第22卷 第2期　　自 动 化 学 报　　Vol. 22, No. 2
1996 年 3 月　　ACTA AUTOMATICA SINICA　　March, 1996

一种新的H$^\infty$-优化方法:梯度方法[1)]

胡庭姝　　陈 力
(上海交通大学自动化系　200030)

摘　　要

提出一种灵活、有效的H$^\infty$-优化方法：梯度方法．利用H$^\infty$-范数与状态空间实现的关系，定义了目标函数 $\rho(\varepsilon,F)$，$\rho(\varepsilon,F)$ 与 H$^\infty$- 范数之间的关系是:

$$\lim_{\varepsilon\to 0}\rho(\varepsilon,F)=1/\|T(s,F)\|_\infty.$$

分析了 $\rho(\varepsilon,F)$ 的可微性,并给出了 $\partial\rho(\varepsilon,F)/\partial F$ 的具体表达式以及使 $\rho(\varepsilon,F)$ 极大化的梯度方法,从而导致 $\|T(s,F)\|_\infty$ 的极小化。实例表明,梯度方法能有效地使 $\rho(\varepsilon,F)$ 上升，并收敛于驻点或终止于不可微点.

关键词: H$^\infty$-范数，梯度方法,可微性,极点配置.

1　引言

H$^\infty$-优化方法在80年代末就已形成一套比较成熟和完整的理论体系．已有的H$^\infty$-优化方法不同于其它传统的规划方法．两类方法(频域方法和时域方法)都是先给定一个正数 γ，然后通过解 Riccati 方程判别是否存在控制器使某一H$^\infty$-性能指标小于 γ．这种方法的优点是能逼近最优解,而不用考虑优化问题的凸性．其缺点是缺乏灵活性,只能优化单一的H$^\infty$-范数,而不能将其它类型的目标函数综合起来,构成带加权的复合目标函数．这往往使得H$^\infty$-优化问题的最优解(或次优解)在其它指标上不能令人满意，如极点远离虚轴,高增益反馈等．如果能加上极点固定在某一位置或区域的限制条件,或将反馈矩阵的范数作为惩罚项加在H$^\infty$-范数之后,那么优化效果就能得到改善,但这样改变后的优化问题是已有的H$^\infty$-优化方法所不能处理的.

规划中最常用的是梯度方法．这种方法使用起来很灵活，只要单个目标函数的梯度能求出,就能求得复合目标函数的梯度,也能处理带约束条件的优化问题．缺点是不能保证求得最优解．这和上述的H$^\infty$-优化方法正相反．但在最优解并不是非求出不可、而只希望目标函数尽可能得到改善的场合,梯度方法则是非常简单而有效的方法.

求H$^\infty$-范数本身就具有一定的难度,因而求H$^\infty$-范数对某一变量的梯度,似乎是不现实的．文中将利用[1]中所得出的关系,定义另一目标函数 $\rho(\varepsilon,F)$，使得

1) 本文的研究获国家自然科学基金资助.
本文于1993年8月23日收到

$$\lim_{\varepsilon\to 0}\rho(\varepsilon,F) = 1/\|T(s,F)\|_\infty.$$

而 $\rho(\varepsilon,F)$ 可通过梯度法得到优化，从而使 $\|T(s,F)\|_\infty$ 极小．这种方法还可用于优化传递函数中的其它变化参数．

2 等价的目标函数 $\rho(\varepsilon,F)$

先讨论如下优化问题:

$$\inf\|H(sI-A-BF)^{-1}G\|_\infty,\ \text{s.t.}\ A+BF\ \text{稳定}. \tag{2.1}$$

设 $A\in R^{n\times n}$, $B\in R^{n\times m}$, $F\in R^{m\times n}$, $G\in R^{n\times p}$, $H\in R^{l\times n}$. 以上目标函数在不同场合可分别代表敏感性指标和鲁棒性指标[2,3]. 这一优化问题的次优解已经能求得. 本文将以此问题作为出发点,建立一套全新的 H$^\infty$-优化方法.

记 $T(s,F)=H(sI-A-BF)^{-1}G$，文 [1] 有如下结论:

$\|T(s,F)\|_\infty = \max\left\{\gamma>0,\ \begin{bmatrix} A+BF & GG'/\gamma^2 \\ -H'H & -(A+BF)' \end{bmatrix}\text{在虚轴上有特征值.}\right\}$ 令 $k=1/\gamma$，并简记 $R(F,k)=\begin{bmatrix} A+BF & kGG' \\ -kH'H & -(A+BF)' \end{bmatrix}$，则上式等效为

$$\frac{1}{\|T(s,F)\|_\infty} = \min\{k>0,\mathrm{Re}\lambda_i[R(F,k)]=0,\ \text{对某一}\ i\}. \tag{2.2}$$

另外,文[1]还指出，当 $k\geqslant 1/\|T(s,F)\|_\infty$ 时，$R(F,k)$ 在虚轴上总有特征值，当 $k<1/\|T(s,F)\|_\infty$ 时，$R(F,k)$ 在虚轴上无特征值．由 $R(F,k)$ 的结构还可以证明，$R(F,k)$ 的特征值关于虚轴对称．

定义：$\rho_0(F):=\min\{k>0,\mathrm{Re}\lambda_i[R(F,k)]=0$, 对某一 $i\}$.

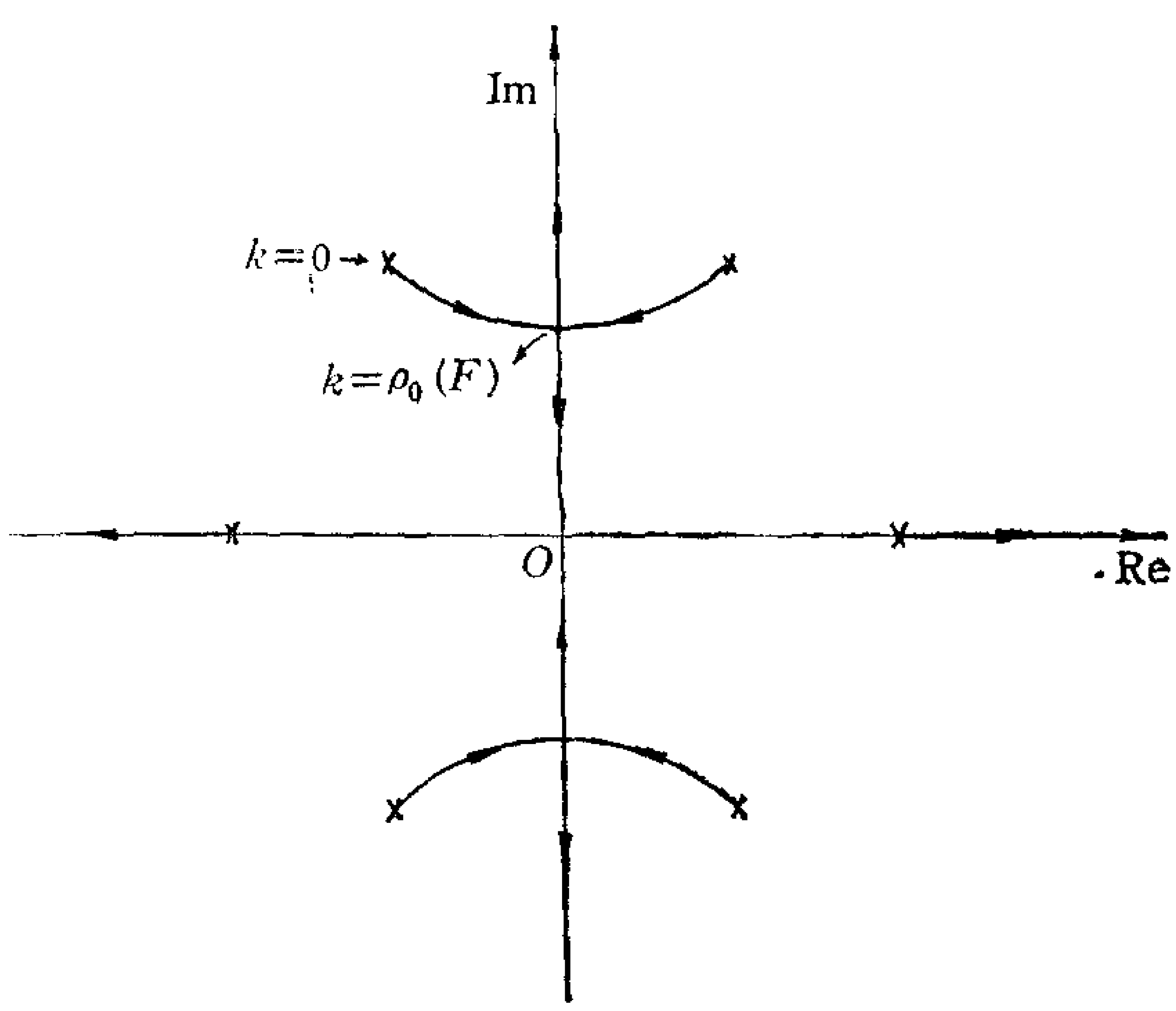


图 1

显然 $\rho_0(F)=1/\|T(s,F)\|_\infty$. 形象地说,$\rho_0(F)$ 等于 $R(F,k)$ 的特征轨迹与虚轴相交的最小 k，见图 1. 由于特征值关于虚轴对称,所以当 $k=\rho_0(F)$ 时,虚轴上的特征值必为重根，这对于求 $\partial\rho_0(F)/\partial F$ 很不利.

为此,定义 $\rho(\varepsilon,F)$ 为

$$\rho(\varepsilon,F):=\min\{k>0,\ \mathrm{Re}\lambda_i[R(F,k)]=-\varepsilon,\ \text{对某一}\ i\}, \tag{2.3}$$

其中 ε 为一小正数,则

$$\lim_{\varepsilon\to 0}\rho(\varepsilon,F)=\rho_0(F).$$

形象地说，$\rho(\varepsilon,F)$ 等于当 $R(F,k)$ 的特征轨迹首次与 $-\varepsilon+j\omega$ 轴相交的 k，见图 2.

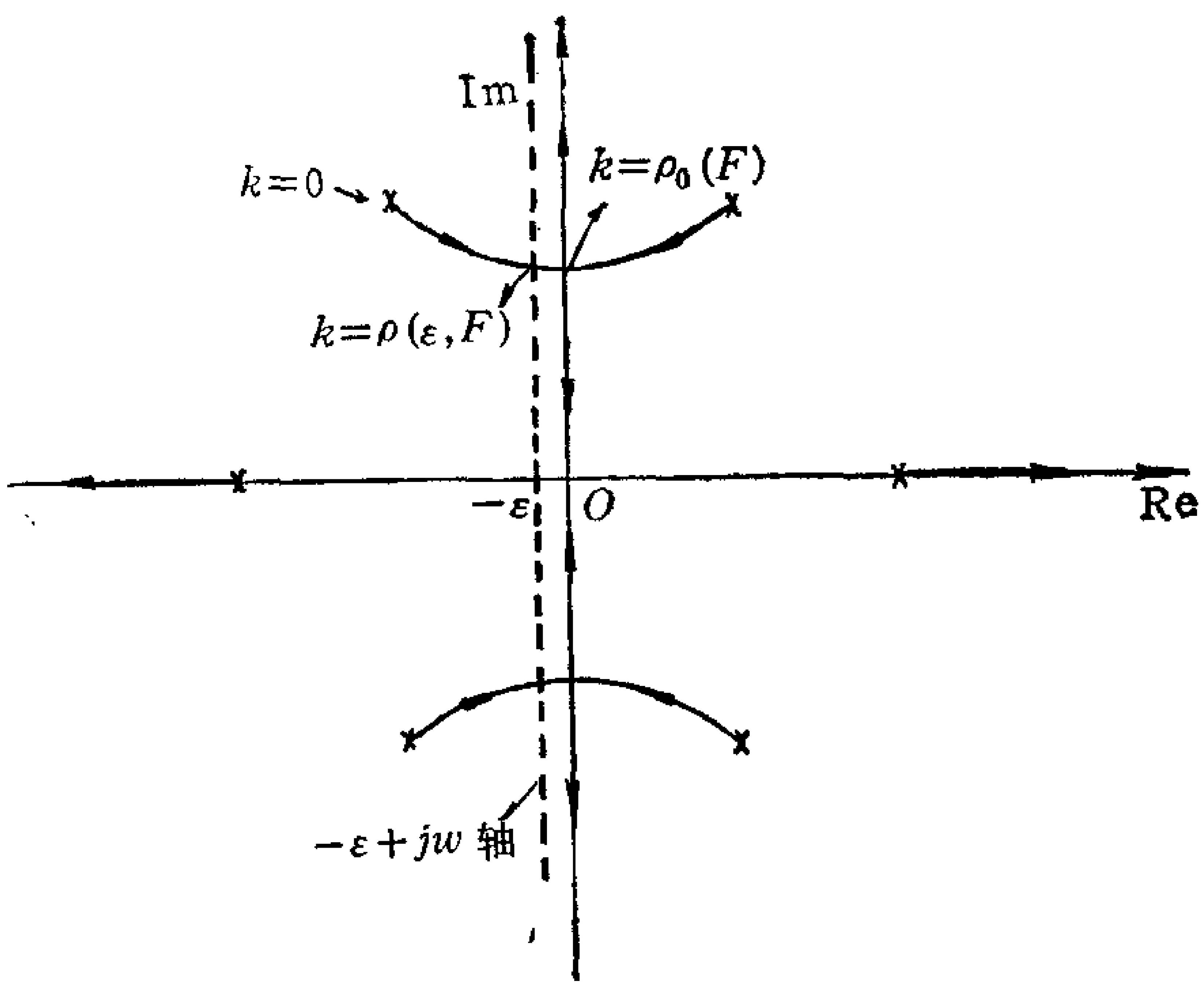


图 2

易知，$\rho(\varepsilon, F) < \rho_0(F)$。$\lim\limits_{\varepsilon\to 0}\rho(\varepsilon,F) = \rho_0(F) = 1/\|T(s,F)\|_\infty$。

至此，可将优化问题 (2.1) 转化为如下形式：

$$\sup_F \{\rho(\varepsilon,F)\} \quad \text{s. t. } A+BF \text{ 稳定.} \tag{2.4}$$

当 ε 足够小时，$\rho(\varepsilon,F)$ 的极大化等效于 $\|T(s,F)\|_\infty$ 的极小化．理论上讲，$\rho(\varepsilon,F)$ 可任意逼近 $1/\|T(s,F)\|_\infty$。但实际上，当 $\varepsilon\to 0$ 时，也会出现奇异性和计算精度问题．这和已有的 H^∞-优化方法相类似。

3　$\rho(\varepsilon,F)$ 的可微性及极大化方法

定理 3.1．对于给定的 F_0，记 $k_0=\rho(\varepsilon,F_0)$．如果以下条件成立：当 $k=k_0$ 时，在 $-\varepsilon+j\omega$ 轴的上半部 $(\omega\geqslant 0)$，$R(F_0,k)$只有唯一的第 i 条特征值轨迹与之相交，且此特征值轨迹满足

$$\partial \mathrm{Re}\,\lambda_i[R(F,k)]/\partial k\big|_{\substack{k=k_0\\F=F_0}} \neq 0,$$

则在点 F_0 的某一邻域 Δ 内，有

$$\frac{\partial\rho(\varepsilon,F)}{\partial F} = -\frac{\partial \mathrm{Re}\,\lambda_i[R(F,k)]}{\partial F}\bigg/\frac{\partial \mathrm{Re}\,\lambda_i[R(F,k)]}{\partial k}\bigg|_{k=\rho(\varepsilon,F)}. \tag{3.1}$$

证明．注意到 $R(F,k)=\begin{bmatrix} A+BF & kGG' \\ -kH'H & -(A+BF)' \end{bmatrix}$ 是关于 F 和 k 的连续可微函数．又由条件知，在 $k=k_0$ 处，λ_i 是 $R(F_0,k_0)$ 的单重特征值．故据特征值及特征向量的连续性可知，在点 $(F_0;k_0)$ 的邻域 D 内，$\partial \mathrm{Re}\,\lambda_i/\partial F$ 与 $\partial \mathrm{Re}\,\lambda_i/\partial k$ 存在且连续。

据多变量隐函数存在定理，对隐函数 $\mathrm{Re}\lambda_i[R(F,k)]=-\varepsilon$ 有以下结论：

1) 在点 $(F_0;k_0)$ 的某一邻域 Δ' 内，对于固定的 ε，方程 $\mathrm{Re}\,\lambda_i[R(F,k)]=-\varepsilon$ 唯一地确定一个函数 $k=f(F)$，且 $k_0=f(F_0)$；

2) $f(F)$ 在 Δ' 内连续；

3) $\dfrac{\partial k}{\partial F} = -\dfrac{\partial \mathrm{Re}\,\lambda_i}{\partial F} \Big/ \dfrac{\partial \mathrm{Re}\,\lambda_i}{\partial k}\Big|_{k=f(F)}$.

由 $k=f(F)$ 在 Δ' 内的连续性、$k_0=\rho(\varepsilon,F_0)$ 及 $\rho(\varepsilon,F)$ 的定义可得，必存在点 F_0; 的邻域 Δ，在其中有 $\rho(\varepsilon,F)f(F)$，代入结论(3)，命题得证。

如果在 $k_0=\rho(\varepsilon,F_0)$ 处，$R(F_0,k_0)$ 有二个以上的特征轨迹与 $-\varepsilon+j\omega$ 轴上半部 $(\omega\geqslant 0)$ 相交，那么一般说来 $\rho(\varepsilon,F)$ 在 F_0 处是不可微的。这种情况等效于 $\bar{\sigma}[T(j\omega,F_0)]$ 在 $0\leqslant\omega<\infty$ 的范围内有两个以上等高的峰值(设 $\varepsilon\approx 0$)。

以下推导求 $\partial\rho(\varepsilon,F)/\partial F$ 的详细算法。

对于给定的 F，利用两分法求 $\rho(\varepsilon,F)$。记 $R[F,\rho(\varepsilon,F)]$ 在 $-\varepsilon+j\omega$ 轴上的特征值为 $-\varepsilon+j\omega_0$，所对应的左、右特征向量分别是 t' 和 v，且 $t'v=1$，则

$$\frac{\partial\lambda_i}{\partial k}=t'\frac{\partial R(F,k)}{\partial k}v,\quad \frac{\partial\lambda_i}{\partial F_{lj}}=t'\frac{\partial R(F,k)}{\partial F_{lj}}v,$$

故 $$\frac{\partial \mathrm{Re}\,\lambda_i}{\partial k}=\mathrm{Re}\left[t'\frac{\partial R(F,k)}{\partial k}v\right],\quad \frac{\partial \mathrm{Re}\,\lambda_i}{\partial F_{lj}}=\mathrm{Re}\left[t'\frac{\partial R(F,k)}{\partial F_{lj}}v\right].$$

将 t', v 分块为 $t'=[t_1', t_2']$, $v=\begin{bmatrix}v_1\\v_2\end{bmatrix}$, $t_1,t_2,v_1,v_2\in R^n$,

再由 $$R(F,k)=\begin{bmatrix}A+BF & kGG'\\ -kH'H & -(A+BF)'\end{bmatrix}$$

可得

$$\frac{\partial \mathrm{Re}\,\lambda_i}{\partial F_{lj}}=\mathrm{tr}\left[\mathrm{Re}(v_1t_1'-t_2v_2')\cdot B\cdot\frac{\partial F}{\partial F_{lj}}\right],$$

因此得

$$\frac{\partial \mathrm{Re}\,\lambda_i}{\partial F}=[\mathrm{Re}(v_1t_1'-t_2v_2')\cdot B]'.$$

同理

$$\frac{\partial \mathrm{Re}\,\lambda_i}{\partial k}=\mathrm{Re}(t_1'GG'v_2-t_2'H'Hv_1). \tag{3.2}$$

最后，由定理 3.1 得

$$\frac{\partial\rho(\varepsilon,F)}{\partial F}=-[\mathrm{Re}(v_1t_1'-t_2v_2')B]'/\mathrm{Re}(t_1'GG'v_2-t_2'H'Hv_1). \tag{3.3}$$

考虑优化问题(2.4)，现已求得 $\partial\rho(\varepsilon,F)/\partial F$，由于满足 $A+BF$ 稳定的 F 是个开集，因此可利用梯度方法使 $\rho(\varepsilon,F)$ 增大，从而使 $\|T(s,F)\|_\infty$ 减小。这种方法可推广至优化传递函数中的其它变化参数。

4 带有极点配置约束的 H$^\infty$-优化问题

考虑以下优化问题：

$$\inf\|H(sI-A-BF)^{-1}G\|_\infty, \tag{4.1}$$

$$\text{s.t. } V^{-1}(A+BF)V=A_1, \tag{4.2}$$

其中 $A_1=\mathrm{diag}[\lambda_1,\lambda_2\cdots\lambda_n]$. $\lambda_1,\cdots\lambda_n$ 为所指定的闭环极点. 如果其中有共轭复数,例如 $\lambda_i,\lambda_{i+1}=\alpha_i\pm j\beta_i$, 则令 $A_1=\mathrm{diag}\left[\lambda_1,\cdots,\begin{bmatrix}\alpha_i & \beta_i\\ -\beta_i & \alpha_i\end{bmatrix},\cdots,\lambda_n\right]$.(4.2) 式即极点配置约束[4],类似文[4]的方法,(4.2)式可通过引入一矩阵函数 f 得以放松.

定义 4.1. 设 (A,B) 可控, A 和 A_1 无相同特征值, 则定义 $f:R^{m\times n}\to R^{m\times n}$ 为: 已知 $U\in R^{m\times n}$, 解 $AV-VA_1=-BU$, 令 $F=UV^{-1}$, 则 f 的定义域 D_f 为使 V 非奇异的 U 的集合, U 在 f 下的象为 F, 即 $F=f(U)$. f 的值域 R_f 为 D_f 在 f 下的象集.

已经证明[4,5], f 的值域就是所有满足 (4.2) 式的 F 的集合, f 的定义域为 $R^{m\times n}$ 中稠的开集. 利用此性质,就可消除约束条件 (4.2), 将 (4.1) 式中的 F 用 $f(U)$ 代替, 而 U 是自由变量.

考虑优化问题

$$\sup_F\{\rho(\varepsilon,F)\},\ \text{s. t. } V^{-1}(A+BF)V=A_1. \tag{4.3}$$

记 $J(U)=\rho(\varepsilon,f(U))$, 则 $\partial J/\partial U$ 可推导如下.

$$\begin{aligned}\frac{\partial J}{\partial U_{pq}}&=\sum_{l=1}^{m}\sum_{j=1}^{n}\frac{\partial J}{\partial F_{lj}}\cdot\frac{\partial F_{lj}}{\partial U_{pq}}\\&=\mathrm{tr}\left[\left(\frac{\partial J}{\partial F}\right)'\frac{\partial F}{\partial U_{pq}}\right]\qquad\left(\frac{\partial J}{\partial F}=\frac{\partial\rho(\varepsilon,F)}{\partial F}\text{ 由 (3.3) 式给出}\right)\\&=\mathrm{tr}\left[\left(\frac{\partial J}{\partial F}\right)'\cdot\left(\frac{\partial UV^{-1}}{\partial U_{pq}}\right)\right]\qquad(F=UV^{-1})\\&=\mathrm{tr}\left[\left(\frac{\partial J}{\partial F}\right)'\left(\frac{\partial U}{\partial U_{pq}}V^{-1}-UV^{-1}\frac{\partial V}{\partial U_{pq}}V^{-1}\right)\right]\\&=\mathrm{tr}\left[\underbrace{V^{-1}\left(\frac{\partial J}{\partial F}\right)'}_{Q_1}\frac{\partial U}{\partial U_{pq}}-\underbrace{V^{-1}\left(\frac{\partial J}{\partial F}\right)'F}_{Q_2}\frac{\partial V}{\partial U_{pq}}\right]\\&=\mathrm{rs}\,Q_1\cdot\mathrm{cs}\frac{\partial U}{\partial U_{pq}}-\mathrm{rs}Q_2\cdot\mathrm{cs}\frac{\partial V}{\partial U_{pq}}\end{aligned}$$

由定义 4.1, V 是 $AV-VA_1=-BU$ 的解,即

$$\mathrm{cs}V=-(I_n\otimes A-A_1'\otimes I_n)^{-1}\cdot(I_m\otimes B)\cdot\mathrm{cs}U.$$

因而 $\dfrac{\partial J}{\partial U_{pq}}=[\mathrm{rs}Q_1+\mathrm{rs}Q_2\cdot(I_n\otimes A-A_1'\otimes I_n)^{-1}(I_m\otimes B)]\cdot\mathrm{cs}\dfrac{\partial U}{\partial U_{pq}}$.

令 X 为方程 $XA=A_1X=Q_2$ 的解,那么 $\mathrm{rs}Q_2\cdot(I_n\otimes A-A_1'\otimes I_n)^{-1}=\mathrm{rs}X$, $\mathrm{rs}Q_2\cdot(I_n\otimes A-A_1'\otimes I_n)^{-1}\cdot(I_m\otimes B)=\mathrm{rs}(XB)$, 因此, $\dfrac{\partial J}{\partial U_{pq}}=\mathrm{rs}(Q_1+XB)\cdot\dfrac{\partial\mathrm{cs}U}{\partial U_{pq}}$.

最后可得

$$\partial J/\partial U=(Q_1+XB)', \tag{4.4}$$

其中 X 满足 $XA-A_1X=Q_2$. $Q_1=V^{-1}(\partial J/\partial F)'$, $Q_2=V^{-1}(\partial J/\partial F)'F$, $\partial J/\partial F=\dfrac{\partial\rho(\varepsilon,F)}{\partial F}$ 由 (3.3) 式给出.

以上 $\partial J/\partial U$ 的推导方法类似 [4] 中各梯度的推导,符号 cs[], rs[] 分别表示列展

开和行展开,"$\otimes$"为 Kronecker 积

5 算法

下面给出求解 (4.1) 式的近似方法,即求解 (4.3) 式的梯度算法.

首先,用两分法求 $\rho(\varepsilon,F)$. 关于 ε 的选择,理论上讲,ε 越趋近于 0, $\rho(\varepsilon,F)$ 越逼近 $1/\|T(s,F)\|_\infty$,但这样会增大计算量和造成算法奇异性. 一般取

$$\varepsilon=\left(\frac{1}{20}\sim\frac{1}{10}\right)\cdot\min_{i\in n}|\mathrm{Re}\lambda_i[A_1]|.$$

具体计算时,要求得精确的 $\rho(\varepsilon,F)$,或是求得使 $\mathrm{Re}\lambda_i[R(F,k)]=-\varepsilon$ 的精确 k 需要花费大量的机时,没有必要。一定范围的误差是允许的. 取小正数 $e,e\ll\varepsilon$,那么,较易求得使 $-\varepsilon-e<\mathrm{Re}\,\lambda_i[R(F,k)]<-\varepsilon+e$ 的 k,因此,有以下算法求近似的 $\rho(\varepsilon,F)$.

设 $R(F,k)$ 的特征值按实部的递增规则排序,即

$$\mathrm{Re}\,\lambda_1\leqslant\mathrm{Re}\,\lambda_2\leqslant\cdots\leqslant\mathrm{Re}\,\lambda_n\leqslant O\leqslant\mathrm{Re}\,\lambda_{n+1}\leqslant\cdots\leqslant\mathrm{Re}\,\lambda_{2n}.$$

注意到 $R(F,k)$ 的特征值关于虚轴对称.

Step 1. 取初值 k.

Step 2. a) 如果 $\mathrm{Re}\lambda_n[R(F,k)]<-\varepsilon-e$,那么用 $2k$ 取代 k,重复此步骤; b) 如果 $\mathrm{Re}\,\lambda_n[R(F,k)]>-\varepsilon+e$,则令 $k_1=k$, $k_2=0$, 转 Step 3); c) $-\varepsilon-e\leqslant\mathrm{Re}\,\lambda_n[R(F,k)]\leqslant-\varepsilon+e$, 转至 Step 5).

Step 3. 令 $k=(k_1+k_2)/2$.

Step 4. a) 如果 $\mathrm{Re}\lambda_n[R(F,k)]<-\varepsilon-e$, 令 $k_2=k$, 转至 Step3); b) 如果 $\mathrm{Re}\lambda_n[R(F,k)]>-\varepsilon+e$, 令 $k_1=k$, 转至 Step 3); c) $-\varepsilon-e\leqslant\mathrm{Re}\,\lambda_n[R(F,k)]\leqslant-\varepsilon+e$, 转至 Step 5).

Step 5. k 即为 $\rho(\varepsilon,F)$ 的估计值. 另外,记满足 $-\varepsilon-e\leqslant\mathrm{Re}\,\lambda_i[R(F,k)]\leqslant-\varepsilon+e$, $\mathrm{Im}\lambda_i[R(F,k)]\geqslant0$ 的特征值的个数为N (N 被认为是 $R[F,\rho(\varepsilon,F)]$ 在 $-\varepsilon+j\omega$ 轴上半部分的特征值个数).

以上算法基于如下事实: 当 $k\geqslant1/\|T(s,F)\|_\infty$ 时, $R(F,k)$ 在虚轴上有特征值,当 $k<1/\|T(s,F)\|_\infty$ 时, $R(F,k)$ 在虚轴上无特征值,因而当 $\varepsilon\to0$ 时有, $k\leqslant\rho(\varepsilon,F)\Longleftrightarrow\mathrm{Re}\lambda_n[R(F,k)]\leqslant-\varepsilon$.

以下是梯度算法:

Step 1. 选择 U_0, 令 $j=0$.

Step 2. 解 $AV-VA_1=-BU_j$, 令 $F=U_jV^{-1}$, 用以上算法求近似的 $\rho(\varepsilon,F)$ 及 N. 如果 $N\geqslant2$, 则认为 $\partial J/\partial U$ 不存在, 转至 Step 5), 如果 $N=1$, 转至 Step 3).

Step 3. 求 $R(F,\rho(\varepsilon,F))$ 的左右特征向量 t' 和 v. 由 (3.3) 式求 $\partial\rho(\varepsilon,F)/\partial F$, 再由 (4.4) 式求 $\partial J/\partial U$. ($J=\rho(\varepsilon,F)$).

Step 4. 一维搜索 h,使 $J\left(U_j + h\dfrac{\partial T}{\partial U}\right)$ 最大。如果 $J\left(U_j + h\dfrac{\partial J}{\partial U}\right) - J(U_j) \leqslant e_1$，转至 Step 5),否则,令 $U_{j+1} = U_j + h\dfrac{\partial J}{\partial U}$，用 $j+1$ 代替 j，转至 Step 2)。

Step 5. 输出终值 U^*，F^*，及 $J^* = \rho(\varepsilon, F^*)$，迭代结束。

6 算例

例 1. 开环系统同 [4] 中的例 1，原出自文 [6]

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

$$A = \begin{bmatrix} -5 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 0 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

给定闭环极点是 $-2 \pm j2, -4$。目标是通过选择状态反馈使复数稳定范围 $1/\|(sI - A - BF)^{-1}\|_\infty$ 在极点配置的约束下最大，与 (4.1) 式所描述的 H$^\infty$-优化问题相同。取 $H = G = I_3$，以及取 $\varepsilon = 0.1$，选择 $U_0 = \begin{bmatrix} 0 & 2 & 1 \\ 1 & 1 & 0 \end{bmatrix}$，$e = 0.02$，那么由梯度方法得到结果见表 1。

表 1

迭代次数	0	1	2	…	8	9
$J = \rho(\varepsilon, F)$	0.12502	1.33398	1.61914	…	1.66211	1.66211
$\|\partial J/\partial U\|_F^2$	2.91461	1.66608	0.28976	…	0.00013	0.00003

迭代在第 9 步停止,此处 $\partial J/\partial U \approx 0$，$F^*$ 为

$$F^* = \begin{bmatrix} 0.02546 & -2.72979 & -2.65177 \\ -3.26085 & -1.34822 & 1.03045 \end{bmatrix},$$

$\|(sI - A - BF^*)^{-1}\|_\infty = 0.60092$，$1/\|(sI - A - BF^*)^{-1}\|_\infty = 1.66411$，略大于 $\rho(\varepsilon, F^*) = 1.66211$。

若取不同的初值 U_0，可得其它驻点,例如

$$F_1^* = \begin{bmatrix} -1.99222 & 0.72296 & -2.64439 \\ -3.27407 & -1.35560 & -3.02955 \end{bmatrix}.$$

这一 F_1^* 所对应的 $\rho(\varepsilon, F_1^*) = 1.66260$，$1/\|(sI - A - BF_1^*)^{-1}\|_\infty = 1.66411$，$\|(sI - A - BF_1^*)^{-1}\|_\infty = 0.60092$，与前一 F^* 的几乎相同。

用本文方法所得结果明显好于文[4,6],其中,文[4]中的 $1/\|(sI - A - BF)^{-1}\|_\infty$ 的最大值是 1.65865，而文 [6] 的最大值是 1.5791。

例 2. 开环系统同文 [4] 的例 2，原出自文 [7]。设计目标同例 1。其中，$n = 5$，$m = 2$，A、B、A_1 同 [4]，$H = G = I_5$。

取 $\varepsilon = 0.01$，$e = 0.002$，$U_0 = \begin{bmatrix} -1 & 0 & 0.2 & 2 & -0.7 \\ 0 & -1 & -3 & 0.3 & 1 \end{bmatrix}$，则梯度算法的结果见表 2.

表 2

迭代次数	0	1	2	…	7	8
$J=\rho(\varepsilon,F)$	0.02210	0.04805	0.05898	…	0.11138	0.11147
$\Vert\partial J/\partial U\Vert_F^2$	0.00283	0.00095	0.00175	…	0.00259	无

在第 8 步，$R(F_8,\rho(\varepsilon,F_8))$ 在 $-\varepsilon+j\omega$ 轴附近有三个特征值：$-0.00977\pm jx$ 和 -0.01129. 因此认为 $\partial J/\partial U$ 不存在,并终止迭代.

$$F_8 = \begin{bmatrix} -21.72166 & 76.53389 & -121.1476 & 80.40894 & -38.9798 \\ -11.48118 & 14.58577 & 2.26744 & -19.05556 & 4.46882 \end{bmatrix}.$$

$\rho(\varepsilon,F_8)=0.11147$，$1/\Vert(sI-A-BF_8)^{-1}\Vert_\infty=0.11150$，$\Vert(sI-A-BF_8)^{-1}\Vert_\infty=8.96861$. 虽然这不是最优解,但梯度方法的效果是显然的，且比文 [4,7] 的结果好. 文 [7]中的 $\Vert(sI-A-BF)^{-1}\Vert_\infty$ 的最小值是 14.353，文 [4] 的最小值是 10.80030.

以上例子都显示了本文方法的有效性及梯度的准确性.

参 考 文 献

[1] 胡庭姝等. 通过传递函数的状态空间实现求 H^∞-范数. 自动化学报,1991,**17**(2): 215—219.

[2] Khargoneker P P, et al. H^∞-Optimal control with state feedback. IEEE. 1988, **AC**—33: 786—788.

[3] Hinrichsen D, et al. Stability radius for structured perturbations and the algebraic Riccati equation. *Systems and Contol Letters*, 1986, **8**: 105—113.

[4] 胡庭姝等,鲁棒设计的新方法——$\bar{\sigma}[p]$ 和 $\bar{\sigma}[v]\bar{\sigma}[v^{-1}]$ 的极小化。自动化学报，1994，**20**(2): 129—137.

[5] Hu T S, et al. The set of feedback matrices that assign the poles of a system. Proc. of MTNS-89, **2**: 129—135.

[6] Dickman A. On the robustness of multivariable linear feedback systems in state space representation, IEEE, 1987, **AC-32**: 407—410.

[7] Kautsky J, et al. Robust pole assignment in linear state feedback. *Int. J. Control*, 1985, **41**: 1129—1155.

A GRADIENT APPROACH TO H$^\infty$-OPTIMIZATION

HU TINGSHU　　CHEN LI

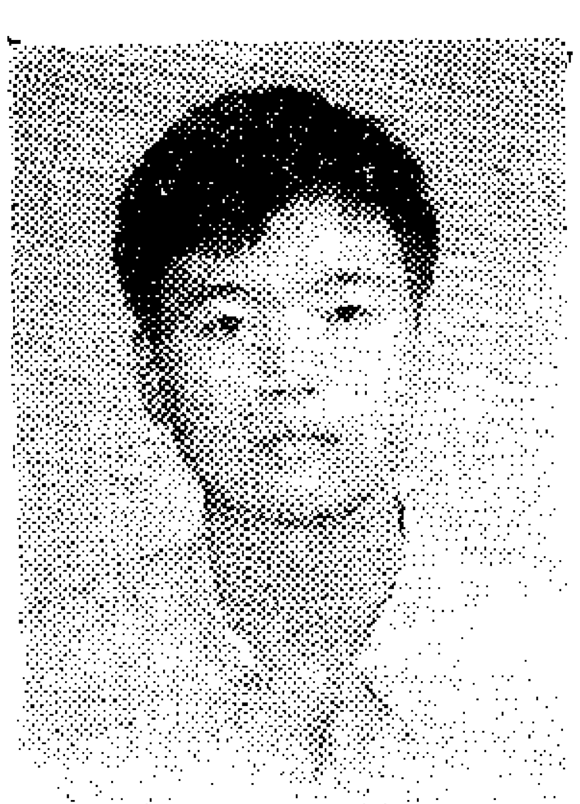
(*Dept. of Autonatic Control, Shanghai Jiaotong Univ. Shanghai* 200030)

ABSTRACT

In this paper, a gradient approach to H$^\infty$-optimization is presented. This new approach is very effective and flexible. Through the relation between the H$^\infty$-norm and state-space representation, an alternative performance index $\rho(\varepsilon, F)$ is defined, with the relation $\lim\limits_{\varepsilon\to 0}\rho(\varepsilon, F) = \|T(s, F)\|_\infty^{-1}$. The differentiability of $\rho(\varepsilon, F)$ with respect to F is investigated and $\partial\rho(\varepsilon, F)/\partial F$ is provided. A gradient algorithm is derived to maximize $\rho(\varepsilon, F)$, and hence to minimize $\|T(s, F)\|_\infty$. Examples show that the gradient algorithm is very effective in increasing $\rho(\varepsilon, F)$. The algorithm converges to stationary points or stops at non-differentiable points.

Key words: H$^\infty$-norm, gradient method, differentiability, pole assignment.

胡庭姝　照片、简介见本刊第 20 卷第 20 期.



陈　力　生于 1973 年. 1993 年毕业于上海交通大学自动控制系,获学士学位,现在该系攻读硕士学位. 研究方向是鲁棒控制.