数理解析研究所講究録
562巻 1985年 1-18

Fault Diagnosis of Linear Systems: Graph Theoretical Approach

Yoshito Ohta

Hajime Maeda

and

Shinzo Kodama

Department of Electronic Engineering

Osaka University, Suita, Osaka 565, Japan

Abstract.

This paper considers a fault diagnosis problem of linear systems from graph theoretical viewpoint. Specifically, we derive conditions for fault distinguishability, and also we derive a fault diagnosis algorithm which decides the faulty element from the system observation. The conditions are given in terms of the system representation graph which depicts the system structure.

1. Introduction.

This paper considers a fault diagnosis problem of linear systems and derives fault distinguishability conditions and a fault diagnosis algorithm.

The fault diagnosis techniques studied in this paper are based

on the fact that when the system is faulty the deviation of the observation vector lies in the certain subspace corresponding to the fault. This method is utilized as a fault diagnosis method for linear electrical networks [1]. Structural conditions for fault distinguishability are obtained for electrical networks [2],[3]. Without restricting the systems to electrical networks, fault diagnosis method for linear systems has been discussed [4],[5].

In this paper, we present a fault diagnosis algorithm and graphical distinguishability conditions. The algorithm is based on linear algebraic consideration, which serves a simple calculation method.

This paper considers the problem in the following manner. Firstly, it is shown that the observation vector lies in the subspace corresponding to the fault when the system is malfunctioning. Then this fact is used to distinguish faults under a certain condition which is made clear in this paper. Based upon these consideration, a fault diagnosis algorithm is proposed. Finally, graphical distinguishability conditions are derived. This leads to a design method of observation point since they clarify the relation between fault distinguishability and the system structure. The conditions are also essential in carrying out the fault diagnosis algorithm.

2. Principle of fault diagnosis.

2.1. System description.

Consider the linear system

$$Ax = 0, \tag{1}$$

where A is a linear map: $\mathcal{V} \to \mathcal{E}$, $\mathcal{V}$ is the n-dimensional state space, $\mathcal{E}$ is the n-dimensional error space (the name is clarified later), and $x \in \mathcal{V}$ is the state vector. We assume that eq.(1) is the complete set for the discription of the system, i.e., A has an inverse map $A^{-1}: \mathcal{E} \to \mathcal{V}$.

Eq.(1) can represent several kind of systems by choosing the map A: a linear stesdy system by a constant $n \times n$ matrix; a linear dynamical system in state-space form with the system matrix F by $sI-F$; etc.

When the system is faulty, eq.(1) cannot be satisfied, and the system is described by

$$Ax = \varepsilon, \tag{2}$$

where $\varepsilon \in \mathcal{E}$ is called the error vector (this clarifies the name of $\mathcal{E}$).

We assume that fault is brought about by the combination of the elementary faults, each of which corresponds to the system component. This implies that we assume that there are the <u>elementary fault vectors</u> $u_1,\ldots,u_t$ (t is the number of the elementary faults) and that the error vector ε is a linear combination of the elementary fault vectors.

When the faulty elements correspond to the elements of the subset J of the index set $\{1,\ldots,t\}$, we say <u>fault J</u> has occured. If fault J occurs, the error vector is contained in the subspace

$$\mathcal{F}(J) := \text{span } \{u_i,\ i \in J\},$$

where span denotes the operation which maps the subset of the linear space to the minimum linear subspace which contains the subset. $\mathcal{F}(J)$ is called the <u>fault subspace</u> (of the fault J). Fault J is called <u>k-fault</u> if $\dim \mathcal{F}(J) = k$. Fault I and fault J are <u>equivalent</u> if $\mathcal{F}(I) = \mathcal{F}(J)$. If fault I and fault J are equivalent, we cannot tell which fault has occured by examining to which subspace the error vector belongs. If $\varepsilon \in \mathcal{F}(J)$, we say the <u>fault equivalent to fault J</u> has occured.

2.2 Principle of fault diagnosis

It is not always possible to observe all the states. Accordingly, we consider direct sum decomposition of the state space $\mathcal{V} = \mathcal{V}_o \oplus \mathcal{V}_{\bar{o}}$. $\mathcal{V}_o$ is called the <u>observation space</u>, and $\mathcal{V}_{\bar{o}}$ is called the <u>nonobservation space.</u> The observation vector is given by

$$y = Cx, \tag{3}$$

where C is the projection map from $\mathcal{V}$ to $\mathcal{V}_o$ along $\mathcal{V}_{\bar{o}}$.

If fault J has occured, then $y \in \mathcal{S}(J) := CA^{-1}\mathcal{F}(J)$ since $\varepsilon \in \mathcal{F}(J)$. We call $\mathcal{S}(J)$ <u>fault observation subspace</u> (corresponding to fault J).

An outline of the fault diagnosis is as follows; we seek the fault observation subspace which contains observation vector y, and decide which fault has occured. To be accurate, there are two cases to be examined; (i) $y \in \mathcal{S}(I) \cap \mathcal{S}(J)$; and (ii) $\mathcal{S}(I) = \mathcal{S}(J)$.

First we consider the case (i). We adopt the criterion that the fault whose fault observation subspace is a minimum subspace containing y has occured. This criterion has some ambiguity

because it ignores two situations; (a) $y \in S(J)$ when fault I occurs, and $S(J) \subsetneqq S(I)$; and (b) $S(I) \not\subseteq S(J)$, $S(I) \not\supseteq S(J)$, and $y \in S(I) \cap S(J)$. The former leads to a misjudgement. When the latter is the case, the uniqueness of the minimum subspace may not be guaranteed. These situations, however, happen only if the error vector lies in the sum of finite number of proper subspaces of the fault subspace, and therefore we can expect these situations are exceptional. This consideration justifies the following hypothesis which we assume to hold.

[Hypothesis H] If fault I occurs, then $y \notin S(I) \cap S(J)$ for every fault J such that $S(I) \not\subseteq S(J)$.

We say fault I and fault J are <u>distinguishable</u> if $S(I) \neq S(J)$, since if this holds we can tell which fault has occured by the observation vector under the Hypothesis H.

The case (ii) is inevitable if $\dim F(\{1,\ldots,t\}) > m$ (m is the dimension of the observation space). This necessitates introducing the concept of k-distinguishability. Before defining the concept, note that equivalent faults are not distinguishable no matter what the observation space is. We do not try to distinguish equivalent faults.

We say the system (A,C) (eq.(1) and eq.(3)) is <u>k-distinguishable</u> if fault I and fault J are distinguishable for arbitrary j-fault J ($j \leqq k$) and arbitrary fault I which is not equivalent to J. It is shown in ch.4 Theorem 1 that if the system is k-distinguishable, then from the observation vector y (a) we can tell whether j-fault ($j \leqq k$) occurs or not, and if the

answer is affirmative (b) we can specify the fault uniquely within the equivalence of faults, i.e., we can specify the fault which is equivalent to the actual fault. A procedure which carry out the judgement (a) and (b) is called <u>k-fault diagnosis method</u>. It is clear that k-distinguishability implies (k-1)-distinguishability ($k > 0$).

A necessary and sufficient condition for the system (A,C) to be k-distinguishable is given in terms of the dimension of the fault subspaces. We make an insubstantial assumption that there is a (k+1)-fault.

[Proposition 1] [2] The system (A,C) is k-distinguishable ($k < n$) if and only if $\dim \mathcal{S}(J) = k+1$ for arbitrary (k+1)-fault J.

3. Fault diagnosis algorithm

In this chapter, we give k-fault diagnosis algorithm based on the principle presented in the previous chapters. We assume that $CA^{-1}u_i$, $(1 \leq i \leq t)$ span the observation space.

[Theorem 1] (k-fault diagnosis algorithm) Suppose that the system (A,C) is k-distinguishable. We assume that the observation vector y satisfies Hypothesis H. Then the following algorithm offer a k-fault diagnosis method for the system (A,C).

Algorithm:

(Step 1) Find a k-cover $\mathbf{K}$ which is a subset of the power set of the fault subspaces satisfying;

(a) for arbitrary $\mathcal{F}(I) \in \mathbf{K}$, I is an m-fault and satisfies

dim (I) = m; and

(b) for arbitrary j-fault J ($j \leqq k$), there exists $\mathcal{F}(I) \in \mathbf{K}$ such that $\mathcal{F}(J) \subset \mathcal{F}(I)$.

(Step 2) For every $\mathcal{F}(I) \in \mathbf{K}$, the linear map $CA^{-1}|_{\mathcal{F}(I)}: \mathcal{F}(I) \to \mathcal{V}_o$ has an inverse map M(I). Define $\mu(I) := M(I)y$. If $\mu(I)$ $\mathcal{F}(J)$ holds for j-fault J ($j \leqq k$) satisfying $\mathcal{F}(J) \subset \mathcal{F}(I)$, then a fault equivalent to fault J has occured.

(Step 3) If the condition in Step 2 does not hold for every $\mathcal{F}(I) \in \mathbf{K}$, then a j-fault ($j > k$) has occured.

The following lemma gurantees the feasibility of Step 1.

[Lemma 1] (k-cover) There exists a k-cover for the system (A,C) if and only if the system is (k-1)-distinguishable.

4. Concrete representation of the fault observation subspaces.

In this chapter, we consider specific linear systems. This enables us to represent the fault observation subspace concretely if the specific discription of the system is given.

(1) Linear steady system

Consider n-set of linear algebraic equations

$$Ax = b \tag{4}$$

where $A \in R^{nxn}$, and $b \in R^n$ is a constant vector. In this case, $\mathcal{V} = \mathcal{E} = R^n$. We assume that every equation in (4) represents the equation which describes the system component. This means that the elementary fault vectors u_i ($1 \leqq i \leqq t$) are the unit vectors e_i ($1 \leqq i \leqq n$) in R^n. The fault observation subspace corresponding to fault $J \subset \{1,\ldots,n\}$ is

$$F(J) = \text{span}\,\{CA^{-1}e_i,\ i \in J\}.$$

(2) Compartmental systems

We consider two types of steady compartmental systems; (a) closed systems:

$$Ax = 0,\ \mathbf{1}^T x = w,\ \mathbf{1}^T = (1,\ldots,1), \tag{5}$$

where A is the compartmental matrix, x is the state vector, and w is a constant; and (b) open systems:

$$Ax + Bu = 0 \tag{6}$$

where A and x are the same as (a), B is the input matrix, and u is a constant vector. The elements of A are related with the rate constants k_{ij}'s;

$$a_{ij} = k_{ij}\ (i \neq j), \tag{7}$$

$$a_{jj} = -k_{0j} - \Sigma_{i \neq j}\, k_{ij}.$$

We say the system is faulty if a part of the rate constants take abnormal values apart from zero-fixed rate constants. This implies that we should choose as the elementary fault vectors

$$e_i - e_j \text{ (for } k_{ij}\text{), and}$$

$$e_j \text{ (for } k_{0j}\text{).}$$

These vectors are called the edge vectors since they have one-to-one correspondence to the edges of the compartmental graph mentioned in ch.5. We avoid double-index and write the edge vectors as $f_1,\ldots,f_t$, where t is the number of edges.

First we consider closed system. When k_{0j}'s (leaks) are zero, the system is called a closed system. We assume that the matrix is irreducible. This can be expressed in terms of the compartmental graph that the graph is strongly connected. Then the

steady state of zero input response follows eq.(5) when the total amount of mass is w. The state space $V = R^n$, and the error space $E = R(A) \oplus R^1$ where $R(\cdot)$ denotes the range space. The system map is $A \oplus 1^T$. The matrix $[A^T\ 1]^T$ has an left inverse $[A^-\ d]$ since A is irreducible. The map $A^-: R(A) \to V$ is independent of the choise of the left inverse. The fault observation subspace corresponding to fault J is

$$F(J) = \text{span}\ \{CA^- f_i,\ i \in J\}. \tag{8}$$

When one of k_{0j}'s is not zero, the system is called an open system. We assume that the matrix [A B] is irreducible, or that every node is accessible from the input nodes on the graph. Also we assume that every node is accesible to environment, i.e., there is a leakage path from every node. Then the steady state follows eq.(6) when we apply constant input u. The state space $V = R^n$, and the error space $E = R(A) = R^n$. The fault observation subspace corresponding to fault J is

$$F(J) = \text{span}\ \{CA^{-1} f_i,\ i \in J\}. \tag{9}$$

(3) Linear dynamical systems in state space form

Consider the linear dynamical system described by state space form

$$\dot{x} = Ax + Bu, \tag{10}$$

where $A \in R^{nxn}$ is the system matrix, $x \in R^n$ is the state vector, $B \in R^{nxr}$ is the input matrix, and $u \in R^r$ is the input vector. Laplace transformation yields

$$(sI-A)x = Bu + x_0. \tag{11}$$

The state x and the input u are elements of ΛR^n and ΛR^r

respectively, where ΛR is the field of truncated R-valued Laurant series of the form

$$z = \Sigma_{t=t_0}^{\infty} z_t s^{-t}$$

with coefficient addition and convolutional product. As in 4.1, we assume every equation in (11) represents the system component. The elementary fault vectors are the unit vectors e_i $(1 \leq i \leq n)$ in ΛR^n. The fault observation subspace corresponding to fault J is

$$F(J) = \text{span}_{\Lambda R} \{C(sI-A)^{-1}e_i,\ i \in J\}. \qquad (12)$$

(4) Linear dynamical systems in descriptor form

Consider the linear dynamical systems in descriptor form

$$E\dot{x} = Ax + Bu, \qquad (13)$$

where $E \in R^{nxn}$, and A, B, x, and u are the same as (a). Laplace transformation yields

$$(sE-A)x = Bu + Ex_0. \qquad (14)$$

The state space, the error space, and the elementary fault vectors are the same as (a). The fault observation subspace corresponding to fault J is

$$F(J) = \text{span}_{\Lambda R} \{C(sE-A)^{-1}e_i,\ i \in J\}. \qquad (15)$$

5. Graphical distinguishability conditions

In this chapter, we first describe the system structure inherent in the system. Then we give the system representation graph which depicts the system structure. Finally, graphical distinguishability conditions are given in terms of the system representation graph.

5.1 System structure

We use the word 'system structure', roughly speaking, to represent how the system components are connected with each other. The following argument clarifies the meaning more clearly.

In the case of linear steady systems, $a_{ij} \neq 0$ (where $A = (a_{ij})$) means that the variable x_j (j-th component of the state vector x) is related to i-th component of the system. Since the elements of A are the functions of the system parameters, zeros in A are assumed to be fixed. Therefore the relation between the variables and the components are inherent in the system in the sense that it is independent of the system parameters.

In the case of compartmental systems, the non-zero rate constant k_{ij} implies that there is material flow from compartment j to compartment i (to environment if $i = 0$). Since the material flows in the system are known a priori, the zero rate constants are fixed. In this manner, the existence of material flows is inherent in the system.

In the case of linear dynamical systems in state space form, $a_{ij} \neq 0$ means that the variable x_j is related to the derivative of x_i. This relation is considered to be inherent in the system as in the case of steady system.

Finally, in the case of descriptor form, $e_{ij} \neq 0$ ($E = (e_{ij})$) or $a_{ij} \neq 0$ means that the variable x_j is related to i-th component of the system. This relation is considered to be inherent in the system.

5.2 System representation graph

It is of crucial importance to utilize graphs both in theory

and application. Indeed, the use of graphs enables us to present graphical distinguishability conditions and to carry out the fault diagnosis algorithm with the aid of graphical techniques.

We use the following graphs. Although we use the same letter to denote the different sets of graphs, there is no ambiguity since the system under consideration is evident.

(1) Linear steady system

We use two representation graphs associated with the matrix A, the Coates graph $G_A(N,E)$ and the bipartite graph $G_b(N_r,N_c,E_b)$. The node set N of G_A has natural correspondence with the index set $\mathbf{n} = \{1,2,\ldots,n\}$ of the state vector x. Without ambiguity, we use the same letter C to denote the set of indices $C \subset N$ corresponding to the observed variable. A directed edge (i,j) belongs to E if and only if $a_{ji} \neq 0$. For $G_b(N_r,N_c,E_b)$, the node set N_r (resp. N_c) has natural correspondence with the index set $\mathbf{n}$ of the row (resp. column) of A. There are self-evident one-to-one correspondences $\pi_r:N \to N_r$ and $\pi_c:N \to N_c$. An undirected edge $(\pi_r(i),\pi_c(j)) \in E_b$ if and only if $(j,i) \in E$. We assume the equations are arranged so that the diagonal elements of A are not zero since it is always possible to do so.

(2) Compartmental systems

We use the compartmental graph $G_c(N,E)$. The node set N has natural correspondence with the set of compartment (and environment if the system is open). A directed edge (i,j) belongs to E if and only if $k_{ji} \neq 0$.

(3) State space form

We use two representation graphs associated with the matrix A, the Coates graph $G_A(N,E)$ and the modified bipartite graph $G_b(N_r,N_c,E_b)$. The Coates graph is the same as for steady systems. For $G_b(N_r,N_c,E_b)$, the node set N_r (resp. N_c) has natural correspondence with the index set $\mathbf{n}$ of the row (resp. column) of A. There are self-evident one-to-one correspondences $\pi_r:N \to N_r$ and $\pi_c:N \to N_c$. An undirected edge $(\pi_r(i),\pi_c(j)) \in E_b$ if and only if either $(j,i) \in E$ or $i = j$.

(4) Descriptor form

We use the bipartite graph $G_b(N_r,N_c,E)$ associated with the matrices A and E. The node sets N_r and N_c are defined similar to those of the bipartite graph for steady systems. An undirected edge $(\pi_r(i),\pi_c(j)) \in E_b$ if and only if either $a_{ij} \neq 0$ or $e_{ij} \neq 0$.

5.3 Graphical distinguishability conditions

We are now ready to present graphical distinguishability conditions. We can point out the following advantages of the graphical representation;

(1) we can represent distinguishability conditions independent of the system parameters;

(2) we can design the measuring points in view of (1); and

(3) it is essential in carrying out the fault diagnosis algorithm at two stages --- to check the k-distinguishability assumed in the algorithm and to find a k-cover defined in the algorithm.

It suffices to derive a graphical condition on which the dimension of the fault observation subspace for arbitrary k-fault is k. This is because if we can measure the dimension of the fault observation subspace on the graph, distinguishability conditions are obtained since Proposition 1 holds.

The conditions are generic conditions since the property must be independent to the system parameters. A generic property holds for almost all parameters [6]. We assume that the system parameters are (1) the nonzero elements of A for linear stedy systems, (2) the nonzero rate constants for compartmental systems, (3) the nonzero elements of A for state space form, and (4) the nonzero elements of E and A for descriptor form.

Previous to stating the results, some graphical terms are in order. In $G_A(N,E)$, a <u>path</u> P from v_0 to v_k is an alternate sequence of nodes and edges,

$$P = (v_0, (v_0,v_1), v_1, (v_1,v_2),\ldots, (v_{k-1},v_k), v_k),$$
$$v_i \in N \ (i=0,\ldots,k), \quad (v_i,v_{i+1}) \in E \ (i=0,\ldots,k-1),$$

where v_0 is the <u>initial node</u>, v_k is the <u>terminal node</u>, and k (the number of edges in the sequence) is the <u>length</u> of the path. If $v_k \in C$, then P is called a <u>path from v_0 to C</u>. A path from B to C is similarly defined. For a path P, $N(P) \subset N$ is the node set and $E(P) \subset E$ is the edge set in the path. A set of paths is said to be <u>node-disjoint</u> if no two paths in the set have a common node. A set of paths $\mathcal{P}$ is said to <u>join</u> $N_1 \subset N$ and $N_2 \subset$

N, if the set of the initial nodes (resp. the terminal nodes) of the paths in $\mathcal{P}$ is equal to N_1 (resp. N_2).

In $G_b(N_r,N_c,E_b)$, a matching between $N_r' \subset N_r$ and $N_c' \subset N_c$ is a matching of which edges are incident to nodes in $N_r' \cup N_c'$. A node $v \in N_r \cup N_c$ is said to be <u>saturated</u> by a matching M if there is an edge in M incident to v.

In $G_c(N,E)$, a rooted tree is a connected subgraph of G_c which has a node called the root of the tre to which every other node in the tree is accessible. An isolated node is regarded as a rooted tree. For $M \subset N$, an M-rooted forest W is a collection of rooted trees of which root is in M, and every node in N is contained in exactly one rooted tree in W. To be accurate, a 'spanning' forest is more appropriate, but we omit 'spanning' for simplicity. The edge set in W is denoted by the same letter W since there is no ambiguity.

(1) Linear steady systems

We can identify fault J with the subset of the node set N, and therefore we use the same letter J to denote the subset corresponding to the fault.

[Theorem 2-1] Let J be a k-fault. Then the following four statements are equivalent.

(i) $k = \text{g-dim } \mathcal{F}(J)$.

(ii) We can join J and $C_1 \subset C$ with $|C_1| = k$ by a set of node-disjoint paths in G_A.

(iii) There is a node set D, $|D| = k$, such that, by removing D in G_A, there is no path from $J \setminus D$ to $C \setminus D$, and

there is no such set whose cardinality is less than k.

(iv) There is a matching between $\pi_r(\bar{J})$ and $\pi_c(\bar{C})$ in G_b by which $\bar{C}$ is saturated, where the bar denotes the complement.

(2) Compartmental systems

We can identify fault J with the subset of the edge set E. We use the same letter to denote the subset.

(2-a) Closed system.

[Theorem 2-2-a] Let J be a k-fault which contains no undirected circuits. Then g-dim $\mathcal{F}(J) = k$ if and only if there are a node q and a $(C \cup \{q\})$-rooted forest W such that $W \cup E$ contains no undirected circuits.

(2-b) Open system.

[Theorem 2-2-b] Let J be a k-fault which contains no undirected circuits. Then g-dim $\mathcal{F}(J) = k$ if and only if there is a $(C \cup \{0\})$-rooted forest W such that $W \cup E$ contains no undirected circuits.

(3) State space form

We can identify fault J with the subset of the node set N, and therefore we use the same letter J to denote the subset corresponding to the fault.

[Theorem 2-3] Let J be a k-fault. Then the following four statements are equivalent.

(i) $k = $ g-dim $\mathcal{F}(J)$.

(ii) We can join J and $C_1 \subset C$ with $|C_1| = k$ by a set of node-disjoint paths in G_A.

(iii) There is a node set D, $|D| = k$, such that, by

removing D in G_A, there is no path from $J \setminus D$ to $C \setminus D$, and there is no such set whose cardinality is less than k.

(iv) There is a matching between $\pi_r(\bar{J})$ and $\pi_c(\bar{C})$ in G_b by which $\bar{C}$ is saturated.

(4) Descriptor form. We can identify fault J with the subset of the node set N, and therefore we use the same letter J to denote the subset coresponding the fault.

[Theorem 2-4] Let J be a k-fault. Then g-dim $F(J) = k$ if and only if there is a matching between $\pi_r(\bar{J})$ and $\pi_c(\bar{C})$ in G_b by which $\bar{C}$ is saturated.

6. Conclusion.

Fault distinguishability conditions and a fault diagnosis algorithm are derived. The distinguishability conditions are given in terms of the system representation graph which depict the system structure.

The algorithm is based on linear algebraic consideration, and this simplifies computational procedures.

The graphical conditions possess the following adantages;

(1) we can represent distinguishability conditions independent of the system parameters;

(2) we can design the measuring points in view of (1); and

(3) it is essential in carrying out the fault diagnosis algorithm at two stages --- to check the k-distinguishability assumed in the algorithm and to find a k-cover defined in the algorithm.

References

[1] R.M.Biernacki, and J.W.Bandler, "Multiple-Fault Location of Analog Circuits," IEEE Trans. Circuits & Syst., CAS-28-5, pp.361-367, 1981.

[2] Z.F.Huang, C.S.Lin, and R.W.Liu, "Node-Fault Diagnosis and a Design of Testability," IEEE Trans. Circuits & Syst., CAS-30-5, pp.257-265, 1983.

[3] C.S.Lin, Z.F.Huang, and R.W.Liu, "Topological Conditions for Single-Branch-Fault," IEEE Trans. Circuits & Syst., CAS-30-6, pp.376-381, 1983.

[4] Y.Ohta, H.Maeda, S.Kodama, and S.Takeda, "Faults Diagnosis and Design of Measuring Set of Linear Systems," submitted to IEEE Trans. Circuits & Syst.

[5] Y.Ohta, H.Maeda, S.Kodama, and T.Itoh, "Fault Diagnosis of Compartmental Systems," Trans. IECE, vol.67-A, No.8, pp.833-840, 1984. (in Japanese.)

[6] R.W.Shields, and J.B.Pearson, "Structural Controllability of Multiinput Linear Systems," IEEE Trans. Automat. Contr., AC-21-2, pp.203-212, 1976.