A NEW CONDITION NUMBER OF THE EIGENVALUE AND ITS APPLICATION IN CONTROL THEORY *

YANG YA-GUANG
(*Zhejiang University, Hangzhou, Zhejing, China*)

Abstract

Since the present condition numbers of the eigenvalue are not convenient for control systems design, this paper puts forward a new condition number and indicates its application in designing robust control systems.

1. Introduction

A most important tend in the development of control theory is the design of robust systems [1]. However, the system robustness is heavily related to the condition number of the eigenvalue. If we can design a system which has not only satisfactory properties (such as having prescribed eigenvalues), but also a good condition number, then, when the model is exactly equal to the actual system, the system will have a good working properties, and when there exist errors between the model and the system, the system performance will not be too bad. This field has attracted many control theory scholars and engineers [2], [3]. But the existing condition numbers of the eigenvalues are not convenient for designing systems, so the present methods (such as [4], [5]) cannot ensure the condition number of the eigenvalue of the system to be an optimal condition number in actual constraints. This paper proposes a new condition number for solving the problem.

2. Main Result

It is well known that, if $X^{-1}AX = J$, where J is a Jordan matrix, then the condition number of the eigenvalue of matrix A becomes $\|X^{-1}\| \cdot \|X\|$. The most popular condition number is the spectral condition number, $K_2(A) = \sigma_1/\sigma_n$, where σ_n, σ_1 are the smallest singular value and the greatest singular value of matrix X, respectively. But σ_1 and σ_n cannot be determined if the matrix X is not determined. He [6] gives another condition number as follows:

$$K(A) = (E[X_k X_1 X_2 \cdots X_{k-1}])^{-1}$$

* Received June 30, 1986.

where $X_i (i = 1, \cdots, k)$ are column vectors of X, which are standardized, $E[X_k X_1 \cdots X_{k-1}] = [G(X_1 \cdots X_k)/G(X_1 \cdots X_{k-1})]^{-1}$, G is a Gram determinant, and the subscript k equals ε-rank of the matrix X. Obviously, the definition is too complex to compute. we give another definition.

Definition. *Suppose A is matrix of order n, and $X_i (i = 1, \cdots, n)$ are standardized eigenvectors or generalized eigenvectors of matrix A. $X = (X_1, X_2, \cdots, X_n)$. Define the condition number of the eigenvalue of matrix A as*

$$K(A) = \frac{1}{|X^T X|}.$$

New, we will prove that the definition is reasonable. First, we introduce the following

Lemma [7]. *Suppose that the real coefficient polynomial $f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$ with $a_n > 0$ and a_{n-k} being the first negative , and B is the greatest value among the absolute values of the negative coefficients. Then*

$$N = 1 + \sqrt[k]{B/a_n}$$

is an upper bound of the roots of $f(x)$.

Theorem. *Suppose X is a matrix composed of standardized eigenvectors and generalized eigenvectors of A. $X = (X_1, X_2, \cdots, X_n)$. Note $\Delta = \det(X^T X)$. Then*

$$\sigma_n^2 \geq \frac{1}{1 + \frac{1}{\Delta}\left(\frac{n}{n-1}\right)^{n-1}}$$

where $\sqrt{\sigma_n^2}$ is the minimal singular value of X.

Proof . Note $\lambda_i = \sigma_i^2$. Since λ_i are eigenvalues of $X^T X$, so

$$(X^T X)Y = Y \cdot \text{ diag } (\sigma_1^2, \sigma_2^2, \cdots, \sigma_n^2), \tag{1}$$

$$\sigma_i^2 \geq \sigma_j^2, \quad i > j, \tag{2}$$

$$\prod_{i=1}^n \sigma_i^2 = \prod_{i=1}^n \lambda_i = X^T X = \Delta. \tag{3}$$

Since X_i are standardized vectors, so

$$\sum_{i=1}^n \lambda_i = \sum_{i=1}^n \sigma_i^2 = \text{ tr } (X^T X) = n. \tag{4}$$

Note

$$K_1 = \Big\{\sigma_i | \sigma_i \text{ satisfy } (1), (2), (3), (4)\Big\},$$
$$K_2 = \Big\{\sigma_i | \sigma_i \text{ satisfy } (3), (4)\Big\}.$$

We have $\min_{\sigma_i \in K_1} \sigma_n^2 > \min_{\sigma_i \in K_2} \sigma_n^2$.

So we only require estimating the lower bound of $\min_{\sigma_i \in K_2} \sigma_n^2$. Constructing the Lagrangian function for $\min_{\sigma_i \in K_2} \sigma_n^2$, we have

$$L = \sigma_n^2 + \beta_0 (\prod_{i=1}^{n} \sigma_i^2 - \Delta) + \beta_1 (\sum_{i=1}^{n} \sigma_i^2 - n), \tag{5}$$

$$\frac{\partial L}{\partial \sigma_i} = 2\beta_0 \prod_{j \neq i}^{n} \sigma_j^2 \sigma_i + 2\beta_1 \sigma_i = 0, \tag{6}$$

$$\frac{\partial L}{\partial \sigma_n} = 2\sigma_n + 2\beta_0 (\prod_{j \neq n} \sigma_j^2 \sigma_n) + 2\beta_1 \sigma_n = 0, \tag{7}$$

$$\frac{\partial L}{\partial \beta_0} = \prod_{i=1}^{n} \sigma_i^2 - \Delta = 0, \tag{8}$$

$$\frac{\partial L}{\partial \beta_1} = \sum_{i=1}^{n} \sigma_i^2 - n = 0, \tag{9}$$

and

$$\begin{aligned} \frac{\partial L}{\partial \sigma_n} - \frac{\partial L}{\sigma_i} &= 2\sigma_n + 2\beta_0 \prod_{j \neq i,n} \sigma_j^2 (\sigma_n \sigma_i^2 - \sigma_n^2 \sigma_i) + 2\beta_1 (\sigma_n - \sigma_i) \\ &= 2\sigma_n + 2[\beta_0 \prod_{j \neq i,n} \sigma_j^2 \sigma_n \sigma_i - \beta_1](\sigma_i - \sigma_n) = 0. \end{aligned} \tag{10}$$

Since $|X^T X| \neq 0$, according to (1), $\sigma_n \neq 0$; so

$$\sigma_i \neq \sigma_n, \quad \forall i \neq n, \tag{11}$$

$$\frac{\partial L}{\partial \sigma_i} - \frac{\partial L}{\partial \sigma_j} = (2\beta_0 \prod_{k \neq i,j} \sigma_k^2 \sigma_j \sigma_i - 2\beta_1)(\sigma_j - \sigma_i) = 0. \tag{12}$$

Therefore

$$\sigma_j = \sigma_i, \quad i \neq j, \quad \forall i, j \neq n, \tag{13}$$

$$\beta_1 / \beta_0 = \prod_{k \neq i,j} \sigma_k^2 \sigma_i \sigma_j, \quad \forall i, j, k \neq n, i \neq j \neq k. \tag{14}$$

By symmetry, equation (14) is equivalent to equation (13). So, all the solutions of (12) are in equation (13).

Substituting (11), (13) into (8) , (9) and letting

$$\lambda = \lambda_i, \quad i \neq n$$

we have

$$\lambda^{n-1} \lambda_n = \Delta, \tag{15}$$

$$(n-1)\lambda+\lambda_n-n=0,\quad \lambda=\frac{n-\lambda_n}{n-1}. \tag{16}$$

Substituting (16) into (15), we have

$$f(\lambda_n)=(n-\lambda_n)^{n-1}\lambda_n-\Delta(n-1)^{n-1}=0. \tag{17}$$

Since equations (6)–(9) are necessary conditions of the problem

$$\min_{\sigma_i\in K_2}\sigma_n^2=\min_{\sigma_i\in K_2}\lambda_n, \tag{18}$$

if λ_n^* is a solution of (18) , then λ_n^* satisfies equations (6)–(9) or equivalently, (17). As λ_n^* is a positive number, it must be greater than or equal to the minimal positive root of (17). Therefore, we only require estimating the lower bound of the minimal positive root of (17).

Considering $\varphi(\lambda_n)=\lambda_n^n f(\frac{1}{\lambda_n})$, if α is an arbitrary positive root of $f(\lambda_n)$, then $\frac{1}{\alpha}$ is a positive root of $\varphi(\lambda_n)$. If N is an upper bound of the positive roots of $\varphi(\lambda_n)$, then $\frac{1}{N}$ is a lower bound of the positive roots of $f(\lambda_n)$.

$$\begin{aligned}\varphi(\lambda_n)&=\lambda_n^n\left[\frac{1}{\lambda_n}(n-\frac{1}{\lambda_n})^{n-1}-\Delta(n-1)^{n-1}\right]=(n\lambda_n-1)^{n-1}-\Delta(n-1)^{n-1}\lambda_n^n\\&=\Delta(n-1)^{n-1}\lambda_n^n-n^{n-1}\lambda_n^{n-1}+n^{n-2}(n-1)\lambda_n^{n-2}+\cdots+(-1)^n=0.\end{aligned} \tag{19}$$

According to the Lemma, we have

$$N=1+\frac{1}{\Delta}\Big(\frac{n}{n-1}\Big)^{n-1}.$$

So, the solution λ_n^* of (18) satisfies

$$\lambda_n^*\geq\frac{1}{1+\frac{1}{\Delta}\Big(\frac{n}{n-1}\Big)^{n-1}}.$$

For the minimal singular value σ_n of matrix X, we have

$$\sigma_n^2\geq\lambda_n^*\geq\frac{1}{1+\frac{1}{\Delta}\Big(\frac{n}{n-1}\Big)^{n-1}}.$$

Obviously, the smaller $\frac{1}{\Delta}$ is, the greater σ_n may be. He [6] indicates that the main factor of the condition number $K(A)=\sigma_1/\sigma_n$ is σ_n. So the definition of this paper is a reasonable definition. Incidentally, the computation of the condition number in this paper is much simpler than that of [4].

3. Application in Control Theory

Suppose the system equation is

$$\dot{X} = AX + BU,$$

and the state feedback control can be expressed in $U = KX$. Then

$$\dot{X} = (A + BK)X.$$

If the system is completely controllable, we can arbitrarily assign the poles of the system. In the case of multi-inputs, K is not unique, so we can make full use of the surplus degrees of freedom to design robust control system.

Suppose that the prescript poles of the system are $\lambda_1, \lambda_2, \cdots, \lambda_n$. Note $\Lambda = \operatorname{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)$. Then the problem of robust control system design becomes the following nonlinear programming problem:

$$\max |Y^T Y| \tag{20}$$

$$\text{subject to} \quad (A + BK)Y = Y\Lambda, \tag{21}$$

$$(y_i, y_i) = 1, \quad i = 1, n \tag{22}$$

where y_i is the ith column vector of matrix Y.

Since (21), (22) are nonlinear equality constraints, Himmelblau [9] points out that (20)–(22) is the most difficult kind of nonlinear programming. But equation (21) can be rewritten as

$$AY + BF = Y\Lambda \tag{21'}$$

where $F = KY$. So the problem becomes

$$\max |Y^T Y|$$

$$\text{subject to} \quad AY + BF = Y\Lambda,$$

$$(y_i, y_i) = 1.$$

Since (21') are linear equality constraints, the problem becomes easier than the original nonlinear programming. Computational examples show that the computing time of problem (20), (21'), (22) is much more shorter than that of problem (20)–(22).

When the optimal solution F^*, Y^* is obtained, the optimal feedback matrix $K^* = F^*(Y^*)^{-1}$.

Example 1 [4].

$$\dot{X} = \begin{pmatrix} 1.38 & -0.2077 & 6.715 & -5.676 \\ -.5814 & -4.29 & 0 & 0.675 \\ 1.067 & 4.273 & -6.654 & 5.893 \\ 0.048 & 4.273 & 1.343 & -2.104 \end{pmatrix} X + \begin{pmatrix} 0 & 0 \\ 5.679 & 0 \\ 1.136 & -3.146 \\ 1.136 & 0 \end{pmatrix} u .$$

We assign the eigenvalues as

$$(A+BK)=(-0.2,-0.5,-5.0566,-8.6659).$$

The proposed problem is solved by using the flexible tolerance method on IBM-PC.

Take the initial values as

$$Y=0,\quad F=0.$$

The optimal solution is

$$Y^*=\begin{pmatrix} -.4243528 & -0.1413062 & -.5204743 & -.7291125 \\ -.9028558 & .01674664 & -.4015174 & -.1769179 \\ .4815997 & .420886 & -.7074955 & -.3060228 \\ -.4919919 & .652804 & .4034852 & -.4170504 \end{pmatrix},$$

$$F^*=\begin{pmatrix} -.5852064 & .1246379 & .02879013 & -.5042462 \\ -.6547011 & .8269487 & .5313507 & .01483982 \end{pmatrix},$$

$$K^*=\begin{pmatrix} .064747 & -.308957 & -.325121 & .334544 \\ 1.038243 & .577936 & .319642 & -.046510 \end{pmatrix},$$

$\det\,[(Y^*)^T(Y^*)]=0.141545.$

In Table 1, the robust indexes obtained by this paper and kautsky et al. [4] are listed.

Table 1

Index	Nonlinear programming method	The methods of [4]		
		0	1	2/3
Y^TY	0.141545	not	0.105296	0.117939
$K_2(A)$	4.755	convergent	5.19[1)]	5.01109[2)]

Example 2 [4].

$$\dot{X}=AX+Bu,$$

1) The value listed in [4] is 3.32. But the author re-computed the value as equals 5.19.

2) The value listed in [4] is 4.54. But the author re-computed the value as equals 5.01109.

$$A = \begin{pmatrix} -0.1094 & 0.0628 & 0 & 0 & 0 \\ 1.306 & -2.132 & 0.9807 & 0 & 0 \\ 0 & 1.595 & -3.149 & 1.547 & 0 \\ 0 & 0.0355 & 2.632 & -4.257 & 1.855 \\ 0 & 0.00227 & 0 & 0.1636 & -0.1625 \end{pmatrix},$$

$$B^T = \begin{pmatrix} 0 & 0.0638 & 0.0838 & 0.1004 & 0.0063 \\ 0 & 0 & 0.1369 & -0.206 & -0.0128 \end{pmatrix}.$$

We set the eigenvalues equal to $(-0.2, -0.5, -1.0, -1 \pm i)$, which includes a complex conjugate pair. So Λ can be written as

$$\Lambda = \begin{pmatrix} -0.2 & 0 & 0 & 0 & 0 \\ 0 & -0.5 & 0 & 0 & 0 \\ 0 & 0 & -1.0 & 0 & 0 \\ 0 & 0 & 0 & -1.0 & 1.0 \\ 0 & 0 & 0 & -1.0 & -1.0 \end{pmatrix}$$

to avoid complex computation.

The optimal solution is

$$Y^* = \begin{pmatrix} -.200752 & .055914 & -.852951 & -.086023 & .017333 \\ .366047 & -.741101 & .273655 & .174146 & -.070746 \\ .642255 & -.571679 & .204050 & .446763 & -.606369 \\ .625816 & -.255142 & .058082 & .361309 & -.593028 \\ -.146869 & -.236069 & -.391397 & -.795067 & .524714 \end{pmatrix},$$

$$|Y^T Y| = 0.000476,$$

$$K^* = \begin{pmatrix} 49.3295 & 108.9142 & -226.2659 & 195.0873 & -41.64777 \\ -23.2589 & 39.71716 & -65.26995 & 50.58586 & 2.25201 \end{pmatrix}.$$

For this example, the robust indexes obtained by this paper and kautsky et al. [4] are given in Table 2.

Table 2

Index	Nonlinear Programming method	The methods of [4]	
		1	2/3
$K_2(A)$	37.09	39.4	66.1

4. Conclusion

Obviously, the robust indexes obtained by this paper are better than those by [4]. Table 1 also shows that the greater $|Y^TY|$ is, the smaller $K_2(A)$ is. So, the new condition number of the eigenvalue put forward in this paper is not only simple to compute but also reasonable. Using the condition number to design the robust control system can yield a good robust index. But the computing time to search for the optimal solution by this method is very long because it is difficult to find an analytic derivative expression of $|Y^TY|$ for higher order systems, and we have to use the methods without evaluating the derivatives. So searching other efficient methods to solve the problem (20), (21'), (22) may be a lot of work to do.

The author is very grateful to Professor Lu Yong-zai and the referee for ging over the manuscript and giving many useful suggestions.

References

[1] R. V. Kokotovic, Recent trends in feedback design: An overview, *Automatica*, 21 : 3 (1985), 225–236.

[2] B. Kouvaritakis , H. Latchman, Singular-value and eigenvalue techniques in the analysis of systems with structured perturbations, *Int. J. Control*, 41 : 6 (1985), 1381–1412.

[3] M. J. Chen, C. A. Desoer, The problem of guaranteeing robust disturbance rejection in linear multivariable feedback systems, *Int. J. Control*, Vol. 37, No. 2 , 305–313.

[4] J. Kautsky, N. K. Nichols , P. Van Dooren, Robust pole assignment in linear state feedback, *Int. J. Control*, 41 : 5 (1985), 1129–1155.

[5] M. M. Fahmy , H. S. Tantawy, Eigenstructure assignment via linear state-feedback control, *Int. J. Control*, 40 : 1 (1984), 161–178.

[6] He Xu-chu, Fundamental Theory and Computational Methods of the Generalized Inverse Matrix, 1985.

[7] Jiang Er-xiong, Linear Algebra, 1978.

[8] J. H. Wilkinson, The Algebraic Eigenvalue Problems, Clarendon Press, 1965.

[9] David M. Himmelblau, Applied Nonlinear Programming, The University of Texas, Austin, Texas, 1972.

[10] Zhu Ben-ren, Jin Mao-yuan, Zhang Sheng-li, Iterative methods for inverse eigenvalue problems, *Math. Num. Sinica*, 8 : 3 (1986), 314–320.