{\lambda}\hat{A} = \sum_{i=1}^{f} \hat{\lambda}_i A_i.$$

Theorem 17 $h_e(\hat{\lambda})$ *is finite for* $\hat{\lambda} \in \partial\Lambda_e$ *iff for every* $y_0 \in \text{Ker}[V^*(B + \hat{\lambda}\hat{A})V]$, $Re\{C_e(\hat{\lambda})^* y_0\} = 0$.

Proof of Theorem 17: The proof involves showing that if any vector in the kernel contributes to the linear term the dual function can be driven to $-\infty$. ∎

The above gives necessary and sufficient conditions for the value of $h_e(\lambda)$ to be finite on the boundary of Λ_e. If the maximum of $h_e(\lambda)$ occurs for $\lambda \in \partial\Lambda_e$ then it is not necessarily true that the gradients, and consequently the constraints, are zero. They may in fact be positive. Consequently, finding the maximum is no longer guaranteed to find a saddlepoint.

6.2 An Intuitive Description

The condition that $y_0 \in \text{Ker}[V^*(B + \hat{\lambda}\hat{A})V]$ can be reformulated as follows. λ defines a scaling on P. Denote this by P_s and examine the partition $\tilde{P}_s$ defined as follows.

$$\tilde{P}_s = \begin{bmatrix} DP_{11}D^{-1} & DP_{12} \\ P_{21}D^{-1} & P_{12} \end{bmatrix}.$$

The condition is then equivalent to the existence of x_s such that

$$\|x_s\| = \|\tilde{P}_s x_s\| \quad \text{and} \quad [0\ I]\tilde{P}_s x_s = 0.$$

Now consider the condition $\text{Re}\{C_e(\hat{\lambda})^* y_0\} = 0$. This can be shown to be equivalent to

$$\text{Re}\{([I\ 0]\tilde{P}_s x_s)^*(P_{s13}u)\} = 0 \tag{10}$$

A sufficient condition for $h_e(\hat{\lambda})$ finite on the boundary $\partial\Lambda_e$, is that there exist a subspace, defined by x_s, such that:

- The subspace corresponds to a singular value of one for the scaled plant $\tilde{P}_s$. Note that this is a necessary condition for the constraints of the model validation problem to be met exactly on $\partial\Lambda_e$.
- The subspace produces a zero output at e and is orthogonal to the contribution of the input u.

The last item can be loosely interpreted as saying that neither the input u nor the output e provides any information on a subspace of vectors achieving the uncertainty block constraints exactly. It is interesting to note that this is always the case when $u = 0$. If the μ analysis problem is formulated in this framework, as it is in Section 6.5, this will always be the case.

6.3 Finding a Kuhn Tucker Saddlepoint on $\partial\Lambda_e$

The previous section demonstrated that on the boundary there exists a subspace which does not affect the output e and is unaffected by the input u. This gives a degree of freedom which can sometimes be exploited to solve the problem on the boundary. This is investigated in this section.

$$\text{If} \quad \max_{\lambda} h_e(\lambda) = L_e(\hat{y}, \hat{\lambda})$$

with $\hat{\lambda} \in \partial\Lambda_e$, then for every vector y_0 in the kernel of $V^*(B+\hat{\lambda}\hat{A})V$,

$$L_e(\hat{y} + y_0, \hat{\lambda}) = L_e(\hat{y}, \hat{\lambda}).$$

The choice of y_0 gives an additional degree of freedom in finding a Kuhn Tucker saddlepoint. Consider the constraint $g_{ei}(y)$ with the addition of y_0.

$$g_{ei}(y+y_0) = y_0^*V^*A_iVy_0 + 2\,\mathrm{Re}\{y_0^*V^*A_i(x_0+Vy)\} -2\,\mathrm{Re}\left\{y_0^*V^*\begin{bmatrix}P_{11}^*\\P_{12}^*\end{bmatrix}T_iP_{13}u\right\} + g_{ei}(y). \quad (11)$$

This in effect defines another problem. Does there exist $y_0 \in \mathrm{Ker}[V^*(B+\hat{\lambda}\hat{A})V]$, such that for $i = 1,\ldots,f$, Equation 11 can be made equal to zero. The following section will show that this problem can be solved in the single uncertainty block case.

6.4 The Single Uncertainty Block Case

λ is a scalar and if $\sigma_{\max}(P_{11}) < 1$ then Λ_e us an open interval on the real line: $(0,\partial\Lambda_e)$. It will be shown that in this case it is always possible to solve the boundary problem of Equation 11.

Consider $\hat{y}$, corresponding to a solution of the Lagrangian on the boundary:

$$h_e(\hat{\lambda}) = \min_y L_e(y,\hat{\lambda}) = L_e(\hat{y},\hat{\lambda}), \quad \hat{\lambda} = \partial\Lambda_e.$$

Partition $\hat{y}$ as $\hat{y} = \hat{y}_1 + y_0$, $y_0 \in \mathrm{Ker}[V^*(B+\lambda A)V]$. Then

$$L_e(\hat{y}_1 + \alpha y_0, \hat{\lambda}) = L_e(\hat{y}_1, \hat{\lambda}), \quad \alpha \in \mathcal{C}.$$

The following two lemmas give the required properties for a solution of Equation 11.

Lemma 18 $g(\hat{y}_1) \geq 0$.

Proof of Lemma 18: The proof involves showing that if this were not true the maximum of the dual function would exist inside Λ_e. ■

Lemma 19 $y_0^*V^*AVy_0 < 0$.

Proof of Lemma 19: On the boundary $y_0^*V^*(B+\lambda A)Vy_0 = 0$. As $y_0V^*BVy_0 \geq 0$, $y_0^*V^*AVy_0 \leq 0$. Equality cannot occur; if it did then there would exist a zero eigenvalue for all λ, making Λ_e empty and contradicting $\sigma_{\max}(P_{11}) < 1$. ■

Now, for $\hat{y}$ minimizing the Lagrangian on the boundary, there is an extra degree of freedom. For $\alpha \in \mathcal{C}$, $g(\hat{y}) = g(\hat{y}_1 + \alpha y_0)$ can be made equal to zero. Equation 11 is simply a quadratic in α. By Lemma 19 it is negative definite, and by Lemma 18 for $\alpha = 0$ Equation 11 has a non negative value. It therefore has a root for $\alpha \in \mathcal{R}$. All three conditions of Theorem 7 are now satisfied and $\hat{y}$ is therefore a solution of the primal problem.

6.5 An Application to a μ Problem

Fan and Tits[6] have applied Lagrange multiplier methods to the calculation of μ. However for a scalar output signal e, the model validation problem and μ can be made equivalent.

If there exists x such that $\|\,[0\ I]\,\tilde{P}x\,\| = 1$, and $\|\,[0\ I]\,x\,\| < 1$, and for $i = 1,\ldots,f$, $\|\,[Q_i\ 0]\,x\,\| = \|\,[Q_i\ 0]\,\tilde{P}x\,\|$, then $\mu > 1$.

In the scalar output case it suffices to choose $e = 1$ as the entire problem can be multiplied by a complex constant of unity magnitude without affecting μ. If $u = 0$ then the generic interconnection structure reduces to that for the μ formulation.

Consider the four block counterexample[4] for which

$$\inf_{D\in\mathcal{D}} \sigma_{\max}(D\tilde{P}D^{-1}) = 1$$

but $\mu < 1$. Then the minimum $\|w\| > 1$ for w meeting the constraints. The model validation techniques will be able to find the minimum $\|w\|$. If it could be found then it would be possible to calculate μ for this problem by a iterative scaling technique. This is hardly surprising, given the demonstrated relationship between the Lagrange multipliers and the D used in the calculation of the upper bound of μ.

Lemma 20 *For scalar $e = 1$, $u = 0$, and a model P such that*

$$\inf_{D\in\mathcal{D}} \sigma_{\max}\left(\begin{bmatrix}D&0\\0&I\end{bmatrix}\tilde{P}\begin{bmatrix}D&0\\0&I\end{bmatrix}^{-1}\right) \geq 1, \quad \max_{\lambda} h_e(\lambda) \leq 1.$$

Proof of Lemma 20: This proof relies on the fact that $\{x_0|V\}$ spans the space in which x lies. $H[L_e(y,\lambda)]$ can then be reformulated as a singular value constraint. ■

Now by Lemma 20 $h_e(\lambda)$ is bounded by one.

$$h_e(\lambda) = \|w\|^2 + \sum_{i=1}^{f}\lambda_i g_{ei}(x) \leq 1.$$

But $\|w\| > 1$. Consequently the constraints can never be met exactly and it is not possible to answer the question is $\mu(\tilde{P}) < 1$.

7 Acknowledgements

The authors wish to thank Andy Packard and Kemin Zhou for many helpful discussions and Harold Stalford for his careful reading of a more complete version of this paper. Errors induced by the severe truncation are ours alone.

References

[1] J.C.Willems, "From time series to linear system, Part I: Finite dimensional linear time invariant linear systems; PartII: Exact modeling; and Part III: Approximate modeling.," *Automatica*, vol. 22, no. 5, pp. 561–580, vol. 22, no.6, pp. 675–694, & vol. 23, no. 1, 1986 & 1987.

[2] J. M. Krause and G. Stein, "A general adaptive control structure," in *Proc. IEEE Control Decision Conf.*, pp. 761–768, 1988.

[3] J. Doyle, "Analysis of feedback systems with structured uncertainties," *IEE Proceedings, Part D*, vol. 133, pp. 45–56, Mar. 1982.

[4] A. K. Packard, *What's new with μ*. PhD thesis, University of California, Berkeley, 1988.

[5] D. A. Wismer and R. Chattergy, *Introduction to Nonlinear Optimization, A Problem Solving Approach*. System Science and Engineering Series, New York: North Holland, 1978.

[6] M. K. H. Fan and A. L. Tits, "m-form numerical range and the computation of the structured singular value," *IEEE Trans. Auto. Control*, vol. AC-33, pp. 284–289, 1988.