International Journal of Enhanced Research in Science, Technology & Engineering
ISSN: 2319-7463, Vol. 5 Issue 7, July-2016

Generalized Model Error Estimators for Nonlinear Systems

Swetha Amit[1], Raol, J. R.[2]

[1]Asst. Professor, Dept. of Telecommunication, MSRIT, and Research Scholar, Jain University, Bangalore
[2]Prof. Emeritus, MSRIT, Bangalore

ABSTRACT

The generalized model error estimators for continuous time and discrete time nonlinear systems are presented. Also, H infinity based semi-robust or adaptive estimators are suggested. The model error estimators are primarily based on the classical least squares criterion, and then the cost function is extended to include the energy term of deterministic discrepancy that is also called model error. The recursive solution to the ensuing two point boundary value problem is obtained by the method of invariant embedding (IE). The continuous time generalized IE based algorithm is illustrated with a numerical example implemented in MATLAB.

Keyword: Model error, Pontryagin's conditions, invariant embedding, generalized estimators, adaptive IE estimators.

1. INTRODUCTION

In several real life practical situations accurate identification of nonlinear terms (parameters) in the model of a dynamic system is required. Interestingly enough, traditionally used KF/EKF (Kalman filter) cannot determine the deficiency or discrepancy in the model of the system used in the filter, since it presupposes availability of accurate statespace model. Assume a situation wherein we are given the measurements from a nonlinear dynamic system and we want to determine the state estimates. In this case, we use extended KF and we need to have the knowledge of the nonlinear function 'f' and 'h'. Any discrepancy in the model will cause model errors that will tend to create mismatch of the estimated states with the true states of the system. There might be some problems [1-3]: i) deviation from the Gaussian assumption might degrade the performance of the algorithm, and ii) the filtering algorithm is dependent on the covariance matrix P of the state estimation error, since this is used for computation of Kalman gain K. These limitations of the KF can be overcome largely by using the method based on principle of model error [1-3]. This approach not only estimates the states of the dynamic system from its measurements, but also the model discrepancy as a time history. The point is that we can use the known (deficient or linear) model in the state estimation procedure, and determine the deterministic discrepancy of the model, using the measurements in the model error estimation procedure. Once the discrepancy time history is available, one can fit another model to it and estimate its parameters using a regression-LS method. Then combination of the previously used model in the state estimation procedure and the new additional model would yield the accurate model of the underlying (nonlinear) dynamic system, which has in fact generated the data.

This approach would be very useful in modelling of the large flexible structures, robotics and many aerospace dynamic systems, which usually exhibit nonlinear behaviour. Often these systems are linearized leading to approximate linear models with useful range of operation but with limited validity at some far away points from the local linearization points. Such linear systems can be easily analysed using simple tools of linear system theory. System identification work generally restricted to such linear and linearized models can lead to modal analysis of the nonlinear systems. However, the linearized models will have limited range of validity for nonlinear practical data, because certain terms are neglected, in the process of linearization and approximation. This will produce inaccurate results, and these linearized models will not be able to predict certain behavioural aspects of the system, like drift.

The approach presented in this paper, would produce/predict accurate state trajectory, even in the presence of deficient/inaccurate model and additionally identify the unknown model (form) as well as its parameters. The method of model error essentially results into a batch estimation procedure, because it is a two-step process. However, realtime solution can be obtained using the method of invariant embedding (IE), in conjunction with real time least squares estimation method (RLS). The method of IE is discussed for continuous as well as discrete time systems. In essence we derive generalizes estimators based on the classical LS criterion and obtain the generalized IE estimators and then specialize these for obtaining the conventional IE estimators. Then we upgrade the classical LS criterion with the robust H infinity norm based criterion and obtain, from these generalized estimators, the semi-robust estimators. We call these as robust because these estimators are based on H infinity norm criteria, and as semi-robust, since these will satisfy only certain theoretical condition on the cost function, and may not satisfy the full robustness condition. Hence, we present novel results on model error estimation in the joint setting of IE and HI (H infinity), and the generalized IE based estimators, which might lead to strictly robust estimators in the joint setting of IE and HI. Performance is illustrated with an example.

2. PHILOSOPHY OF MODEL ERROR AND PONTRYAGIN'S CONDITIONS

Our main aim is to determine the model error, or the so called deterministic discrepancy based on the available noisy measurements for a given nonlinear dynamic system; and it is assumed that the experimental/real data are from a nonlinear system, however, we fit only a primarily known model that might be deficient, i.e. this postulated model is not true model. Let the mathematical description of the nonlinear system be given as

$$\dot{x} = f(x(t), u(t), t) + d(t) \tag{1}$$

The un-modelled (not modelled) disturbance is represented by d(t), which is assumed to be piecewise continuous. This is not the process noise term as in the/of the KF theory. Hence, like the output error method, this approach cannot as such handle the true process noise. However, the aim here is different. In control theory, the term d(t) would represent a control force or input, u(t), which is determined using an optimisation method by minimizing the cost function

$$J = \sum_{k=1}^{N} [z(k) - h(\hat{x}(k), k)]^T R^{-1} [z(k) - h(\hat{x}(k), k)] + \int_{t_0}^{t_f} d^T(t) Q d(t) dt \tag{2}$$

It is assumed that $E\{v(k)\} = 0$; $E\{v(k)v^T(k)\} = R(k)$, which is assumed to be known. Here, 'h' is the measurement model. The weighting matrix Q plays an important role as a tuning device for the estimator. One natural way to arrive at Q is to choose it such that the equality is satisfied

$$R(k) = [z(k) - h(\hat{x}(k), k)][z(k) - h(\hat{x}(k), k)]^T \tag{3}$$

In (3), R(.) is the postulated covariance matrix of the measurement noise and the right hand side is the measurement covariance matrix computed using the difference between the actual measurements and the predicted measurements, obtained from the estimator. This equality is called the covariance constraint. The main merit of the present approach is that it obtains state estimates in the presence of un-modelled effects as well as accurate estimates of these effects. Except on R, no statistical assumptions are required. The criteria used for estimation are based on LS and one can after some transformations obtain recursive estimator like KF. In the process, the model itself is improved, since this estimate of the un-modelled effects can be further modelled and the new model can be obtained: accurate model (of the original system) = deficient model + model fitted to the discrepancy (i.e. un-modelled effects). The problem of determination of the model error, deficiency or discrepancy is via minimization of the cost functional (2) which gives rise to the so called two point boundary value problem (TPBVP) [1-4]. The dynamic system is given

$$\dot{x} = f(x(t), u(t), t); \qquad x(t_0) = x_0 \tag{4}$$

Then, define a composite performance index as

$$J = \phi(x(t_f), t_f) + \int_{t_0}^{t_f} \psi(x(\tau), u(\tau), \tau) d\tau \tag{5}$$

In (5), the first term is the cost penalty on the final value of the state $x(t_f)$, and the term $\psi(.)$ is the cost penalty governing the deviation of $x(t)$ and $u(t)$ (in general a control input) from their desired timehistories. The aim is to

determine this input $u(t)$, in the interval $t_0 \le t \le t_f$, such that the performance index J is minimized, subject to the constraint of (4), which states that the state should follow integration of (4) with the input thus determined.
We use the concept of Lagrange multiplier to handle the constraint within the functional J

$$J_a = \phi(x(t_f), t_f) + \int_{t_0}^{t_f} [\psi(x(\tau), u(\tau), \tau) + \lambda^T(-f(x(\tau), u(\tau), \tau) + \dot{x})] d\tau \tag{6}$$

In (6), λ is the Lagrange multiplier and it facilitates the inclusion of the condition (4), which is the constraint on the state of the dynamical system, the point is that in the process of determination of $u(t)$ by minimization of J_a, the condition of (4) should not be violated. In the absence of the constraint, the u might be such that the system does not follow the dynamics, which otherwise it is supposed to follow/obey. The Lagrange multipliers are known as adjoint variables or co-states. Since, in the sequel, we will have to solve the equations for the Lagrange multipliers, simultaneously with those of state equations, we prefer to use the 'co-state' terminology. If the condition of (4) is strictly satisfied, then essentially (5) and (6) are identical and the same, but now the dynamic constraint is taken care. Equation (6) can be rewritten as

$$J_a = \varphi(x(t_f), t_f) + \int_{t_0}^{t_f} [H_a(x(\tau), u(\tau), \tau) - \dot{\lambda}^T(\tau) x(\tau)] d\tau + (\lambda^T x)_{t_f} - (\lambda^T x)_{t_0} \tag{7}$$

Here, $H_a = \psi(x(\tau), u(\tau), \tau) - \lambda^T(\tau) f(x(\tau), u(\tau), \tau)$ (8)

In (7), the last three terms related to λ are obtained from the term $\int_{t_0}^{t_f} \lambda^T \dot{x} d\tau$ of (6) by integrated by parts, and the remaining two terms of (6) are clubbed and we get (8) to define it as Hamiltonian. From (7), we obtain, by using the concept of differentials

$$\delta J_a = 0 = \left((\frac{\partial \varphi}{\partial x})^T \delta x \right)_{t_f} + \lambda^T \delta x \Big|_{t_f} - \lambda^T \delta x \Big|_{t_0} + \int_{t_0}^{t_f} \left[\left((\frac{\partial H}{\partial x})^T - \dot{\lambda}^T \right) \delta x + (\frac{\partial H}{\partial u})^T \delta u \right] d\tau \tag{9}$$

From (9), the socalled Pontryagin's necessary conditions are obtained as

$$\lambda(t_f) = -\frac{\partial \varphi}{\partial x}\Big|_{t_f} \tag{10}$$

$$\frac{\partial H}{\partial x} = \dot{\lambda} \tag{11}$$

and $\dfrac{\partial H}{\partial u} = 0$ (12)

Here, $\delta x(t_0) = 0$, assuming that the initial conditions $x(t_0)$ are independent of $u(t)$. Equation (10) is called the transversality condition. In the TPBV problem the boundary condition for state is specified at t_0 and for the co-state, λ it is specified at t_f (10), then from (8) and (11), we obtain

$$\dot{\lambda} = \left(\frac{\partial H}{\partial x} \right) = -\left(\frac{\partial f}{\partial x} \right)^T \lambda + \left(\frac{\partial \psi}{\partial x} \right) \tag{13}$$

$$\frac{\partial H}{\partial u} = 0 = -\left(\frac{\partial f}{\partial u} \right)^T \lambda + \left(\frac{\partial \psi}{\partial u} \right) \tag{14}$$

Because of the split boundary conditions, the problem as such is relatively to solve. One method to solve the TPBVP is to start with guesstimate on $\lambda(t_0)$ and use $x(t_0)$ to integrate forward to the final time t_f; then verify the boundary condition $\lambda(t_f) = -\left.\frac{\partial\phi}{\partial x}\right|_{t_f}^{T}$. If the condition is not satisfied, then iterate once again with new $\lambda(t_0)$ (or start from the end and traverse backward and verify other conditions) and so on until the convergence of the algorithm is obtained. We discuss the method of invariant embedding that obtains the recursive solution to the TPBV problem,

3. INVARIANT EMBEDDING APPROACH

In many cases, it is useful to analyse a general process/solution of which our original problem is one particular case [4,5], and the method of invariant embedding belongs to this category. It means that the particular solution we are seeking is embedded in the general class and after the general solution is obtained, our particular solution can be obtained by using the special conditions, which we have kept invariant, in final analysis. Let the resultant equations from the TPBVP be given as (1) and (13), in general

$$\dot{x} = \Phi(x(t), \lambda(t), t) \tag{15}$$

$$\dot{\lambda} = \Psi(x(t), \lambda(t), t) \tag{16}$$

We see that the dependencies for Φ and Ψ on $x(t)$ and $\lambda(t)$ arise from the form of (1), (13) and (14), hence, here, we have a general TPBVP with associated boundary conditions as: $\lambda(0) = a$ and $\lambda(t_f) = b$. Now, though the terminal condition $\lambda(t_f) = b$ and time are fixed, we consider them as free variables, this makes the problem more general, which anyway includes our specific problem. We know from the nature of the TPBVP that the terminal state $x(t_f)$ depends on t_f and $\lambda(t_f)$. Therefore, this dependency can be represented as

$$x(t_f) = r(c, t_f) = r(\lambda(t_f), t_f) \tag{17}$$

with $t_f \rightarrow t_f + \Delta t$, we obtain by neglecting higher order terms

$$\lambda(t_f + \Delta t) = \lambda(t_f) + \dot{\lambda}(t_f)\Delta t = c + \Delta c \tag{18}$$

We also get, using (16) in (18)

$$c + \Delta c = c + \Psi(x(t_f), \lambda(t_f), t_f)\Delta t \tag{19}$$

and therefore, we get

$$\Delta c = \Psi(r, c, t_f)\Delta t \tag{20}$$

Additionally, we get, like (18)

$$x(t_f + \Delta t) = x(t_f) + \dot{x}(t_f)\Delta t = r(c + \Delta c, t_f + \Delta t) \tag{21}$$

and hence, using (15) in (21), we get

$$\begin{aligned} r(c + \Delta c, t_f + \Delta t) &= r(c, t_f) + \Phi(x(t_f), \lambda(t_f), t_f)\Delta t \\ &= r(c, t_f) + \Phi(r, c, t_f)\Delta t \end{aligned} \tag{22}$$

Using Taylor's series, we get:

$$r(c + \Delta c, t_f + \Delta t) = r(c, t_f) + \frac{\partial r}{\partial c}\Delta c + \frac{\partial r}{\partial t_f}\Delta t \tag{23}$$

Comparing (22) and (23), we get

$$\frac{\partial r}{\partial t_f}\Delta t + \frac{\partial r}{\partial c}\Delta c = \Phi(r,c,t_f)\Delta t \tag{24}$$

or, using (20) in (24), we obtain

$$\frac{\partial r}{\partial t_f}\Delta t + \frac{\partial r}{\partial c}\Psi(r,c,t_f)\Delta t = \Phi(r,c,t_f)\Delta t \tag{25}$$

Equation (25) simplifies to

$$\frac{\partial r}{\partial t_f} + \frac{\partial r}{\partial c}\Psi(r,c,t_f) = \Phi(r,c,t_f) \tag{26}$$

We see that (26) links the variation of the terminal condition $x(t_f) = r(c,t_f)$ to the state and co-state differential functions, (15) and (16), and now in order to find an optimal estimate $\hat{x}(t_f)$, we need to determine $r(b,t_f)$ as

$$\hat{x}(t_f) = r(b,t_f) \tag{27}$$

Equation (26) can be transformed to an initial value problem by using approximation

$$r(c,t_f) = S(t_f)c + \hat{x}(t_f) \tag{28}$$

Substituting (28) in (26), we get

$$\frac{dS(t_f)}{dt_f}c + \frac{d\hat{x}(t_f)}{dt_f} + S(t_f)\,\Psi(r,c,t_f) = \Phi(r,c,t_f) \tag{29}$$

Next, expanding Φ and Ψ about $\Phi(\hat{x},b,t_f)$ and $\Psi(\hat{x},b,t_f)$, we obtain

$$\begin{aligned}\Phi(r,c,t_f) &= \Phi(\hat{x},b,t_f) + \Phi_{\hat{x}}(\hat{x},b,t_f)(r(c,t_f) - \hat{x}(t_f)) \\ &= \Phi(\hat{x},b,t_f) + \Phi_{\hat{x}}(\hat{x},b,t_f)S(t_f)c\end{aligned} \tag{30}$$

and $$\Psi(r,c,t_f) = \Psi(\hat{x},b,t_f) + \Psi_{\hat{x}}(\hat{x},b,t_f)S(t_f)c \tag{31}$$

Utilizing expressions of (30) and (31), in (29), we obtain a composite state estimation equation

$$\begin{aligned}&\frac{dS(t_f)}{dt_f}c + \frac{d\hat{x}(t_f)}{dt_f} + S(t_f)[\Psi(\hat{x},b,t_f) + \Psi_{\hat{x}}(\hat{x},b,t_f)S(t_f)c] \\ &\qquad = \Phi(\hat{x},b,t_f) + \Phi_{\hat{x}}(\hat{x},b,t_f)S(t_f)c\end{aligned} \tag{32}$$

The equation (32) is in essence a sequential state estimation algorithm but a composite one involving $\hat{x}$, $S(t_f)$, and c, hence it should be separated by substituting the specific expressions for Φ and Ψ in (32). We can do this when we specify the dynamic systems for which we need to obtain the estimators. This we do in next section after arriving at TPBVP for a specific problem at hand, and then using (32).

4. GENERALIZED CONTINUOUS TIME ALGORITHM

Let nonlinear dynamic system be represented by

$$\dot{x} = f(x(t),t) + d(t) \tag{33}$$

$$z(t) = Hx(t) + v(t) \tag{34}$$

We form the generalized cost functional based on the LS principle, and the deterministic discrepancy, i.e. model error energy, since we want to estimate the states of the system, and the model error, we assume that we are using only the postulated (deficient) model to start with

$$J = \int_{t_0}^{t_f} [(z(t) - Hx(t))^T R^{-1}(z(t) - Hx(t))] + \alpha\ (d^T(t)\ Q\ d(t))]\ dt \tag{35}$$

In (35) d(t) is the model discrepancy to be estimated simultaneously with x(t), and R(t) is the spectral density matrix of noise (the covariance). Also, in the second term we have introduced an arbitrary parameter α that would generalize the model error estimator. We reformulate J by using Lagrange multiplier in order to incorporate the constraint of the system dynamics

$$J_a = \int_{t_0}^{t_f} [(z(t) - Hx(t))^T R^{-1}(z(t) - Hx(t)) + \alpha(d^T(t)\ Q\ d(t)) + \lambda^T(\dot{x}(t) - f(x(t),t) - d(t))]\ dt \tag{36}$$

Comparing with (7) and (8), we get

$$\begin{aligned} H_a &= (z(t) - Hx(t))^T R^{-1}(z(t) - Hx(t)) + \alpha d^T(t)Qd(t) - \lambda^T(f(x(t),t) + d(t)) \\ &= \psi - \lambda^T f_m(x(t), d(t), t) \end{aligned} \tag{37}$$

In (37), the first two terms define the function $\psi = (z(t) - Hx(t))^T R^{-1}(z(t) - Hx(t)) + \alpha d^T(t)Qd(t)$. By utilizing this function and the Hamiltonian, and applying the Pontryagin's conditions, we obtain

$$\dot{\lambda} = \frac{\partial H}{\partial x} = \frac{\partial \psi}{\partial x} - \lambda^T \frac{\partial f_m}{\partial x} \tag{38}$$

$$\dot{\lambda} = \left(\frac{\partial \psi}{\partial x}\right) - \left(\frac{\partial f_m}{\partial x}\right)^T \lambda = -f_{\hat{x}}^T \lambda - 2H^T R^{-1}(z(t) - Hx(t)) \tag{39}$$

and $$0 = \frac{\partial H}{\partial d} = 2\alpha Q d - \lambda \rightarrow \qquad d = \frac{1}{2}\alpha^{-1}Q^{-1}\lambda \tag{40}$$

Thus our two-point boundary value problem is

$$\begin{aligned} \dot{x} &= f(x(t),t) + d(t) \\ \dot{\lambda} &= -f_{\hat{x}}^T \lambda - 2H^T R^{-1}(z(t) - Hx(t)) \\ d &= \frac{1}{2}\alpha^{-1}Q^{-1}\lambda \end{aligned} \tag{41}$$

Now comparing with (15) and (16), we obtain

$$\Phi(x(t), \lambda(t), t) = f(x(t),t) + d(t) \tag{42}$$

and $$\Psi(x(t), \lambda(t), t) = -f_{\hat{x}}^T \lambda - 2H^T R^{-1}(z(t) - Hx(t)) \tag{43}$$

We also have

$$\Psi_{\hat{x}} = 2H^T R^{-1} H - \left[\frac{\partial}{\partial \hat{x}}(f_{\hat{x}}^T)\lambda\right] \tag{44}$$

and $$\Phi_{\hat{x}} = f_{\hat{x}} \tag{45}$$

Substituting (42) to (45) in (32) and considering t_f as the running time 't', we obtain

$$\dot{S}(t)\lambda + \dot{\hat{x}}(t) + S(t)[-f_{\hat{x}}^{T}\lambda - 2H^{T}R^{-1}(z(t) - Hx(t)) + 2H^{T}R^{-1}HS(t)\lambda - (\frac{\partial}{\partial\hat{x}}(f_{\hat{x}}^{T}\lambda)S(t)\lambda]$$
$$= f(x(t),t) + \frac{1}{2}\alpha^{-1}Q^{-1}\lambda + f_{\hat{x}}S(t)\lambda \tag{46}$$

We separate terms related to λ from (46) to obtain

$$\dot{\hat{x}} = f(x(t),t) + 2S(t)H^{T}R^{-1}(z(t) - Hx(t)) \tag{47}$$

$$\dot{S}(t)\lambda = S(t)f_{\hat{x}}^{T}\lambda + f_{\hat{x}}S(t)\lambda - 2S(t)H^{T}R^{-1}HS(t)\lambda + \frac{1}{2}\alpha^{-1}Q^{-1}\lambda + S(t)\frac{\delta}{\delta\hat{x}}(f_{\hat{x}}^{T}\lambda)S(t)\lambda \tag{48}$$

Dividing (48) by λ and with $\lambda \to 0$, we get

$$\dot{S}(t) = S(t)f_{\hat{x}}^{T} + f_{\hat{x}}S(t) - 2S(t)H^{T}R^{-1}HS(t) + \frac{1}{2}\alpha^{-1}Q^{-1} \tag{49}$$

Wehave the explicit expression for the model error (discrepancy)by comparing (47) to (33)

$$\hat{d}(t) = 2S(t)H^{T}R^{-1}(z(t) - Hx(t)) \tag{50}$$

Equations (47), (49) and (50) give the generalized (from the model error point of view) invariant embedding based model error estimation algorithm for continuous time system of (33) and (34) in a recursive form. The (49) is often called matrix Riccati equation, like the one in Kalman-Bucy filter.In order to implement the estimation algorithm, we need to solve the matrix differential (49). We can use the following transformation $a = S\,b$

(51)

and using (49)

$$\dot{S}\,b = S\,f_{\hat{x}}^{T}b + f_{\hat{x}}S\,b - 2SH^{T}R^{-1}HSb + \frac{1}{2}\alpha^{-1}Q^{-1}b \tag{52}$$

or $$\dot{S}\,b + 2SH^{T}R^{-1}HSb - Sf_{\hat{x}}^{T}\,b = f_{\hat{x}}\,a + \frac{1}{2}\alpha^{-1}Q^{-1}b \tag{53}$$

We also have $\dot{a} = \dot{S}\,b + S\,\dot{b}$ and $\dot{S}\,b = \dot{a} - S\,\dot{b}$. Using $\dot{S}\,b$ in eqn. (53) and defining $\dot{b}$ as in (54), we obtain

$$\dot{b} = -f_{\hat{x}}^{T}b + 2H^{T}R^{-1}H\,a \tag{54}$$

$$\dot{a} = \frac{1}{2}\alpha^{-1}Q^{-1}b + f_{\hat{x}}\,a \tag{55}$$

Equations (54) and (55) can be solved by using the transition matrix method, as is done in the code for the example.We note here that Q is the weighting matrix for the model error term. It provides normalization to the second part of the cost function (36).

5. GENERALIZED DISCRETE TIME ALGORITHM

Let the nonlinear system be given as

$$X(k+1) = g(X(k),k) \tag{56}$$

$$Z(k) = h(X(k),k) \tag{57}$$

Here 'g' is the vector valued function and Z is the vector of observables defined in the interval $t_0 < t_j < t_N$. Equations (56) and (57) are rewritten to express explicitly the model error (discrepancy)

$$x(k+1) = f(x(k),k) + d(k) \tag{58}$$

$$z(k) = h(x(k),k) + v(k) \tag{59}$$

Here 'f' is the nominal model, which is as such a deficient model. The vector v is measurement noise with zero mean and covariance matrix R, the variable 'd' is the model discrepancy, which is determined by minimizing the criterion

$$J = \sum_{k=0}^{N} \{[z(k) - h(x(k),k)]^T R^{-1} [z(k) - h(x(k),k)] + \alpha d^T(k) Q d(k) \tag{60}$$

Minimization should obtain two things: $\hat{x} \to X$ and estimate $\hat{d}(k)$ for $k = 0,\ldots,N$ .By incorporating the constraint (58) in (60), we obtain

$$\begin{aligned} J_a = \sum_{k=0}^{N} \{[z(k) - h(x(k),k)]^T R^{-1} [z(k) - h(x(k),k)] + \alpha d^T(k) Q d(k) \\ + \lambda^T [x(k+1) - f(x(k),k) - d(k)] \end{aligned} \tag{61}$$

The Euler-Lagrange conditions yield

$$\hat{x}(k+1) = f(\hat{x}(k),k) + \frac{1}{2}\alpha^{-1} Q^{-1} \lambda(k) \tag{62}$$

$$\lambda(k-1) = f_{\hat{x}}^T(\hat{x}(k),k)\lambda(k) + 2H^T R^{-1} [z(t) - H\hat{x}(k)] \tag{63}$$

with $H(k) = \left.\dfrac{\partial h(x(k),k)}{\partial x(k)}\right|_{x(k)=\hat{x}(k)}$ and $d(k) = \dfrac{1}{2}\alpha^{-1} Q^{-1} \lambda(k)$. Equations (62) and (63) constitute a TPBVP which is solved by using invariant embedding method. The resulting generalized discrete time IE recursive algorithm is given as

$$\hat{x}(k+1) = f_{\hat{x}}(\hat{x}(k),k) + 2S(k+1)H^T(k+1)R^{-1}[z(k+1) - h(\hat{x}(k+1),k+1)] \tag{64}$$

$$S(k+1) = \left[\mathrm{I} + 2P(k+1)H^T(k+1)R^{-1}H(k+1)\right]^{-1} P(k+1) \tag{65}$$

$$P(k+1) = f_{\hat{x}}(\hat{x}(k),k)\; S(k)\; f_{\hat{x}}^T(\hat{x}(k),k) + \frac{1}{2}\alpha^{-1} Q^{-1} \tag{66}$$

and $$\hat{d}(k) = 2S(k)H^T(k)\, R^{-1}[z(k) - h(\hat{x}(k),k)] \tag{67}$$

6. CONVENTIONAL INVARIANCE EMBEDDING ESTIMATORS

Now, we can obtain the conventional IE estimators by choosing the value of $\alpha = 1$ in the generalized IE estimators. Conventional continuous time IE estimator is given as

$$\dot{\hat{x}} = f(x(t),t) + 2S(t)H^T R^{-1}(z(t) - Hx(t)) \tag{68}$$

$$\dot{S}(t) = S(t) f_{\hat{x}}^T + f_{\hat{x}} S(t) - 2S(t) H^T R^{-1} H S(t) + \frac{1}{2} Q^{-1} \tag{69}$$

$$\hat{d}(t) = 2S(t)H^T R^{-1}(z(t) - Hx(t)) \tag{70}$$

Conventional discrete time IE estimator is given as

$$\hat{x}(k+1) = f_{\hat{x}}(\hat{x}(k),k) + 2S(k+1)H^T(k+1)R^{-1}[z(k+1) - h(\hat{x}(k+1),k+1)] \quad (71)$$

$$S(k+1) = \left[I + 2P(k+1)H^T(k+1)R^{-1}H(k+1)\right]^{-1} P(k+1) \quad (72)$$

$$P(k+1) = f_{\hat{x}}(\hat{x}(k),k)\, S(k)\, f_{\hat{x}}^T(\hat{x}(k),k) + \frac{1}{2}Q^{-1} \quad (73)$$

and $$\hat{d}(k) = 2S(k)H^T(k)\, R^{-1}[z(k) - h(\hat{x}(k),k)] \quad (74)$$

7. ROBUST ESTIMATION OF MODEL ERROR IN H INFINITY SETTING

Filtering in/of dynamical systems have historically been accomplished through the application of the KF (EKF), while it has been rather successful in general, in some practical applications it has been have found that model uncertainty is problematic. To resolve this issue, filters based on alternative performance criteria have been developed. The design of filters with accurate and predictable performance led to the filters which are termed as robust, and H-infinity (H_∞) norm based H-infinity filters (HIFs) belong to this class of robust estimators; H_∞ filter places a (upper) bound on the (error) energy (variance) gain from the deterministic inputs to the filter error [6]. The problem of estimation of deterministic model error, d, is traditionally solved by formulating it as a TPBVP as we have seen previous sections. This with the use of invariant embedding method (IEM) gives the recursive solution for the state estimates, and we obtain the model errors also explicitly. Robust H_∞filters can estimate the parameters even under some uncertainties with yet acceptable performance.

Performance Norm

A quantitative treatment of the performance and robustness of control systems requires the introduction of appropriate signal and system norms, which give measures of the magnitude of the involved signals and system operators. Consider a LTI system with the disturbance to error transfer function. The system and variables are defined as: i) $G_{\tilde{z}d}$as the TF matrix from the system disturbance, say d, to the estimate error; ii) zasoutput; and iii) $\tilde{z} = z - \hat{z}$as estimate of the error.The L_2-norm of an error signal vector is given as [6]

Continuous signals $$\|\tilde{z}\|_2^2 = \frac{1}{2\pi}\int_{-\infty}^{\infty} \tilde{z}(j\omega)^H \tilde{z}(j\omega)d\omega = \int_{-\infty}^{\infty} \tilde{z}(t)^H \tilde{z}(t)dt \quad (75)$$

Discrete signals $$\|\tilde{z}\|_2^2 = \sum_{n=-\infty}^{\infty} \tilde{z}(n)^H \tilde{z}(n) = \frac{1}{2\pi j}\oint \tilde{z}(e^{j\omega T})^H \tilde{z}(e^{j\omega T}) \frac{de^{j\omega T}}{e^{j\omega T}} \quad (76)$$

However, the H∞-norm of an error system matrix is

Continuous $$\|G_{\tilde{z}d}\|_\infty^2 = \sup_{\|d\|_2 \neq 0} \frac{\|G_{\tilde{z}d}\|_2^2}{\|d\|_2^2} = \sup_\omega \sigma^{-2}(G_{\tilde{z}d}(j\omega)) \quad (77)$$

Discrete $$\|G_{\tilde{z}d}\|_\infty^2 = \sup_{\omega T \epsilon[-\pi,\pi]} \sigma^{-2}(G_{\tilde{z}d}(e^{j\omega T})) \quad (78)$$

G(jw)is the maximum gain at the frequency ω [6]. Consider a stable system with transfer function, G(s). The H_∞ norm is defined as

$$\|G\|_\infty = \max_\omega |G(jw)| \quad (79)$$

In the event that the maximum may not exist, the H_∞ norm of the transfer function matrix G(s) is given by

$$\|G\|_\infty = \sup_\omega \|G(jw)\| \quad (80)$$

For an H_∞ criteria, the TF from the input disturbances to the estimator error shall be required to have a system gain that conforms to an upper bound

$$||G||_{\infty} < \gamma^2 \quad (81)$$

Hence, the performance bound criteria can be written as

$$\sup_{\sim} \frac{||z-\hat{z}||_2^2}{||d||_2^2} < \gamma^2 \quad (82)$$

On simplifying,

$$\sup_{\sim} ||z-\hat{z}||_2^2 - \gamma^2 ||d||_2^2 < 0 \quad (83)$$

Constraint on cost function

Let the mathematical description of dynamic system be given as

$$\dot{x} = f(x(t), u(t), t) + d(t) \quad (84)$$

$$z(t) = Hx(t) + v(t) \quad (85)$$

We have seen that the basic cost function is given by

$$J = \int_{t_0}^{t_f} \left[\left(z(t) - Hx(t)\right)^T R^{-1} \left(z(t) - Hx(t)\right) + (d^T(t)Qd(t)) \right] dt \quad (86)$$

We can combine the conventional LS cost functional and the HI norm based criterion (83) by appending H-infinity norm to the basic cost function to obtain

$$J = \sup ||z-\hat{z}||_2^2 - \gamma^2 ||d||_2^2 < 0 \quad (87)$$

We see from (83) and (87), that to obtain the formal and proper robust estimator we need to have the cost function J strictly negative, and not only necessarily minimum. Here, we invoke the generalized IE estimator criterion

$$J = \int_{t_0}^{t_f} \left[\left(z(t) - Hx(t)\right)^T R^{-1} \left(z(t) - Hx(t)\right) + \alpha(d^T(t)Qd(t)) \right] dt \quad (88)$$

In essence the last terms related to only ‘d’ in (86), (87), and (88) are exactly the same we specify certain new conditions on the cost function

i) If J is strictly negative, we obtain the formal and proper robust estimators using (87) at a γ level
ii) If J is not necessarily strictly negative and it has attained some minimum we obtain generalized IE estimators using (88) at an α level
iii) If J is again some minimum, but with $\alpha = 1$, we obtain the conventional LS based IE estimators.
iv) In (88), if we put $\alpha = -\gamma^2$ in (88), then we can get semi robust estimators at γ level, and we can even call them as (some) adaptive IE based estimators. In this case if we practically (during implementation with real data) get the (actual value of) cost function less than zero then we have obtained practically the robust estimator.

We have already obtained, in the preceding sections, the IE based estimators that would satisfy the condition at ii) and iii).

Semi robust-adaptive IE estimators

Obtaining the IE+HI based formal and proper robust estimators based on condition i) might be rather difficult and hence, we try to obtain the estimators that satisfy the condition iv) on J. We also observe that the semi robust/adaptive (SRA) estimator equations can be easily obtained by using the condition $\alpha = -\gamma^2$, in the generalized estimators. Continuous time SRA estimator is given as

$$\dot{\hat{x}} = f(x(t),t) + 2S(t)H^T R^{-1}(z(t) - Hx(t)) \tag{89}$$

$$\dot{S}(t) = S(t) f_{\hat{x}}^T + f_{\hat{x}} S(t) - 2S(t)H^T R^{-1} HS(t) - \frac{1}{2}\gamma^{-2} Q^{-1} \tag{90}$$

We also, have the explicit expression for the model error (discrepancy)

$$\hat{d}(t) = 2S(t)H^T R^{-1}(z(t) - Hx(t)) \tag{91}$$

And correspondingly (55) becomes

$$\dot{a} = -\frac{1}{2}\gamma^{-2} Q^{-1} b + f_{\hat{x}}\, a$$

Discrete time SRA estimator is given as

$$\hat{x}(k+1) = f_{\hat{x}}(\hat{x}(k),k) + 2S(k+1)H^T(k+1)R^{-1}[z(k+1) - h(\hat{x}(k+1),k+1)] \tag{92}$$

$$S(k+1) = \left[\mathrm{I} + 2P(k+1)H^T(k+1)R^{-1}H(k+1)\right]^{-1} P(k+1) \tag{93}$$

$$P(k+1) = f_{\hat{x}}(\hat{x}(k),k)\; S(k)\; f_{\hat{x}}^T(\hat{x}(k),k) - \frac{1}{2}\gamma^{-2} Q^{-1} \tag{94}$$

and $$\hat{d}(k) = 2S(k)H^T(k)\, R^{-1}[z(k) - h(\hat{x}(k),k)] \tag{95}$$

8. PROCEDURE OF MODEL FITTING TO THE DISCREPANCY AND RESULTS

Once we determine the time history of the discrepancy, d(t), we need to fit a mathematical model to it in order to estimate the parameters of this model by using a regression method. Assume that the original model of the system is given as

$$z(k) = a_0 + a_1 x_1 + a_2 x_1^2 + a_3 x_2 + a_4 x_2^2 \tag{96}$$

Since, we would not know the accurate model of the original system, we would use only a deficient model in the system state equations

$$z(k) = a_0 + a_1 x_1 + a_3 x_2 + a_4 x_2^2 \tag{97}$$

The above equation is deficient by the term $a_2 x_1^2$. When we apply the invariant embedding model error estimation algorithm to determine the discrepancy, we will obtain time history of 'd', when we use the deficient model (97); and once the 'd' is estimated, a model can be fitted to this 'd' and its parameters estimated. In all probability, the estimate of the missing term will be obtained

$$d(k) = \hat{a}_2 \hat{x}_1^2 \tag{98}$$

In (98) $\hat{x}_1$ is the estimate of state from the model error estimation algorithm. In order to decide which term should be added, a correlation test can be used. Then the total model can be obtained as

$$\hat{z}(k) = a_0 + a_1 \hat{x}_1 + \hat{a}_2 \hat{x}_1^2 + a_3 \hat{x}_2 + a_4 \hat{x}_2^2 \tag{99}$$

Under the condition that the model error estimation algorithm has converged, we will get $\hat{x} \to x$ and $\hat{a}_i \to a_i$, thereby obtaining the correct or adequately accurate model of the system [5].The performance of the generalized nonlinear continuous time IE estimator at an alpha level is demonstrated now. We consider the following system, and utilize the continuous time generalized estimator,

$$\dot{x}_1(t) = 2.5\cos(t) - 0.68x_1(t) - x_2(t) - 0.0195x_2^3(t) \quad (100)$$

$$\dot{x}_2(t) = x_1(t) \quad (101)$$

We estimate the model discrepancy in (101) by eliminating the terms from it turn: i) X_2^3, and ii) X_1, X_2, X_2^3
Then, we fit a model to the discrepancy thus estimated

$$d(t) = a_1x_1(t) + a_2x_2(t) + a_3x_2^3(t) \quad (102)$$

to estimate the parameters of the continuous-time nonlinear system. Data are generated by integrating (100) and (101) for a total of 14 sec. using a sampling time =0.05 sec. We consider case in which a deficient model is formulated by removing the term $x_2^3(t)$ from the equation (102). The deficient model is then used in the robust invariant embedding model error estimation algorithm as 'f' and the model discrepancy d(t) is estimated. The parameters are estimated from the model discrepancies using LS method. The numerical result is shown in Table 1. It is to be noted that in all the cases, from the estimated model discrepancy, the parameter that is removed from the model is estimated.Figure 1 shows the model error and state time history match. The match is very good and it indicates that the model discrepancy is estimated accurately by the algorithm. This example shows that the generalized IE method at some alpha level gives satisfactory results.

Table 1 Generalized continuous time TE estimator at alpha level

Parameter	a_1 X_1	a_2 X_2	a_3 X_2^3	Terms removed
True Value	0.68	1	0.0195	-
Estimated value (J=0.00016722)	(0.68)	(1)	0.0188	X_2^3
Conventional IE method (J=0.0012)	(0.68)	(1)	0.0182	X_2^3

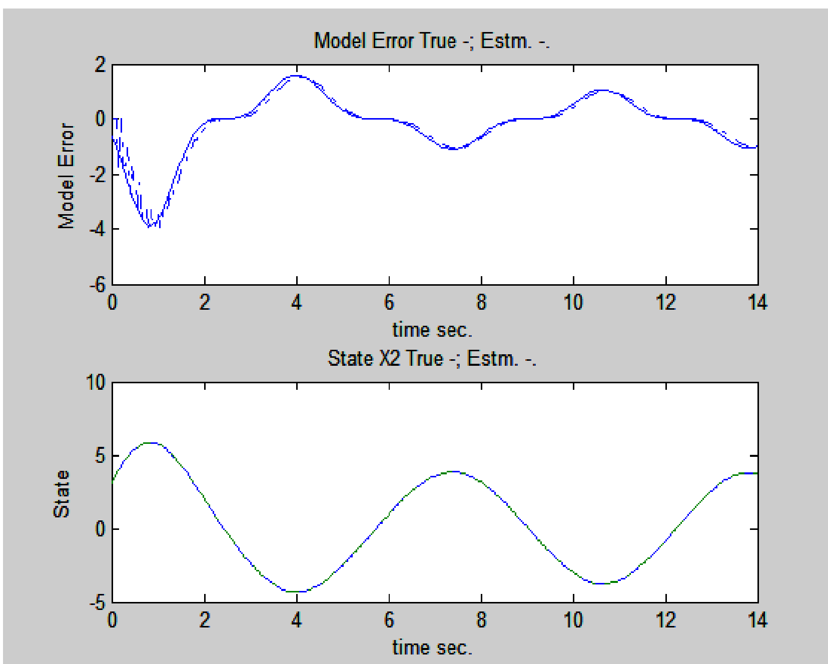

Figure 1. Model error/state time history match by continuous time IE estimator at alpha level

CONCLUDING REMARKS

In this paper we have presented generalized continuous and discrete time invariant embedding estimators for determination of model errors in nonlinear systems. Also, we have indicted the possibility of extending these to HI based robust estimators. The conventional IE based estimators are obtained as a special case, and the performance is compared with the generalized estimator with simulated data form a nonlinear system. The results are encouraging, however further work with more dynamic systems would be useful.

REFERENCES

[1]. Mook, D. J., and Junkins, J. L. Minimum model error estimation for poorly modelled dynamic systems. AIAA 25th aerospace sciences meeting (AIAA-87-0173). Reno, Nevada, USA, 1987.
[2]. Mook, D.J., Estimation and identification of nonlinear dynamic systems, AIAA Jl., Vol. 27, No. 7, pp. 968-974, USA, July 1989.
[3]. Crassidis, J.L, Markley, F.L. and Mook, D.J., A real time model error filter and state estimator, In Proc. of AIAA Conf. On Guidance, Navigation and Control, Arizona, USA, Paper No. AIAA-94-3550-CP, Aug. 1-3, 1994.
[4]. Datchmendy, D.M and Sridhar, R., Sequential estimation of states and parameters in noisy nonlinear dynamical systems, Trans. Of the ASME, Jl. of Basic Engineering, pp. 362-368, USA, June 1966.
[5]. Parameswaran, V. and Raol, J.R., Estimation of model error for nonlinear system identification, IEE Proc. Control Theory and Applications, Vol. 141, No. 6, pp. 403-408, Nov., 1994.
[6]. Paul Lawrence Rawicz. H∞/H2/Kalman Filtering of Linear Dynamical Systems via Variational Techniques with Applications to target tracking, Ph.D. thesis, Drexel University, December 2000.