第 11 卷第 5 期
1994 年 10 月

控制理论与应用
CONTROL THEORY AND APPLICATIONS

Vol. 11, No. 5
Oct., 1994

鲁棒极点配置的优化算法*

王耀青
(华中理工大学动力系·武汉，430074)

摘要：本文研究了一类线性多变量系统的鲁棒极点配置问题，提出了一种基于目标函数优化的鲁棒极点配置算法．解决问题的基本思想是，利用极点配置的参数化表示结果，以系统输出的 L_∞ 范数作为极点配置寻优的准则函数，在这个准则下，通过解优化问题的途径确定极点配置问题中的自由参数，从而达到鲁棒极点配置的目的．

关键词：线性系统；鲁棒极点配置；准则函数；特征值

1 引 言

考虑如下线性连续时间系统

$$\dot{x} = Ax + Bu + Dw, \tag{1.1}$$

$$y = Cx. \tag{1.2}$$

其中状态 $x \in \mathbb{R}^n$，控制 $u \in \mathbb{R}^m$，扰动 $w \in \mathbb{R}^p$，输出 $y \in \mathbb{R}^l$，A, B, C, D 为适当维数的矩阵．极点配置问题指的是，设计状态反馈控制器

$$u = -Kx, \tag{2}$$

使得闭环控制系统

$$\dot{x} = (A - BK)x + Dw \tag{3}$$

具有一组指定的特征值 $\lambda_{ci}, i = 1, 2, \cdots, n$.

对于这个问题，在矩阵对 (A, B) 可控的条件下，λ_{ci} 总是可以任意配置，而且当 $1 < m \leqslant n$ 时，配置 λ_{ci} 的矩阵 K 是不唯一的．这就使我们考虑是否可以在配置系统极点的同时，利用系统潜在的设计自由度，设计出具有某一特定性能的闭环控制系统，如闭环系统特征值对系统矩阵 $(A - BK)$ 中可能引入扰动的不敏感性．这就是本文设计方法思想的来源．

研究鲁棒极点配置的方法很多，如文献[1～3]．他们多以系统的条件数作为鲁棒极点配置的准则函数，通过极小化条件数实现系统的极点配置．本文将根据 L_∞ 范数的定义，将系统输出 L_∞ 范数作为鲁棒极点配置的准则函数，利用多输入系统极点配置问题中的自由度或自由参数极小化准则函数，从而达到鲁棒极点配置的目的．

2 预备知识

2.1 基本概念与定义

考虑由(1)～(3)所描述的闭环控制系统．输出 L_∞ 范数定义为

$$\| y(t) \|_\infty^2 \triangleq \sup_{t \geqslant 0} y^{\mathrm{T}}(t) y(t). \tag{4}$$

* 国家教委博士点专项基金资助课题．
本文于1993年5月29日收到．1994年2月8日收到修改稿．

系统扰动 $w(t)$ 的 L_2 范数定义为

$$\| w(t) \|_2^2 \triangleq \int_0^\infty w^{\mathrm{T}}(t) W^{-1} w(t) \mathrm{d}t. \tag{5}$$

式中 T 表示转置，W 为给定的加权矩阵. 下面的引理将给出扰动 $w(t)$ 的 L_2 范数和系统输出 $y(t)$ 的 L_∞ 范数之间的关系，它将为下文从间接的途径达到极小化输出 L_∞ 范数提供了理论基础.

引理 1[4] 对于线性连续时间系统(3)，若矩阵 $(A-BK)$ 渐近稳定，矩阵对 $(A-BK, D)$ 可控，且 $x(0)=0$，则有

$$\| y(t) \|_\infty^2 \leqslant \sigma[Y] \| w(t) \|_2^2. \tag{6}$$

其中 $\sigma(Y)$ 为矩阵 Y 的最大奇异值，$Y=CPC^{\mathrm{T}}$，且 P 满足

$$P(A-BK)^{\mathrm{T}} + (A-BK)P + DWD^{\mathrm{T}} = 0. \tag{7}$$

即 P 为可控格兰姆矩阵 $P = \int_0^\infty e^{(A-BK)t} DWD^{\mathrm{T}} e^{(A-BK)^{\mathrm{T}}t} \mathrm{d}t.$

2.2 极点配置方法的参数化表示

为了简便起见，本文假定 $\lambda_{ci} \overline{\in} \lambda(A)$，$i=1,2,\cdots,n$，并且只考虑 λ_{ci} 互异的情形.

引理 2 考虑由方程(1)所描述的线性连续时间系统，当矩阵对 (A,B) 可控时，使(3)具有特征值 $\lambda_{ci}(i=1,2,\cdots,n)$ 的反馈系数矩阵 K 可以参数化表示为

$$K = -U\mathrm{diag}[\alpha(\lambda_{c1})\ \alpha(\lambda_{c2})\ \cdots\ \alpha(\lambda_{cn})]X^{-1}. \tag{8}$$

当 $\lambda_{ci}=\lambda_{cj}^*$ 时，取 $U_i=U_j^*$，$*$ 表示复数共轭. 式中

$$U = [U_1\ U_2\ \cdots\ U_n],$$
$$X = [X_1\ X_2\ \cdots\ X_n], \quad X_i = CHA_iU_i, \quad i=1,2,\cdots,n,$$
$$C = [B\ AB\ \cdots\ A^{n-1}B], \quad A_i = [I_m\ \lambda_{ci}I_m\ \cdots\ \lambda_{ci}^{n-1}I_m]^{\mathrm{T}}.$$

X_i 为矩阵 $(A-BK)$ 属于其特征值 λ_{ci} 的右特征矢量，H = 第一行为 $[a_1I_m\ a_2I_m\ \cdots\ a_nI_m]$ 的左上三角 Toeplitz 矩阵，$a_i(i=1,2,\cdots,n)$ 为系统(1)的开环特征多项式 $\alpha(s)$

$$\alpha(s) = \det(sI_n - A) = a_ns^n + a_{n-1}s^{n-1} + \cdots + a_1s + a_0 \tag{9}$$

的系数；$a_n=1$，$I_{\tau\times\tau}$ 表示 $\tau\times\tau$ 维的单位阵.

证 因矩阵 K 使矩阵 $(A-BK)$ 具有特征值 $\lambda_{ci}(i=1,2,\cdots,n)$，根据特征值的定义有

$$(A-BK)X_i = \lambda_{ci}X_i, \quad i=1,2,\cdots,n,$$
$$X_i = -(\lambda_{ci}I_n - A)^{-1}BKX_i, \quad \lambda_{ci} \overline{\in} \lambda(A). \tag{10}$$

因为 $\alpha(\lambda_{ci})(\lambda_{ci}I_n-A)^{-1}B=CHA_i$，所以，(10)式可以表示为

$$X_i = -[\alpha(\lambda_{ci})]^{-1}CHA_iKX_i. \tag{11}$$

定义 $K=-U\mathrm{diag}[\alpha(\lambda_{c1})\ \alpha(\lambda_{c2})\ \cdots\ \alpha(\lambda_{cn})][X_1\ X_2\ \cdots\ X_n]^{-1}$，即(8)式成立时，(11)式可以简化表示为

$$X_i = CHA_iU_i, \quad i=1,2,\cdots,n. \tag{12}$$

这就证明了(8)式的正确性. 当 $\lambda_{ci}=\lambda_{cj}^*$ 时，假定 $i=1, j=2$，取 $U_1=U_2^*=\overline{U}_2+\overline{U}_1j$，则(8)式经变换矩阵

$$T = \text{block-daig}\left\{\begin{bmatrix}\frac{1}{2}j & \frac{1}{2}\\ -\frac{1}{2}j & \frac{1}{2}\end{bmatrix}, 1, \cdots, 1\right\}$$

变换后可以表示为

$$K = -[\overline{U}_1\ \overline{U}_2\ U_3 \cdots U_n]\text{block-diag}\left\{\begin{bmatrix}\mathrm{Re}\,\alpha(\lambda_{c1}) & -\mathrm{Im}\,\alpha(\lambda_{c1})\\ \mathrm{Im}\,\alpha(\lambda_{c1}) & \mathrm{Re}\,\alpha(\lambda_{c1})\end{bmatrix}, \alpha(\lambda_{c3}), \cdots, \alpha(\lambda_{cn})\right\}$$

$$\cdot\left[CH(\mathrm{Im}\,\Lambda_1\ \mathrm{Re}\,\Lambda_1\ \Lambda_3 \cdots \Lambda_n)\text{block-diag}\left\{\begin{bmatrix}\overline{U}_1 & -\overline{U}_2\\ \overline{U}_2 & \overline{U}_1\end{bmatrix}, U_3, \cdots, U_n\right\}\right]^{-1}.$$

这就保证了 $K\in \mathbf{R}^{l\times n}$.

性质 1　(8)式中 U_i 乘以一个非零的常数，矩阵 K 的计算结果不变.

性质 2　当 $m=1$时，对于任何非零的 U_i，下面的等式恒成立，即

$$[\alpha(\lambda_{ci}) + KCH\Lambda_i]U_i = 0,\quad i = 1,2,\cdots,n.$$

这个性质表明单输入系统的极点配置是唯一的，在 λ_{ci}确定的情况下，系统无鲁棒极点配置而言.

性质 3[5]　当矩阵对(A,B)可控，且 λ_{ci}互异时，几乎对于所有的 $U_i\in \mathbf{C}^{n\times 1}, X_i, i=1,2,\cdots,n$，在 C^n 上总是线性无关的.

性质3为我们下一部分进行优化问题求解提供了有利条件.

3　鲁棒极点配置方法

上一节我们已经分别介绍了 L_∞范数和 L_2范数的定义形式以及它们之间的相互关系，并给出了极点配置的参数化表示形式及其有关性质. 从输出 L_∞范数$\|y(t)\|_\infty^2$来看，除扰动因素 $w(t)$之外，影响它的人为因素是 W 的选取和 $\sigma(Y)$. 也就是说，通过改变 W 和 $\sigma(Y)$可以达到改变$\|y(t)\|_\infty^2$的目的. 本文将不研究 W 的选取问题，并设 $W=I_p$. 对 W 选取方法有兴趣者可以参阅文献[4]. 这里将只考虑如何通过改变 $\sigma(Y)$实现极小化$\|y(t)\|_\infty^2$的目的. 根据本文上一节的定义 $Y=CPC^{\mathrm{T}}$ 可知，由于 $\sigma(Y)\leqslant\sigma^2(C)\sigma(P)$. 显然，利用 K 中的自由参数使 $\sigma(Y)$极小化，实质上等价于极小化 $\sigma(P)$. 根据这一关系，并考虑到矩阵 P 的对称正定性，鲁棒极点配置问题可以归结为以下问题的求解

$$\min_{U_i} J = \min_{U_i}\lambda_{\max}[P],\quad i = 1,2,\cdots,n. \tag{13}$$

$$\text{s.t.}\ \begin{cases}(A-BK)P + P(A-BK)^{\mathrm{T}} + DD^{\mathrm{T}} = 0, & (14.1)\\ K = -[\alpha(\lambda_{c1})U_1\ \alpha(\lambda_{c2})U_2 \cdots \alpha(\lambda_{cn})U_n][CH(\Lambda_1U_1\ \Lambda_2U_2 \cdots \Lambda_nU_n)]^{-1}. & (14.2)\end{cases}$$

为方便起见，在进行问题求解之前，先假定 $\lambda_{ci}\in\mathbf{R}, i=1,2,\cdots,n$.

定理 1　对于由方程(13)，(14.1)和(14.2)所定义的一类优化问题，其一阶梯度$\frac{\partial J}{\partial U_i}$可以表示为

$$\frac{\partial J}{\partial U_i} = 2[\alpha(\lambda_{ci})I_m + KCH\Lambda_i]^{\mathrm{T}}B^{\mathrm{T}}VPX^{-\mathrm{T}}[\underbrace{0 \cdots 0}_{i-1}\ 1\ 0 \cdots 0]^{\mathrm{T}} \tag{15}$$

$i=1,2,\cdots,n$. 式中 V 满足方程

$$(A-BK)^{\mathrm{T}}V + V(A-BK) + v_mv_m^{\mathrm{T}} = 0 \tag{16}$$

v_m 为矩阵 P 属于其最大特征值 $\lambda_{\max}[P]$的右特征矢量.

证 根据(13)式及 v_m 的定义,J 可以表示为

$$J = v_m^{\mathrm{T}} P v_m = \mathrm{tr}[v_m v_m^{\mathrm{T}} P]$$

当变量 U 具有增量 ΔU 时,定义

$$\Delta \bar{U} = \Delta U \mathrm{diag}[\alpha(\lambda_{c1})\ \alpha(\lambda_{c2})\ \cdots\ \alpha(\lambda_{cn})] = [0\ \cdots\ 0\ \alpha(\lambda_{ci})\Delta U_i\ 0\ \cdots\ 0]$$

$$\Delta X = CH[0\ \cdots\ 0\ \Lambda_i \Delta U_i\ 0\ \cdots\ 0], \quad i \in [1, n],$$

$$\bar{U} = [U_1\ U_2\ \cdots\ U_n]\mathrm{diag}[\alpha(\lambda_{c1})\ \alpha(\lambda_{c2})\ \cdots\ \alpha(\lambda_{cn})].$$

则矩阵 K 的增量 ΔK 满足方程

$$\Delta \bar{U} + K\Delta X + \Delta K X = 0,$$

$$\Delta K = -[\Delta \bar{U} + K\Delta X]X^{-1} = -\{0\ \cdots\ 0\ \alpha(\lambda_{ci})I_m + KCH\Lambda_i]\Delta U_i\ 0\ \cdots\ 0\}X^{-1}. \tag{17}$$

矩阵 P 的增量 ΔP 可以由方程(14.1)的增量方程

$$(A - BK)\Delta P + \Delta P(A - BK)^{\mathrm{T}} - B\Delta KP - (B\Delta KP)^{\mathrm{T}} = 0 \tag{18}$$

来求得.另一方面,J 的增量 ΔJ 可以表示为

$$\Delta J \cong \mathrm{tr}[(v_m + \Delta v_m)(v_m + \Delta v_m)^{\mathrm{T}}(P + \Delta P)] - \mathrm{tr}[(v_m v_m^{\mathrm{T}})P] \tag{19}$$

当 ΔU_i 足够小时,可以认为$(v_m+\Delta v_m)\cong v_m$,因此有

$$\Delta J = \mathrm{tr}[v_m v_m^{\mathrm{T}} \Delta P]. \tag{20}$$

根据文献[6],增量方程(18)的解为

$$\Delta P = -\sum_{j=1}^{n}\sum_{k=1}^{n} v_{jk}(A - BK)^{j-1}[B\Delta KP + (B\Delta KP)^{\mathrm{T}}](A^{\mathrm{T}} - K^{\mathrm{T}}B^{\mathrm{T}})^{k-1}. \tag{21}$$

式中 $v_{jk}(=v_{kj})$为一常数,且只与矩阵$(A-BK)$和矩阵对$[(A-BK),D]$的可控性有关.将(17)式和(21)式代入(20)式得

$$\Delta J \cong \mathrm{tr}\{\sum_j\sum_k v_{jk}v_m v_m^{\mathrm{T}}(A - BK)^j[B(\Delta\bar{U} + K\Delta X)X^{-1}P + PX^{-\mathrm{T}}(\Delta\bar{U} + K\Delta X)^{\mathrm{T}}B^{\mathrm{T}}](A^{\mathrm{T}} - K^{\mathrm{T}}B^{\mathrm{T}})^k$$

$$= 2\mathrm{tr}\{\sum_j\sum_k v_{jk}v_m v_m^{\mathrm{T}}(A - BK)^j[B(\Delta\bar{U} + K\Delta X)X^{-1}P](A^{\mathrm{T}} - K^{\mathrm{T}}B^{\mathrm{T}})^k$$

$$= 2\mathrm{tr}\{\sum_j\sum_k v_{jk}X^{-1}P(A^{\mathrm{T}} - K^{\mathrm{T}}B^{\mathrm{T}})^k v_m v_m^{\mathrm{T}}(A - BK)^j B(\Delta\bar{U} + K\Delta X)\}$$

$$= 2[\underbrace{0\ \cdots\ 0\ \ 1}_{i}\ \ 0\ \cdots\ 0]X^{-1}P\underbrace{\sum_j\sum_k v_{jk}(A^{\mathrm{T}} - K^{\mathrm{T}}B^{\mathrm{T}})^k v_m v_m^{\mathrm{T}}(A - BK)^j}_{V}$$

$$\cdot B[\alpha(\lambda_{ci})I_m + KCH\Lambda_i]\Delta\bar{U}_i$$

$$= 2([0\ \cdots\ 0\ \ 1\ \ 0\ \cdots\ 0]X^{-1}PVB[\alpha(\lambda_{ci})I_m + KCH\Lambda_i]\Delta\bar{U}_i, \quad i \in [1, n].$$

因此有

$$\frac{\partial J}{\partial U_i} = 2[\alpha(\lambda_{ci})I_m + KCH\Lambda_i]^{\mathrm{T}}B^{\mathrm{T}}VPX^{-\mathrm{T}}[\underbrace{0\ \cdots\ 0\ \ 1}_{i}\ \ 0\ \cdots\ 0]^{\mathrm{T}}.$$

式中 V 是方程(16)的唯一对称非负定解.

根据定理1和性质2,当 $m=1$时,$\partial J/\partial U_i=0$.这说明对于 $m=1$的系统根本无鲁棒极点配置而言.所以说,本文的结论只能用于 $m>1$的线性多变量控制系统.

鲁棒极点配置算法如下:

1) 给定 U 的初值 U_0；

2) 根据(8)和(14)式计算矩阵 X, K 和 P；

3) 计算 v_m，并根据(16)式计算矩阵 V；

4) 计算 $\partial J/\partial U_i$，若 $\sum_{i=1}^{n}\|\partial J/\partial U_i\|_F$ 小于某一给定精度 ε 或 J 趋于某一定值时，转至6)；

5) 定义 $U=U-\begin{bmatrix}\dfrac{\partial J}{\partial U_1} & \dfrac{\partial J}{\partial U_2} & \cdots & \dfrac{\partial J}{\partial U_n}\end{bmatrix}\mathrm{diag}[h_1\ h_2\ \cdots\ h_n]$，转至2)；

6) 停止求解，矩阵 K 即为所求.

说明　①在算法中 $h_i>0$为寻优步长，合适的 h_i 对算法的收敛性具有很大的好处. ②上述寻优结果为局部极值，也就是说，计算结果与 U_0和 h_i 的取值有关.

4　设计举例

考虑一线性多变量连续时间系统

$$\dot{x}=\begin{bmatrix}-1 & 0 & 0\\ 1 & 0 & -1\\ 0 & 4 & 0\end{bmatrix}x+\begin{bmatrix}1 & 0\\ 0 & 1\\ 0 & 0\end{bmatrix}u+w,$$

设计反馈控制器 $u=-Kx$，使得 $\lambda_{c1}=-2, \lambda_{c2}=-3, \lambda_{c3}=-4$.

首先取 $U=\begin{bmatrix}1 & 0 & 1\\ 0 & 1 & 0\end{bmatrix}$，$h_1=0.05, h_2=0.1, h_3=0.2$；此时，矩阵 X, K, P 和 V 分别为

$$X=\begin{bmatrix}8 & 0 & 20\\ -2 & 6 & -4\\ 4 & -8 & 4\end{bmatrix},\quad K=\begin{bmatrix}2.4286 & -11.4286 & -8.5714\\ 0.9286 & 5.5714 & 0.9286\end{bmatrix},$$

$$P=\begin{bmatrix}2.2781 & -0.1645 & 1.0722\\ -0.1645 & 0.1309 & -0.1250\\ 1.0722 & -0.1250 & 0.6723\end{bmatrix},\quad V=\begin{bmatrix}0.1218 & 0.2871 & 0.2870\\ 0.2871 & 1.5424 & 1.3274\\ 0.2870 & 1.3274 & 1.1658\end{bmatrix}.$$

经过17次迭代运算，最后的结果为

$$X=\begin{bmatrix}7.9698 & -3.5136 & -4.0266\\ -3.6866 & 6.7180 & 12.7998\\ 7.3732 & -8.5730 & -12.7998\end{bmatrix},\quad K=\begin{bmatrix}0.7457 & -0.8686 & -0.5995\\ 0.1816 & 7.2543 & 2.5117\end{bmatrix},$$

$$P=\begin{bmatrix}0.3019 & 0.0250 & 0.0332\\ 0.0250 & 0.1323 & -0.1250\\ 0.0332 & -0.1250 & 0.44166\end{bmatrix},\quad V=\begin{bmatrix}0.0213 & 0.0339 & 0.0667\\ 0.0339 & 0.0817 & 0.1265\\ 0.0667 & 0.1265 & 0.3638\end{bmatrix}.$$

定义[1]　$k_f(X)=\|X\|_F\|X^{-1}\|_F$；$k_2(K)=\|K\|_2=\sqrt{\lambda_{\max}(K^{\mathrm{T}}K)}$；则有

初值时：$J^*=2.8295$，　$k_f(X)=25.1257$，　$k_2(K)=15.2941$，

结束时：$J^*=0.4668$，　$k_f(X)=22.4583$，　$k_2(K)=7.7269$.

5　结束语

如何定义一种鲁棒性准则，且便于问题的求解是鲁棒控制器设计问题中的关键. 本文所给出的算法显然具有这种特点，这就是本文的优点所在. 由于定义鲁棒性准则的方法很多，如何将一种定义方法同另一种定义方法相联系，特别是自由参数优化设计方法与 H_∞，LQ 最优控制等设计方法之间的内在联系，并加以比较，这将是有待于进一步研究的内容.

参考文献

[1] Kautsky, J., Nichols, N. K. and Van Dooren, P.. Robust Pole Assinment in Linear State Feedback. Int. J. Control, 1985, 41(5):1129—1155

[2] 杨亚光,吕勇哉.鲁棒设计的时域方法.控制理论与应用,1988,5(4):1—9

[3] Byers, B. and Nash, S. G.. Approaches to Robust Pole Assignment. Int. J. Control, 1989, 49(1):97—117

[4] Skelton, R. Z. and Zhu, G.. Optimal L_∞ Bounds for Disturbance Robustness. IEEE Trans. Automat. Contr., 1992, 37 (10)

[5] O'Reilly, J. and Fahmy, M. M.. The Minimum Number of Degrees of Freedom in State Feedback Control. Int. J. Control, 1985, 41(3):749—768

[6] DeZouza, E. and Bhattacharyya, S. P.. Controllability, Observability and the Solution of $AX-XB=C$. Linear Algebra Appl., 1981, 39(2):167—188

An Optimization Algorithm for Robust Pole Assignment

WANG Yaoqing

(Department of Power Engineering, Huazhong University of Science & Technology·Wuhan, 430074, PRC)

Abstract: The problem of robust pole assignment has been studied for a class of linear multivariable systems in the paper, and a robust pole assignment algorithm is proposed based on the objective functional optimization. The main idea of the paper is that the feedback gain matrix can be parametized in terms of a set of free variables, and the variables can be determined by solving the optimization problem under the consideration of L_∞ norm of system output as the objective functional to be minimized. As a result, the robust controller is determined.

Key words: linear systems; robust pole assignment; criterion function; eigenvalues

本文作者简介

王耀青 1961年生.1983年毕业于天津纺织工学院自动化系.1986年在中国科学院自动化研究所取得硕士学位;1989年在浙江大学获得工学博士学位.现任华中理工大学副教授.主要研究兴趣为线性多变量控制系统设计中自由参数特性的研究,鲁棒控制器设计方法的研究以及LQ最优控制问题的研究等.