第16卷 第5期　　自 动 化 学 报　　Vol. 16, No. 5
1990 年 9 月　　ACTA AUTOMATICA SINICA　　Sept., 1990

鲁棒性指标分析及其在多变量控制系统极点配置中的应用

杨亚光　吕勇哉
(浙 江 大 学)

摘　要

本文详细分析了文献[3, 4, 6]中采用的鲁棒性指标的优缺点，进而提出了一种改进的鲁棒极点配置方法. 其收敛性比文[8]的结果好，计算复杂性明显优于文献[4,6,11]的方法.

关键词——鲁棒性指标，多变量系统，极点配置.

一、引　言

鲁棒控制系统的设计已被众多的著名学者公认为是目前最重要的研究领域之一[1,2]. 自 Cavin 等人[3]提出利用多余自由度设计鲁棒极点的方法以来，时域鲁棒极点配置不仅受到了控制理论学者的广泛重视，也引起了数学界的注意[4]. 为了方便鲁棒系统设计，引入了一些鲁棒指标，进而提出了各种最优鲁棒极点配置方法. 遗憾的是，有些指标没有严格的数学证明说明其合理性和局限性. 本文针对 Cavin 给出的指标，详细地讨论了这些问题，并给出了一个有效的鲁棒极点配置方法.

二、Cavin 的指标及相关的工作

由特征值的摄动理论知道，特征值的条件数就是最好的鲁棒性指标. 设 $A\cdot X=X\cdot J$，A 是任意实矩阵，J 是约当阵，X是特征向量和广义特征向量构成的矩阵，$\lambda_i(A)$ 是 A 的特征值，对任意摄动阵 E，$\lambda_i(A+E)$ 满足[5]

$$|\lambda_i(A)-\lambda_i(A+E)|^{r_i}\leqslant h\sum_{l=0}^{r_i-1}(1+h)^l, \tag{1}$$

这里 r_i 是最大约当块的阶数，$h=\|X^{-1}\|_\infty\cdot\|X\|_\infty\cdot\|E\|_\infty$. 显然，$\|X^{-1}\|_\infty\cdot\|X\|_\infty$ 越小，$\lambda_i(A)$ 对摄动阵 E 的变化越小. 鉴于 $\|\cdot\|_2$ 范数有类似的性质，最常用的还是谱条件数

$$k_2(X)=\|X^{-1}\|_2\cdot\|X\|=\sigma_1/\sigma_n, \tag{2}$$

这里 σ_1、σ_n 分别是X的最大、最小奇异值.

本文于 1988 年 3 月 22 日收到.

由于这些条件数在鲁棒极点配置的计算中都很不方便，Cavin 等人[3]提出了如下指标：

$$k_c(X) = \mathrm{tr}[I - X^TX]^2, \tag{3}$$

并给出了实极点的鲁棒配置方法. 涂健和周军[6]利用这一指标讨论了复极点的配置和二次型调节器问题. 遗憾的是他们都没有证明指标的合理性. 孙继广[4]最近给出了一个非常有效的算法并利用了如下指标：

$$k_s = \|I - X^TX\|_F, \tag{4}$$

这里 $\|\cdot\|_F$ 表示 Frobenius 范数. 由文[7]知，k_c 和 k_s 等价. 尽管在仿真中，这一指标非常有效，但仍然没有给出合理性证明. 下一节将证明指标的合理性和局限性. 在此基础上，第四节给出了一种更为有效的改进算法.

三、鲁棒性指标 $K_c(X)$ 的分析

这一节将给出几个主要定理.

定理 1. 假定X的列 x_j, $j = 1, \cdots, n$ 是标准化向量，则有

$$k_2^2(X) \leqslant \min\left\{n, 1 + \sqrt{\frac{n-1}{n}k_c}\right\}\Big/\max\left\{0, 1 - \sqrt{\frac{n-1}{n}k_c}\right\}. \tag{5}$$

证明. 记 $\{\lambda_1, \cdots, \lambda_n\}$ 是 (X^TX) 的特征值，σ_i, $i = 1, \cdots, n$ 是X的奇异值，那么

$$\lambda_i = \sigma_i^2, \tag{6}$$

$$(X^TX)\boldsymbol{y}_i = (\lambda_i\boldsymbol{y}_i). \tag{7}$$

于是

$$(X^TX - I)\boldsymbol{y}_i = (\lambda_i - 1)\boldsymbol{y}_i, \tag{8}$$

$$(X^TX - I)^2\boldsymbol{y}_i = (\lambda_i - 1)^2\boldsymbol{y}_i, \tag{9}$$

所以 $(\lambda_i - 1)^2$ 是 $(I - X^TX)^2$ 的特征值. 记 $Y = (\boldsymbol{y}_1, \cdots, \boldsymbol{y}_n)$ 因此

$$(X^TX)Y = Y\cdot\mathrm{diag}(\lambda_1, \cdots, \lambda_n), \tag{10}$$

且

$$\lambda_i \geqslant \lambda_j, \; i > j. \tag{11}$$

因为X是标准化列向量的矩阵，所以

$$\sum_{i=1}^n \lambda_i = n. \tag{12}$$

由(9)式得

$$\sum_{i=1}^n (\lambda_i - 1)^2 = \mathrm{tr}[X^TX - I]^2 = k_c. \tag{13}$$

记 $k_1 = \{\lambda_i | \lambda_i$ 满足(10),(11),(12),(13)$\}$, $k_2 = \{\lambda_i | \lambda_i$ 满足(12),(13)$\}$，显然X的最小奇异值满足

$$\sigma_n^2 = \min_{\lambda_i \in k_1}\lambda_n \geqslant \min_{\lambda_i \in k_2}\lambda_n \triangleq \lambda_n^*. \tag{14}$$

对

$$\min_{\lambda_i \in k_2} \lambda_n, \tag{15}$$

构造拉格朗日函数得

$$L = \lambda_n + \beta_0\left(\sum_{i=1}^{n} \lambda_i - n\right) + \beta_1\left(\sum_{i=1}^{n} (\lambda_i - 1)^2 - k_c\right), \tag{16}$$

于是

$$\frac{\partial L}{\partial \lambda_n^*} = 1 + \beta_0 + 2\beta_1(\lambda_n^* - 1) = 0, \tag{17}$$

$$\frac{\partial L}{\partial \lambda_i} = \beta_0 + 2\beta_1(\lambda_i - 1) = 0, \quad i \neq n, \tag{18}$$

$$\frac{\partial L}{\partial \beta_0} = \sum_{i=1}^{n} \lambda_i - n = 0, \tag{19}$$

$$\frac{\partial L}{\partial \beta_1} = \sum_{i=1}^{n} (\lambda_i - 1)^2 - k_c = 0, \tag{20}$$

$$\frac{\partial L}{\partial \lambda_n^*} - \frac{\partial L}{\partial \lambda_i} = 1 + 2\beta_1(\lambda_n^* - \lambda_i) = 0, \quad i \neq n. \tag{21}$$

由(21)，$\beta_1 \neq 0$，$\lambda_n^* \neq \lambda_i$.　　(22)

由(22)式及(18)式，得

$$\lambda_i = \lambda_j, \quad i \neq n, \quad j \neq n, \tag{23}$$

于是令

$$\lambda = \lambda_i, \quad i \neq n, \tag{24}$$

由(19)式得

$$\lambda = (n - \lambda_n^*)/(n - 1). \tag{25}$$

(25)式代入(20)式并化简，得

$$\lambda_n^* = 1 \pm \sqrt{\frac{n-1}{n} k_c}.$$

由(12)、(14)式知，$\lambda_n^* \leqslant \sigma_n^2 \leqslant 1$，因此

$$\lambda_n^* = 1 - \sqrt{\frac{n-1}{n} k_c}. \tag{26}$$

因为 $\sigma_n \geqslant 0$，于是有

$$\sigma_n^2 \geqslant \max\left\{1 - \sqrt{\frac{n-1}{n} k_c}, \ 0\right\}. \tag{27}$$

对极大奇异值 λ_1 有

$$\sigma_1^2 = \max_{\lambda_i \in k_c} \lambda_1 \leqslant \max_{\lambda_i \in k_2} \triangleq \lambda_1^*. \tag{28}$$

类似前面的推导，并注意到(19)式，有

$$\sigma_1^2 \leqslant n, \tag{29}$$

容易证明

$$\sigma_1^2 \leqslant \min\left\{1 + \sqrt{\frac{n-1}{n}k_c},\ n\right\}. \tag{30}$$

由(27)、(30)式，结论是明显的．证毕．

定理 1 表明，在 $k_c \in \left[0, \frac{n}{n-1}\right)$ 区间内，k_c 越小则 $k_2(X)$ 越小，也就是说矩阵 A 对任意摄动阵 E 来说，特征值的鲁棒性越好．

在实际设计中，X 的列向量标准化意味着增加 n 个等式约束，将导致问题求解的复杂化．为此给出：

定理 2. 对任意 n 阶方阵 X 都有

$$k_2^2(X) \leqslant (1 + \sqrt{k_c(X)})/\max\{1 - \sqrt{k_c(X)},\ 0\}. \tag{31}$$

证明．令 $k_2' = \{\lambda_i | \lambda_i$ 满足(13)式$\}$，用 k_2' 代替 k_2，完全类似定理 1 的证明，容易得到(31)式．证毕．

注 1. 由于去掉了列向量标准化的限制，$k_c(X)$ 的适用范围由 $\left[0, \frac{n}{n-1}\right)$ 缩小到 $[0, 1)$．同时 X 的列向量没有标准化时，k_c 不必满足(12)式，只需满足(13)式，因此 k_c 没有上界限制，故这种情况下作为合理的鲁棒性指标，k_c 的适用范围很小．因此，文[6]的算法具有较大的局限性，其算例之所以很好是因为它们的 $k_c(X)$ 均为零．

定理 3. 在 X 的列向量标准化的情况下，

$$k_c(X) \leqslant n \cdot (n-1). \tag{32}$$

证明．因为 X 是标准化列向量方阵，故只需考虑

$$\max \sum_{i=1}^{n} (\lambda_i - 1)^2 = k_c, \quad \text{sub. to} \sum_{i=1}^{n} \lambda_i = n.$$

由动态规划的最优性原理

令
$$f_2(a) = \max_{0 \leqslant \lambda \leqslant 2}\{(\lambda - 1)^2 + (a - \lambda - 1)^2\},$$

得
$$\lambda = 0, f_2(a) = (a-1)^2 + 1|_{a=2} = (2-1)^2 + 1.$$

令
$$\begin{aligned} f_3(a) &= \max_{0 \leqslant \lambda \leqslant 3}\{(\lambda - 1)^2 + f_2(a - \lambda)\} \\ &= \max_{0 \leqslant \lambda \leqslant 3}\{(\lambda - 1)^2 + (a - \lambda - 1)^2 + 1\}, \end{aligned}$$

得
$$\lambda = 0,\ f_3(a) = (a-1)^2 + 2|_{a=3} = (3-1)^2 + 2.$$

设
$$\begin{aligned} f_{n-1}(a) &= \max_{0 \leqslant \lambda \leqslant n-1}\{(\lambda - 1)^2 + f_{n-2}(a - \lambda)\} \\ &= (a-1)^2 + n - 2|_{a=n-1} = (n-2)^2 + (n-2), \end{aligned}$$

由归纳法

$$\begin{aligned} f_n(a) &= \max_{0 \leqslant \lambda \leqslant n}\{(\lambda - 1)^2 + f_{n-1}(a - \lambda)\} \\ &= \max_{0 \leqslant \lambda \leqslant n}\{(\lambda - 1)^2 + (a - \lambda - 1)^2 + (n-2)\}, \end{aligned}$$

得

$$\lambda = 0,\ f_n(a) = (a-1)^2 + n - 1|_{a=n} = (n-1)^2 + n - 1 = (n-1)n.$$

此时 $\lambda_1 = n$，$\lambda_2 = \lambda_3 = \cdots = \lambda_n = 0$．证毕．

注 2. 定理 1 给出 $k_c(X)$ 的适用范围是 $\left[0, \dfrac{n}{n-1}\right)$，而 $k_c(X)$ 的可能取值范围是 $[0, n(n-1))$。显然 $n \geqslant 2$ 时有 $n(n-1) \geqslant n/(n-1)$，是否在区间 $[n/(n-1), n(n-1))$ 中 $k_c(X)$ 的大小不能反映 $k_2(X)$ 的大小，也就是不能反映 A 的特征值对摄动矩阵的鲁棒性呢？回答是肯定的，举一例就可以说明。

例 1. 当 $k_c(X) = n/(n-1)$ 时，设 X 是列向量标准化的矩阵，取

$$\lambda_1 = \lambda_2 = \cdots = \lambda_{n-1} = n/(n-1), \quad \lambda_n = 0.$$

显然，

$$\sum_{i=1}^{n} \lambda_i = n, \text{ 且 } \sum_{i=1}^{n} (\lambda_i - 1)^2 = \frac{n}{n-1}, \text{ 而}$$

$$k_2(X) = \sigma_1^2/\sigma_n^2 = \lambda_1/\lambda_n = \infty,$$

因此，文[4]的方法虽然考虑了 X 的列向量标准化，使得指标的适用范围增大，但仍有其局限性。从数值实例来看，其实例之所以成立，是因为 $k_c(X) \leqslant n/(n-1)$。

四、鲁棒极点配置的设计方法

设有多变量系统

$$\dot{\boldsymbol{x}} = A\boldsymbol{x} + B\boldsymbol{u}, \tag{33}$$

由状态反馈得

$$\dot{\boldsymbol{x}} = (A + Bk)\boldsymbol{x} + B\boldsymbol{v} \tag{34}$$

由于多变量反馈阵 k 的非唯一性，可以利用多余的自由度在保证闭环极点的情况下，使闭环系统的鲁棒性达到最好。记

$$(A + Bk)Q = \Lambda Q, \tag{35}$$

这里 $\Lambda = \mathrm{diag}(\lambda_1, \cdots, \lambda_n)$，$\lambda_i$，$i = 1, \cdots, n$ 是指定的闭环极点，Q 是闭环特征向量构成的矩阵。Kautsky 等人[8]给出了如下重要的结果。

定理 4[8]. 对于指定的闭环特征值 λ_i，闭环系统 $(A + Bk)$ 的对应于 λ_i 的特征向量 q_i 属于子空间

$$\varphi_i = \mathscr{N}\{\boldsymbol{u}_1^T(A - \lambda_i I)\}, \tag{36}$$

这里 $\mathscr{N}(\cdot)$ 表示零空间，u_1 由下式决定：

$$B = [U_0 \boldsymbol{u}_1] \begin{bmatrix} Z \\ 0 \end{bmatrix}, \tag{37}$$

其中 $U = [U_0, \boldsymbol{u}_1]$ 是正交阵，Z 是非奇阵。

鲁棒极点配置的主要步骤可以归结为

Step A. 用 QR 分解 B 和 $[\boldsymbol{u}_1^T(A - \lambda_i I)]^T$ 阵得

$$B = [U_0, \boldsymbol{u}_1] \begin{bmatrix} Z \\ 0 \end{bmatrix},$$

$$[\boldsymbol{u}_1^T(A - \lambda_i I)]^T = [\hat{S}_i, S_i] \begin{bmatrix} R_i \\ 0 \end{bmatrix}. \tag{38}$$

容易证明 S_i 是 φ_i 的一组基,故

$$q_i = S_i D_i, \tag{39}$$

D_i 是 r 阶的列向量,r 是 B 的列数.

Step X. 选择向量 $q_i = S_i D_i \in \varphi_i$, 使得 $Q = (q_1, q_2, \cdots, q_n)$ 是鲁棒的.

Step F. $$K = Z^{-1} U_0^T (Q \Lambda Q^{-1} - A). \tag{40}$$

很明显, Step X 是最重要的一步. 文[8]的方法是轮流优化 q_i, 这样的收敛点肯定满足

$$f(q_1^*, \cdots, q_n^*) = \min_{q_i \in \varphi_i} f(q_1^*, \cdots, q_{i-1}^*, q_i, q_{i+1}^*, \cdots, q_n^*)$$

对任意 i 都成立. 然而文[9]指出仅满足上述条件的收敛点不一定是一个最优点. 文[4]对 Step X 提出了一个修正的办法, 但由于 q_i 的标准化使得问题变成了一个非线性等式约束的非线性规划问题,而此类问题是非线性规划中最为复杂的一类问题[10],求解相当麻烦. 为此本文提出一个修正的指标.

显然, Step X 可以表示成

$$\min \operatorname{tr}[I - Q^T Q]^2, \tag{41}$$

$$\text{sub. to } q_i^T q_i = 1. \tag{42}$$

(41)式化简得

$$\operatorname{tr}[I - Q^T Q]^2 = \sum_{i=1}^{n} \sum_{j \neq i}^{n} [(q_i^T q_j)^2 + (1 - q_i^T q_i)^2]. \tag{43}$$

由罚函数原理,可以得修正的鲁棒性指标

$$k_y = \sum_{i=1}^{n} \sum_{j \neq i}^{n} [(q_i^T q_j)^2 + \beta(1 - q_i^T q_i)^2] \tag{44}$$

代替(43)式. 在 β 充分大时,优化(44)式等效于求解非线性规划(41)、(42)式, 但计算简单得多. 在 $\beta \to \infty$ 时,显然 $k_y \in [0, n(n-1)]$ 且适用范围在 $[0, n/(n-1))$ 区间, 大大优于 $k_c \in [0, \infty)$, 且只适用在[0,1) 区间内. 由于主要计算任务在求解无约束优化问题上,算法有效性的关键也在这里. 因此,比较用非线性规划的文[4]、[6]和[11]来说,本文的方法具有计算量小、适用范围大的特点. 如前述,文[8]的方法不一定能收敛到最优点. 下面给出本方法的计算步骤.

定理 5. 设

$$k_y = \sum_{i=1}^{n} \sum_{j \neq i}^{n} [(D_i^T S_i^T S_j D_j)^2 + \beta(1 - D_i^T S_i^T S_i D_i)^2], \tag{45}$$

则

$$\frac{dk_y}{dD_i} = 4\left[\sum_{j \neq i}^{n} (D_i^T S_i^T S_j D_j)^2 S_j^T S_i + \beta(D_i^T D_i - 1) D_i\right]. \tag{46}$$

证明. 代入(39)到(44)式,并注意到(38)式,有 $(S_i^T S_i) = I$, 即可得证.

由此, Step X 可以用各种无约束梯度方法方便地求解,如共轭梯度法、变尺度法等. 以共轭梯度法为例,鲁棒极点配置方法的步骤归纳如下:

1) 由 QR 分解方法得

$$B = [U_0 \boldsymbol{u}_1] \begin{bmatrix} Z \\ 0 \end{bmatrix},$$

$$[\boldsymbol{u}_1^T(A - \lambda_i I)]^T = (\hat{S}_i, S_i) \begin{bmatrix} R_i \\ 0 \end{bmatrix},$$

其中 S_i 是 φ_i 空间的一组正交基,整个计算约需 $O(n^3m)$ 次运算[8].

2）给定初始 $D^{(0)} = (D_1^{(0)}, \cdots, D_n^{(0)})$，收敛误差允许范围 ε.

3）计算

$$\boldsymbol{P}^{(0)} = -\left(\frac{dk_y(D^{(0)})}{dD_1}, \cdots, \frac{dk_y(D^{(0)})}{dD_n}\right)^T \triangleq -\nabla k_y^{(0)},$$

$$\boldsymbol{P}^{(k+1)} = -\nabla k_y^{(k+1)} + r_k \boldsymbol{P}^{(k)},$$

$$r_k = \frac{\nabla k_y^T(D^{(k+1)}) \nabla k_y(D^{(k+1)})}{\nabla k_y^T(D^{(k)}) \nabla k_y(D^{(k)})}.$$

4）求

$$\begin{cases} D^{(k+1)} = D^{(k)} + \alpha_k P^{(k)}, \\ \alpha_k = \min\limits_{\alpha} k_y(D^{(k)} + \alpha P^{(k)}). \end{cases}$$

5）若 $\|k_y(D^{(k+1)})\| \leqslant \varepsilon$，则转 6;若 $k < n \times r$ 时返回 3)．当 $k = n \times r$ 时，令 $D^{(0)} = D^{(n\times r)}$ 返回 2)．

6）$Q = (\boldsymbol{q}_1, \cdots, \boldsymbol{q}_n)$，$\boldsymbol{q}_i = S_i D_i$，$K = Z^{-1} U_0^T (Q \Lambda Q^{-1} - A)$．

五、数值实例

例 2[8].

$$B^T = \begin{bmatrix} 0 & 5.679 & 1.136 & 1.136 \\ 0 & 0 & -3.146 & 0 \end{bmatrix},$$

$$A = \begin{bmatrix} 1.38 & -0.2077 & 6.715 & -5.676 \\ -0.5814 & -4.24 & 0 & 0.675 \\ 1.067 & 4.273 & -6.654 & 5.893 \\ 0.048 & 4.273 & 1.343 & -2.104 \end{bmatrix},$$

指定闭环极点为$(-0.2, -0.5, -5.566, -8.666)$．

取 $\beta = 200$，由上面给出的算法得最优解为

$$K = \begin{bmatrix} 0.064747 & -3.089 & -0.325 & 0.3345 \\ 1.038243 & 0.5779 & 0.319 & -0.0465 \end{bmatrix},$$

$$Q = \begin{bmatrix} -0.424 & -0.9028 & 0.481 & -0.492 \\ 0.1413 & 0.0167 & 0.42 & 0.652 \\ -0.52 & -0.401 & -0.707 & 0.403 \\ -0.729 & -0.177 & -0.306 & -0.417 \end{bmatrix},$$

$k_c \doteq 1.05$，$k_y \doteq 1.24$，$n/(n-1) = 1.33$．显然用 k_c 作为鲁棒性指标有可能失败,因为它的适用范围在$[0,1)$,而 k_y 的适用范围在$[0,1.33)$,所以问题的解是合理的.

结 语

鲁棒极点配置问题是八十年代以来引起广泛重视的控制理论新课题. 到目前为止，虽有一些杰出的工作，但还有许多工作要做。根据对偶原理，假若多输出系统完全能观，本文的方法容易推广到鲁棒观测器设计中去。显见本文的指标仍然存在着不足，寻找更方便设计的合理指标及算法无疑是仍然具有吸引力的课题。

参 考 文 献

[1] EDITORIAL, Challenges to Control: An Collective View, *IEEE Trans on AC*, **32**(1987), 275—285.

[2] Kokotovic, R. V., Recent Trends in Feedback Design: An Overviews, *Automatic*, **21** (1985), 225—236.

[3] Cavin, R. K. et al., Robust and Well-conditioned Eigen-structure Assignment Via Sylvester's Equations, American Control Conference, 1982, 1053—1057.

[4] Sun Jiguang, On Numerical Methods for Robust Pole Assignment in Control System Design, *Journal of Computational Mathematics*, **5**(1987), 119—134.

[5] 蒋尔雄等，线性代数，人民教育出版社，1978 年第一版，449—451.

[6] 涂健，周军，二次型最优控制系统鲁棒特征结构配置的计算机辅助设计，控制与决策，**2**(1987), 4, 22—26.

[7] 黄琳，系统与控制中的线性代数，科学出版社，1984, 241—244.

[8] Kautsky, J. et al., Robust Pole Assignment in Linear State Feedback, *Int. J. Control*, **41**(1985), 1129—1155.

[9] 魏权龄等，直接最优化的算法与不动点，系统科学与数学，**1**(1981), 2, 81—98.

[10] Himmelblau, Applied Nonlinear Programming, The University of Taxas, Austin Texas, 1972.

[11] Dickman, A., On the Robustness of Multivariable Linear Feedback Systems in State-space Representation, *IEEE Trans. on AC*, **32**(1987), 3, 407—410.

THE ANALYSIS OF A ROBUST INDEX AND ITS APPLICATION IN ROBUST POLE ASSIGNMENT OF MULTIVARIABLE SYSTEMS

YANG YAGUANG LU YONGZAI

(*Zhejiang University*)

ABSTRACT

The advantages and drawbacks of the robust index applied in references[3,4,6] are analysed in this paper. A robust pole assignment method is then proposed. This method involves much less computation than those in references[4,6,11] and its convergence property is better than that in reference [8].

Key words——Robust index; multivariable systems; pole assignment.