HUITIEME COLLOQUE SUR LE TRAITEMENT DU SIGNAL ET SES APPLICATIONS

NICE du 1er au 5 JUIN 1981

DETECTION ET PREDICTION D'IMPERFECTIONS DANS LES SYSTÈMES DYNAMICS PAR EVALUATION DES PARAMÈTRES DE DERIVE.

ESTIMATION OF PARAMETERS DRIFT FOR DETECTION AND PREDICTION OF FAULTS IN DYNAMIC SYSTEMS.

L. Kazovsky, A. Cohen, I. Marshand

Department of Electrical Engineering, Ben-Gurion Uni. of the Negev, Beer-Sheva, Israel

RESUME

La detection et la prediction d'imperfections dans un systeme dynamic peuvent être obtenues a l'aide de l'estimation des trajectoires des parametres, dans l'espace des parametres. Durant un fonctionnement normal, les parametres exhibent un mouvement aleatoire autour de leur valeurs nominales. La derive d'un (ou de plusieurs) parametre peut signaler une imperfection. L'estimation de la frequence du parametre de derive a ete obtenue a l'aide de deux algorithmes developpes dans cette etude. La performance des algorithmes developpes a aussi ete estimee.

SUMMARY

System faults may be detected and predicted via estimation of trajectories of the system dynamic parameters in the parameter space. During normal operation the parameters exhibit random motion around their nominal values. A drift of one (or more) parameters may indicate a fault. Two algorithms are developed here for the estimation of the parameter drift rate. Performance of the developed algorithms is also evaluated.

ESTIMATION OF PARAMETERS DRIFT FOR DETECTION AND PREDICTION OF FAULTS IN DYNAMIC SYSTEMS.

L. Kazovsky, A. Cohen, I. Marshand

1. INTRODUCTION

Detection and prediction of faults in dynamic systems is an important aspect of preventive maintenance. It has been applied to both physical[1] and biomedical[2] systems. In this paper we describe two algorithms for detection of drifts in the systems parameters. The decision whether a significant parameters drift exists, is an essential part of fault detection and prediction algorithms. The algorithms imply estimation of the trajectories of the system dynamic parameters in the parameter space.

2. MODEL DESCRIPTION AND PROBLEM STATEMENT

Consider a dynamic system described by its state equations

$$\underline{x}(k+1)=\underline{f}[\underline{x}(k),\underline{u}(k),\underline{\beta}(k)]$$
$$\underline{y}(k)=\underline{g}[\underline{x}(k),\underline{u}(k),\underline{\beta}(k)] \qquad k=0,1,\dots \qquad (1)$$
$$\underline{x}(0)=\underline{x}_0$$

where $\underline{x}(k)$ is the state vector; $\underline{u}$ and $\underline{y}$ are the input and output vectors, respectively; $\underline{\beta}(k)$ is the system's parameters vector; $\underline{x}_0$ is the vector of initial conditions; k represents a discrete time. $\beta(k)$ may be identified on line using a suitable identification algorithm[3]. Denote as $\hat{\underline{\beta}}(k)$ the estimated value of $\beta(k)$. During normal operation, $\hat{\underline{\beta}}(k)$ would exhibit a random motion around a nominal value. When a fault develops, $\hat{\underline{\beta}}(k)$ shifts from the normal operating point; this condition will be referred to as a drift. Linear approximation of the trajectory of $\hat{\underline{\beta}}(k)$ yields

$$\hat{\underline{\beta}}(k)=\underline{\beta}(k_0)+(k-k_0)\underline{a}+\underline{n}(k) \qquad (2)$$

where $\underline{n}(k)$ is a zero mean random vector resulting from the random motion of the parameters and the identification errors; $\underline{a}$ is a vector of drift rates; and k_0 is an arbitrary time point. The forgoing model enables construction of the fault detection algorithm shown in Fig. 1. The algorithm consists of the parameter identifier, trend rate estimator and threshold comparator. The parameter identifier (described elsewhere[3]) computes $\hat{\underline{\beta}}(k)$ and transfers it to the trend rate estimator. The estimator computes $\hat{\underline{a}}$ - the estimate of $\underline{a}$ - and transfers it to the comparator. The comparator compares $\hat{\underline{a}}$ with the threshold vector $\underline{\varepsilon}$ and accepts decisions as follows:

$$\hat{\underline{a}} \geq \underline{\varepsilon}: \text{ drift from the normal condition}$$
$$\hat{\underline{a}} < \underline{\varepsilon}: \text{ normal condition} \qquad (3)$$

In the following sections we describe two different algorithms for trend rate estimation and evaluate their respective performance.

3. HEURISTIC ESTIMATOR

Equation (2) may be rewritten as follows:

$$\hat{\underline{\beta}}(i+k_0)=\underline{\beta}(k_0)+i\underline{a}+\underline{n}(i+k_0) \qquad (4)$$

where $i \triangleq k-k_0$. Consider the sum of N points on the estimated trajectory of $\hat{\underline{\beta}}(\cdot)$,

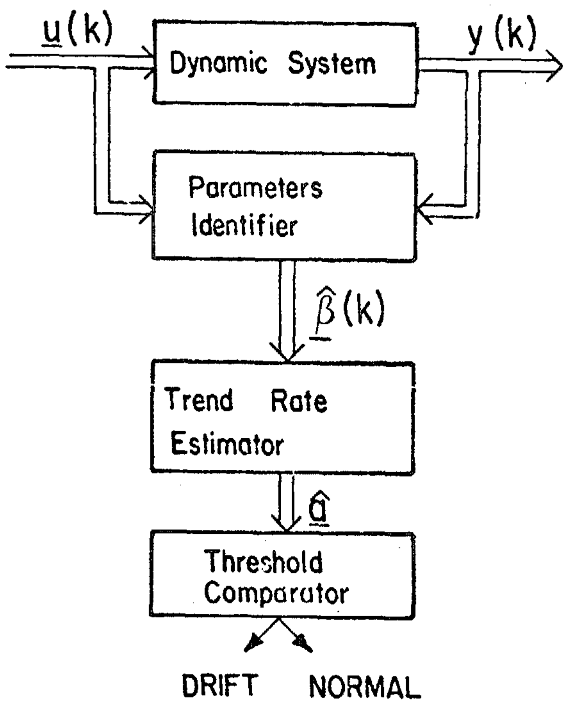

Fig. 1. Block diagram of the suggested system.

as defined by (4):

$$\sum_{i=1}^{N}\hat{\underline{\beta}}(i+k_0)=N\underline{\beta}(k_0)+\underline{a}\sum_{i=1}^{N} i+\sum_{i=1}^{N}\underline{n}(i+k_0)=$$
$$=N\underline{\beta}(k_0)+\underline{a}\frac{N(N+1)}{2}+\sum_{i=1}^{N}\underline{n}(i+k_0) \qquad (5)$$

Define a heuristic estimation algorithm as follows:

$$\hat{a}=\frac{2}{N(N+1)}\left[\sum_{i=1}^{N}\hat{\beta}(k_0+i)-N\hat{\beta}(k_0)\right] \qquad (6)$$

Note that this algorithm requires (N+1) additions and two multiplications. Let us investigate properties of the algorithm. Equation (6) can be rewritten as

$$\hat{\underline{a}}=\underline{a}+\frac{2}{N(N+1)}\sum_{i=1}^{N}\underline{n}(k_0+i)-\frac{2}{N+1}\underline{n}(k_0) \qquad (7)$$

Since $\underline{n}(\cdot)$ has a zero mean, (7) yields

$$E[\hat{\underline{a}}]=\underline{a} \qquad (8)$$

i.e., our heuristic estimator is unbiased. The covariance matrix of the estimator is defined as

$$\underline{K} \triangleq E[(\hat{\underline{a}}-\underline{a})(\hat{\underline{a}}-\underline{a})^T] \qquad (9)$$

where $(\cdot)^T$ denotes a transposition. Diagonal elements of $\underline{K}$ are variances of the errors:

$$\sigma^2_{\hat{a}} \triangleq K_{jj}=\frac{4}{N^2(N+1)^2}E\left[\left(\sum_{i=1}^{N}n_j(k_0+i)-Nn_j(k_0)\right)^2\right]=$$
$$=\frac{4}{N^2(N+1)^2}\left[\sum_{m=1}^{N}\sum_{i=1}^{N}R_n(m-i)-2N\sum_{i=1}^{N}R_n(i)+N^2\sigma_n^2\right] \qquad (10)$$

where $R_n(\cdot)$ is the noise autocorrelation function and σ_n^2 is the noise variance. Note that $\sigma_{\hat{a}}^2, R_n(\cdot)$ and σ_n^2 can be different for different j; however, we do not indicate it explicitly in order not to muddle up the notation. Let us consider two practically important particular cases:

Uncorrelated noise

$$R_n(i) = \sigma_n^2 \delta(i), \quad \forall i \tag{11}$$

Substituting (11) in (10) yields

$$\frac{\sigma_{\hat{a}}^2}{\sigma_n^2} = \frac{4}{N(N+1)} \tag{12}$$

Exponentially correlated noise

$$R_n(i) = \sigma_n^2 \exp(-\alpha|i|), \quad \forall i \tag{13}$$

Substituting (13) in (10) yields

$$\frac{\sigma_{\hat{a}}^2}{\sigma_n^2} = \frac{4}{N^2(N+1)^2}\Big[N + \frac{2\exp(-\alpha)}{1-\exp(-\alpha)}(N-1) + \frac{2\exp(-\alpha)(1-\exp(-\alpha N))}{(1-\exp(-\alpha))^2} - 2N\exp(-\alpha)\cdot\frac{1-\exp(-\alpha N)}{1-\exp(-\alpha)} + N^2\Big] =$$

$$= \frac{4}{N^2(N+1)^2}\Big\{N(N+1) + \frac{2\exp(-\alpha)}{1-\exp(-\alpha)}\big[N\exp(-\alpha N)-1\big] + \frac{2\exp(-\alpha)\big[1-\exp(-\alpha(N-1))\big]}{\big[1-\exp(-\alpha)\big]^2}\Big\} \tag{14}$$

4. MAXIMUM LIKELIHOOD ESTIMATOR

Assume components of $\underline{\beta}(k)$ be mutually independent. This condition need not be valid in actual application. However, it yields a useful algorithm, and allows us to treat each of the components of $\underline{\beta}(k)$ separately. Denote the j-th component of $\underline{\hat{\beta}}(k)$ as $\hat{\beta}_j(k)$. Hence,

$$\hat{\beta}_j(k) = \beta_j(k_0) + a_j(k-k_0) + n_j(k) \tag{15}$$

where a_j and $n_j(k)$ are scalars. Assume $\{n_j(k)\}$ be zero mean Gaussian random variables with known covariance matrix, and compute an N-dimensional vector $\underline{r}$:

$$\underline{r} \triangleq [r_1 \;\; r_2 \ldots r_k \ldots r_N]^T \tag{16}$$

where $r_k \triangleq \hat{\beta}_j(k) - \hat{\beta}_j(k-1) =$

$$= a_j + n_j(k) - n_j(k-1), \quad k = 1, 2, \ldots N \tag{17}$$

Notice that $\underline{r}$ is a Gaussian vector whose mean is equal to $a_j \underline{I}$ ($\underline{I}$ denotes a unity vector), and conditional covariance matrix is

$$\underline{\Lambda} \triangleq E[(\underline{r} - a_j\underline{I})(\underline{r}^T - a_j\underline{I}^T)] \tag{18}$$

$\underline{\Lambda}$ can be easily found from the covariance matrix of $\{n(k)\}$, as we shall see shortly.

The conditional probability density of $\underline{r}$ is

$$P(\underline{r}|a_j) = [(2\pi)^{N/2}|\underline{\Lambda}|^{1/2}]^{-1}\exp[-\tfrac{1}{2}(\underline{r}^T - a_j\underline{I}^T)\underline{\Lambda}^{-1}(\underline{r} - a_j\underline{I})] \tag{19}$$

The ML estimate satisfies the condition[4]

$$\frac{\partial}{\partial a_j} \ln P(\underline{r}|a_j)\Big|_{a_j=\hat{a}_j} = 0 \tag{20}$$

Substituting (19) in (20) yields

$$\frac{\partial}{\partial a_j}\left[\frac{1}{2}\underline{r}^T\underline{\Lambda}^{-1}\underline{r} - \frac{1}{2}a_j\underline{I}^T\underline{\Lambda}^{-1}\underline{r} - \frac{a_j}{2}\underline{r}^T\underline{\Lambda}^{-1}\underline{I} + \frac{1}{2}a_j^2\underline{I}^T\underline{\Lambda}^{-1}\underline{I}\right]\Bigg|_{a_j=\hat{a}_j} = 0 \tag{21}$$

But Λ^{-1} is a symmetrical matrix. Consequently, (21) yields

$$\frac{\partial}{\partial a_j}\left[-a_j\underline{r}^T\underline{\Lambda}^{-1}\underline{I} + \frac{a_j^2}{2}\underline{I}^T\underline{\Lambda}^{-1}\underline{I}\right]\Bigg|_{a_j=\hat{a}_j} = 0 \tag{22}$$

Solving (22), we obtain

$$\hat{a}_j = \sigma^2 \cdot \underline{r}^T\underline{\Lambda}^{-1}\underline{I} \tag{23}$$

where $\sigma^2 \triangleq (\underline{I}^T\underline{\Lambda}^{-1}\underline{I})^{-1}$ should be computed before the test begins. (23) may be also rewritten as

$$\hat{a}_j = \sigma^2 \sum_{k=1}^{N} r_k g_k \tag{24}$$

where $\{g_k\}$ are the components of the vector $\underline{g} \triangleq \underline{\Lambda}^{-1}\underline{I}$.

Let us find now the variance of the ML estimate $\hat{a}_j$. Applying the operator of mathematical expectation to (23) we obtain

$$E[\hat{a}_j] = \sigma^2 \cdot \underline{I}^T \cdot \underline{\Lambda}^{-1}E[\underline{r}] = a_j \cdot \sigma^2\underline{I}^T\underline{\Lambda}^{-1}\underline{I} = a_j \tag{25}$$

We see that $\hat{a}_j$ is an unbiased estimate. In addition,

$$\frac{\partial \ln P(\underline{r}|a_j)}{\partial a_j} = \underline{r}^T\underline{\Lambda}^{-1}\underline{I} - a_j\underline{I}^T\underline{\Lambda}^{-1}\underline{I} = (\hat{a}_j - a_j)\sigma^{-2} \tag{26}$$

where the first part follows from (19) and the second from (23). If (25) and (26) are both satisfied, then the estimate is effective[4]. Hence, its variance may be found from the Kramer-Rao theorem:

$$\sigma_{\hat{a}}^2 = \{-E[\frac{\partial^2 \ln P(\underline{r}|a_j)}{\partial a_j^2}]\}^{-1} \tag{27}$$

Substituting (19) in (27), we obtain

$$\sigma_{\hat{a}}^2 = \{-E[-\underline{I}^T\underline{\Lambda}^{-1}\underline{I}]\}^{-1} = (\underline{I}^T\underline{\Lambda}^{-1}\underline{I})^{-1} = \sigma^2 \tag{28}$$

The physical meaning of the constant σ^2 is now clear; it is the variance of the ML estimator. Now let us again consider two practically important particular cases.

Uncorrelated noise.

In this case, defined by eq. (11), $\underline{\Lambda} = 2\sigma_n^2\underline{A}$ where

$$\underline{A} \triangleq \begin{bmatrix} 1 & -0.5 & & & 0 \\ -0.5 & 1 & -0.5 & & \\ & -0.5 & 1 & & \\ & & -0.5 & & -0.5 \\ 0 & & & -0.5 & 1 \end{bmatrix} \qquad (29)$$

The matrix A is a symmetric tridiagonal Toeplitz matrix. To calculate $\sigma^2_{\hat{a}}$ (eq. 18) one requires the sum of components of the Matrix A^{-1} in (29). An algorithm for inverting a general tridiagonal Toeplitz matrix is given by Roebuck and Barnett[5]. The matrix $\underline{A}$ in[5] is more general, it is defined as:

$a_{i,i}=1$, $a_{i=1,i}=a$, $a_{i,i-1}=c$. $i=1,2,\dots N$ and all other $a_{ij}=0$. This matrix is inverted by the following algorithm:

1. Set $x_0=1$; $x_1=1-ac$ and $x_k=x_{k-1}-ac\,x_{k-2}$
2. for $k=1,2,\dots N$ form: (30)

$$b_0 = 1$$

$$b_k = \begin{cases} a^{-k}\,x_{k-1}; & i \le j \\ c^{-k}\,x_{k-1}; & i > j \end{cases} \qquad (31)$$

The inverse of the matrix $A^{-1}=\{\alpha_{i,j}\}$ is given

$$\alpha_{i,j} = \begin{cases} (ab_N)^{-1}\,b_{i-1}\,b_{N-i}; & i = j \\ (ab_N)^{-1}\,(-1)^{i+j}\,b_{i-1}\,b_{N-j}; & i < j \\ (cb_N)^{-1}\,(-1)^{i+j}\,b_{j-1}\,b_{N-i}; & i > j \\ \gamma_{N+1-j,N+1-i}; & \forall\, i,j \end{cases} \qquad (32)$$

For the covariance matrix of interest (eq. 29) we have $a = c = -1/2$ hence equation (30) becomes

$$x_k = x_{k-1} - 1/4\,x_{k-2} \qquad (33)$$

Solving this difference equations yields

$$x_k = \frac{k+2}{2^{k+1}} \qquad (34)$$

Then equation (31) yields

$$b_k = (-1)^k\,(k+1) \qquad (35)$$

and the inverse matrix is

$$\alpha_{i,j} = \begin{cases} \dfrac{2i\,(N-j+1)}{N+1}; & i \le j \\ \dfrac{2j\,(N-i+1)}{N+1}; & i \ge j \end{cases} \qquad (36)$$

The variance of the estimate is given from (28), (29) and (36):

$$\frac{\sigma^{-2}_{\hat{a}}}{\sigma^{-2}_n} = \Big(\sum_{i=1}^{N} \frac{i\,(N-i+1)}{N+1} + \sum_{j=1}^{N-1}\sum_{i=j+1}^{N} \frac{j\,(N-i+1)}{N+1} =$$

$$= \frac{N(N+1)}{2} - \sum_{i=1}^{N} i^2 + \frac{2}{N+1}\sum_{j=1}^{N-1}\sum_{i=j+1}^{N} j(N-i+1) =$$

$$= \frac{N(N^2+1)}{2} - \frac{1}{N+1}\sum_{i=1}^{N} i^2 - \frac{2N+1}{N+1}\sum_{j=1}^{N-1} j^2 + \frac{1}{N+1}\sum_{j=1}^{N-1} j^3 \qquad (37)$$

To calculate the last three terms consider the Euler-Mclaurin Sum Formula

$$\sum_{x=1}^{N-1} f(x) = \int_1^N f(x)\,dx + \sum_{i=1}^{\alpha} \frac{Bi}{i!} f^{(i-1)}(x)\Big|_{x=1}^{x=N} \qquad (38)$$

where Bi are the Bernoulli coefficients.

For $f(x) = x^2$ (38) yields

$$\sum_{j=1}^{N-1} j^2 = \frac{1}{3}(N^3-1) - \frac{1}{2}(N^2-1) + \frac{1}{6}(N-1) \qquad (39)$$

and for $f(x) = x^3$

$$\sum_{j=1}^{N-1} j^3 = \frac{1}{4}(N^4-1) - \frac{1}{2}(N^3-1) + \frac{1}{4}(N^2-1) \qquad (40)$$

hence

$$\frac{\sigma^2_{\hat{a}}}{\sigma^2_n} = \frac{12(N+1)}{N}\,\frac{1}{N^3+4N+5N+2} = \frac{12}{N(N+1)(N+2)} \qquad (41)$$

Comparing the last equation with equation (12) shows that the ratio of the variance of the ML estimator to that of the heuristic estimator (for uncorrelated noise) -

$$\frac{(\sigma^2_{\hat{a}})_{ML}}{(\sigma^2_a)_H} = \frac{3(N+1)^2}{N^3+4N^2+5N+2} = \frac{3}{N+2} \qquad (42)$$

Exponentially Correlated Noise

For exponential correlation (equation (13)) the covariance matrix Ω is given by

$$\Omega_{i,j} = \begin{cases} \sigma_n^2\cdot\delta\cdot\gamma\cdot\exp(-\alpha|i-j|); & i \ne j \\ \sigma_n^2\cdot\gamma; & i = j \end{cases} \qquad (43)$$

where $\gamma = 2(1-\exp(-\alpha))$; $\delta = (2-\frac{\gamma}{2}-(\frac{\gamma}{2})^{-1})\cdot\gamma^{-1}$.

$$\Lambda=\gamma\sigma_n^2\begin{bmatrix} 1 & \delta\exp(-\alpha) & \delta\exp(-2\alpha)\ldots\delta\exp(-\alpha(N-1)) \\ \delta\exp(-\alpha) & 1 & \delta\exp(-\alpha)\ldots\delta\exp(-\alpha(N-2)) \\ \vdots & & \\ \vdots & & 1 \quad \delta\exp(-\alpha) \\ \delta\exp(-\alpha(N-1))\ldots & \ldots\delta\exp(-\alpha) & 1 \end{bmatrix} \quad (44)$$

This is a symmetric Toeplitz matrix. Its inverse can be calculated numerically but no close form is available.

Note that the M.L. estimate is given by:

$$\hat{a}_i = (\underline{I}^T\underline{\Lambda}^{-1}\underline{I})^{-1}\ \underline{r}^T\underline{\Lambda}^{-1}\underline{I} =$$

$$= (\underline{I}^T\underline{\Lambda}^{-1}\underline{I})^{-1}\sum_{k=1}^{N} r_k g_k \quad (45)$$

where g_k are the components of the vector

$\underline{g} = \underline{\Lambda}^{-1}\underline{I}$

assuming that the quantities $(\underline{I}^T\Lambda^{-1}\underline{I})^{-1}$ and $\underline{\Lambda}^{-1}\underline{I}$ are precalculated, the calculation of the estimate requires N multiplications and N-1 additions.

For comparison recall that the heuristic estimator used only two multiplications and N+1 additions.

The variances of the two estimators as a function of N, are shown in Fig. 2.

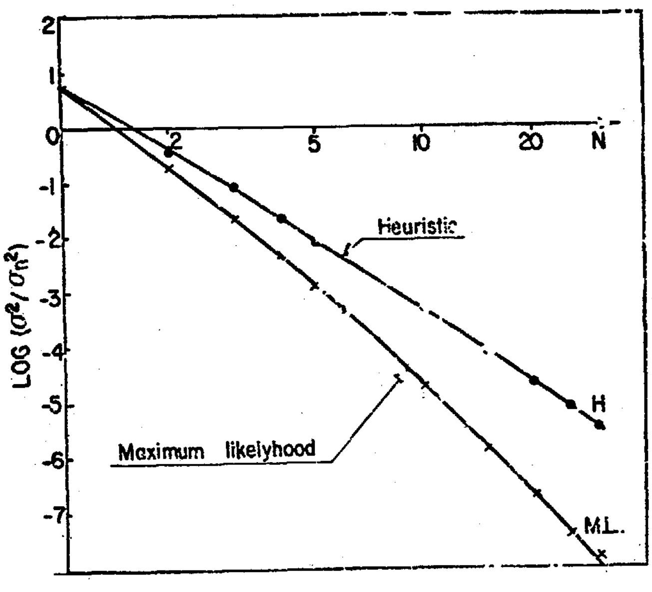

Figure 2: Variance of the two estimators versus the number of samples.

5. EXAMPLE

To demonstrate the applicability of the parameter drift estimation to fault detection, consider the following simple example. A linear discrete system, H(z), is given by

$$H(z) = \frac{1+\beta_1 Z^{-1}}{1+\beta_2 Z^{-1}} \quad (46)$$

where $\underline{\beta}^T = [\beta_1, \beta_2]$; $\beta_1 = -0.8 + 1.6U(k-k_0)$; $\beta_2 = 1-2U(k-k_0)$ and $U(\cdot)$ is the unit step function. The system is under normal operating conditions with $\underline{\beta}^T = [-0.8,\ 1]$ at times $k<k_0$, (point 1 in Fig. 3). It has undergone a

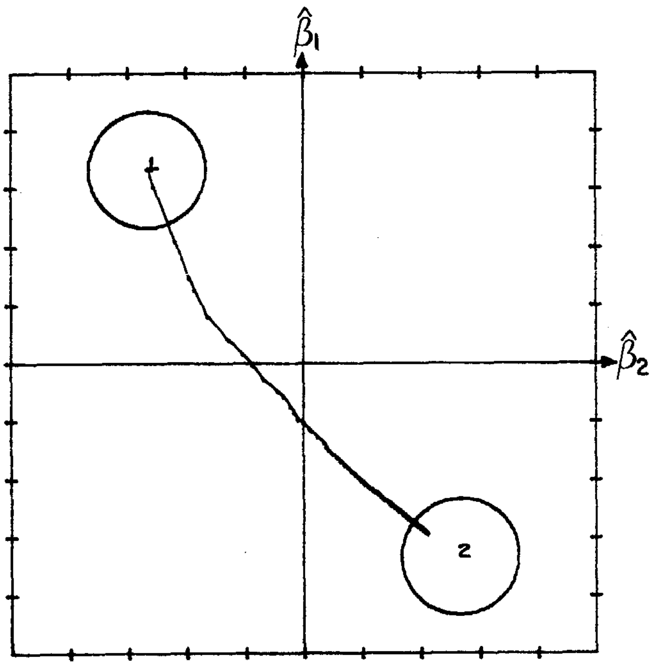

Figure 3: Parameters trajectory in the two dimensional space.

fault at $k = k_0$ and its location in the parameter space has shifted to $\underline{\beta}^T = [0.8,-1]$ (point 2 in Figure 3). Figure 3 shows the two dimensional parameter space with the normal operating region 1, the parameters trajectory and the fault region 2. The parameter vector $\underline{\beta}$ was estimated by a recursive least squares algorithm. Figure 4 shows the parameters estimates as a function of time. The heuristic algorithm was used to estimate the drifts $\hat{a}_1$ and $\hat{a}_2$. These are shown in Figure 5. Detection of the fault is provided by applying a threshold to $\hat{a}_1$ and $\hat{a}_2$. Defining the drift vector

$$\underline{\hat{a}}^T = [\hat{a}_1,\ \hat{a}_2] \quad (47)$$

We use equation (4) to predict the motion of the system in the parameter space at time (k+i)

$$\beta p(k+i) = \underline{\hat{\beta}}(k)+i\cdot\underline{\hat{a}}(k) \quad (48)$$

Here $\beta p(k+i)$ is the predicted location of the system at time k+i, based on the estimated parameters $\underline{\hat{\beta}}(k)$ and drifts $\underline{\hat{a}}(k)$ at time k.

ESTIMATION OF PARAMETERS DRIFT FOR DETECTION AND PREDICTION OF FAULTS IN DYNAMIC SYSTEMS.

L. Kazovsky, A. Cohen, I. Marshand.

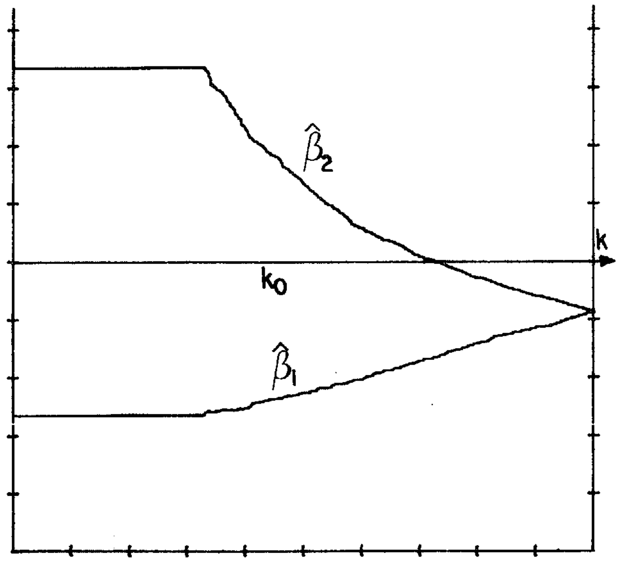

Figure 4: Estimated parameters $\hat{\beta}_1$ and $\hat{\beta}_2$ vs time.

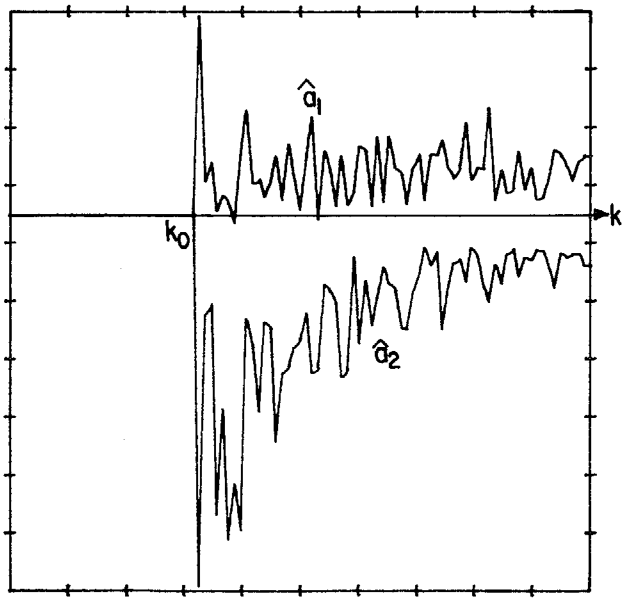

Figure 5: Estimated drifts $\hat{a}_1$, $\hat{a}_2$ vs time.

6. CONCLUSIONS

A method for detection and prediction of faults in dynamic systems has been presented. The method is based on the assumption that the fault causes the dynamic parameters of the system to change from a given "normal" region to a "new" fault location. The systems parameters are constantly identified using an on-line identification algorithm. The identified parameters are observed to detect a drift. When a drift is detected the system is considered to be under a fault. The estimated drift is used to predict the future trajectory of the parameters and thus yield information concerning the type of fault the system is moving at and the estimated time of arrival at the faulty location.

Two estimation algorithms for the estimation of the drift were discussed. Both algorithms yield unbiased estimators. The variance of the heuristic estimator is inversely proportional to N^2 while that of the ML is inversely proportional to N^3 (for uncorrelated noise). The heuristic estimator, however, requires much less computations.

REFERENCES

1. Oe, S., T. Soeda and T. Nukamizo, "A Method of Predicting Failure or Life for Stochastic Systems by Using Auto Regressive Models", Int. J. Syst. Sci., 1980, Vol. 11, pp. 1177-1188.

2. Cohen, A., "Parameter Estimation for Medical Diagnosis", Proc. of the 3rd IFAC Symp. on Identification and System Parameter Estimation, The Hague, Netherlands, June 1973, pp. 239-241.

3. Hastings-Hames, R. and M. W. Sage, "Recursive Generalized Least Squares Procedure for On Line Identification of Process Parameters", Proc. IEE, 1969, Vol. 116, pp. 2057-2062.

4. H. L. Van Trees, "Detection, Estimation and Modulation Theory", Part I, N.Y., Wiley, 1969.

5. Roebuck, P. A. and S. Barnett, "A Survey Toeplitz and Related Matrices", Int. J. of Syst. Sci., 1978, Vol. 9, pp. 921-934.