

Robust
Robust

鲁棒控制

——线性矩阵不等式 处理方法

俞立 著



清华大学出版社

鲁棒控制

——线性矩阵不等式处理方法

俞立著

清华大学出版社

(京)新登字 158 号

内 容 简 介

本书结合作者的研究工作,详细介绍了基于线性矩阵不等式的不确定系统鲁棒控制的概念、理论及设计方法。主要包括目前应用广泛的线性矩阵不等式的概念、理论、算法及相关软件;基于线性矩阵不等式处理方法的线性时不变系统性能分析和综合方法;重点介绍了不确定系统的模型、鲁棒性能分析、鲁棒 H_2/H_∞ 控制、LMI区域及相应的区域极点配置方法、结合二次型性能指标的保性能控制、鲁棒方差控制、时滞系统的分析与鲁棒控制器设计、不确定系统的鲁棒滤波问题及鲁棒滤波器设计。

本书反映了近年来鲁棒控制领域中的最新研究成果,系统介绍了线性矩阵不等式这一有效工具,它在应用中的典型处理方法及MATLAB软件中的LMI工具箱。

本书可作为从事自动控制工作的科研人员、工程技术人员以及高等院校自动化及其他相关专业教师、高年级学生和研究生的参考用书。

版权所有,翻印必究。

本书封面贴有清华大学出版社激光防伪标签,无标签者不得销售。

书 名:鲁棒控制——线性矩阵不等式处理方法
作 者:俞 立 著
出 版 者:清华大学出版社(北京清华大学学研大厦,邮编 100084)
<http://www.tup.tsinghua.edu.cn>
责任编辑:朱英彪
印 刷 者:北京通州区大中印刷厂
发 行 者:新华书店总店北京发行所
开 本:787×1092 1/16 印张:17.75 字数:403 千字
版 次:2002年12月第1版 2002年12月第1次印刷
书 号:ISBN 7-302-05854-7/O·269
印 数:0001~4000
定 价:26.00 元

前 言

在实际工业控制中，各种工业生产过程、生产设备以及其他众多的被控对象，其动态特性一般都难以用精确的数学模型来描述。有时即使能获得被控对象的精确数学模型，但由于过于复杂，使得难以对其进行有效的控制性能分析和综合，因此必须进行适当的简化。另一方面，随着生产过程中工作条件和环境的变化，控制系统中元器件的老化或损坏，被控对象本身的特性也会随之发生变化。所有这些因素使得描述被控对象的数学模型和实际对象之间不可避免地具有误差。因此，在工程实践中，采用基于精确数学模型是现代控制理论方法所设计的控制系统往往难以具有所期望的性能，甚至连系统的稳定性都难以得到保证。鲁棒控制理论结合系统模型参数不确定性和外部扰动不确定性的考虑，研究系统的鲁棒性能分析和综合问题，弥补了现代控制理论需要对象精确数学模型的缺陷，使得系统的分析和综合方法更加有效、实用。

鲁棒控制自提出以来，很快受到了人们的广泛重视和研究，取得了一系列的研究结果和方法，并在一些工程领域中获得了成功的应用。特别地，随着线性矩阵不等式及求解凸优化问题的内点法的提出，为许多控制问题的分析和求解提供了有效工具。MATLAB 软件中线性矩阵不等式工具箱的推出使得各种线性矩阵不等式问题的求解更加方便、直接，从而，进一步推动了线性矩阵不等式处理方法在系统和控制领域中的应用。

本书系统介绍了线性矩阵不等式的概念、性质、求解线性矩阵不等式相关问题的算法以及 MATLAB 软件中的线性矩阵不等式工具箱。结合作者的研究工作，介绍了基于线性矩阵不等式的鲁棒控制性能分析和综合方法，并指出了一些在系统和控制中有广泛应用价值的线性矩阵不等式典型处理方法。作者努力将鲁棒控制的最新研究成果和方法反映在本书中，但限于篇幅，书中所包含的内容仅仅是鲁棒控制研究成果的很少一部分。限于作者的水平，书中不妥和错误之处在所难免，恳请广大读者批评指正。

本书中介绍的作者研究工作及本书的撰写得到了国家自然科学基金和教育部高校优秀青年教师教学科研奖励计划的资助，在此表示衷心的感谢。

作者
2002 年 5 月
杭州

目 录

第 1 章 引言	1
第 2 章 线性矩阵不等式	6
2.1 线性矩阵不等式的表示式	6
2.1.1 线性矩阵不等式的一般表示	6
2.1.2 可转化成线性矩阵不等式表示的问题	7
2.1.3 复线性矩阵不等式的处理	10
2.1.4 非严格线性矩阵不等式	11
2.2 一些标准的线性矩阵不等式问题	11
2.3 求解线性矩阵不等式问题的算法	14
2.3.1 椭球法	15
2.3.2 内点法	16
2.4 关于矩阵不等式的一些结论	18
2.4.1 矩阵变量的消去法	18
2.4.2 S-procedure	20
第 3 章 系统性能分析	23
3.1 连续时间系统	23
3.1.1 系统增益指标	23
3.1.2 H_2 性能	29
3.1.3 H_∞ 性能	31
3.2 离散时间系统	34
第 4 章 控制系统综合	41
4.1 H_∞ 控制	41
4.1.1 状态反馈 H_∞ 控制	42
4.1.2 输出反馈 H_∞ 控制	44
4.2 H_2 控制	59
4.3 H_2/H_∞ 控制	61
4.4 设计示例	64
第 5 章 不确定系统的分析与综合	68
5.1 不确定模型	68

5.1.1	不确定状态空间模型	68
5.1.2	不确定线性分式模型	72
5.2	鲁棒稳定性分析	75
5.2.1	二次稳定性	75
5.2.2	仿射二次稳定性	79
5.3	鲁棒性能分析	83
5.4	鲁棒 H_2/H_∞ 控制	86
5.4.1	问题的描述和准备	86
5.4.2	H_2/H_∞ 控制器设计	91
第 6 章	区域极点配置	96
6.1	LMI 区域	97
6.1.1	LMI 区域的描述	97
6.1.2	D-稳定性分析	100
6.2	具有闭环区域极点约束的状态反馈控制器设计	104
6.3	鲁棒 D-稳定性分析	107
6.3.1	无结构不确定性	109
6.3.2	结构不确定性	114
6.4	输出反馈控制器设计	119
第 7 章	保性能控制	122
7.1	连续系统的保性能控制	122
7.2	离散系统的保性能控制	127
7.3	具有闭环极点约束的保性能控制	131
7.3.1	鲁棒性能分析	132
7.3.2	二次 D-保性能控制器设计	135
第 8 章	鲁棒方差控制	141
8.1	连续系统的鲁棒方差控制	141
8.1.1	系统性能分析	141
8.1.2	状态反馈控制器设计	144
8.1.3	输出反馈控制器设计	146
8.2	离散系统的鲁棒方差控制	152
第 9 章	时滞系统的分析和综合	158
9.1	时滞系统的稳定性	158
9.1.1	时滞独立的稳定性条件	159
9.1.2	时滞依赖的稳定性条件	160
9.1.3	Lurie 时滞系统的稳定性分析	163

9.2 时滞系统的鲁棒稳定性分析	169
9.2.1 时滞独立的鲁棒稳定性条件	169
9.2.2 时滞依赖的鲁棒稳定性条件	174
9.3 不确定时滞系统的保性能控制	178
9.3.1 鲁棒性能分析	178
9.3.2 状态反馈保性能控制器设计	183
9.3.3 输出反馈保性能控制器设计	186
9.3.4 不确定离散时滞系统的保性能控制	193
9.4 时滞系统的 H_∞ 控制	199
9.4.1 时滞系统的 H_∞ 性能分析	199
9.4.2 H_∞ 控制器设计	202
9.4.3 不确定离散时滞系统的鲁棒 H_∞ 控制	207
第 10 章 滤波器设计	213
10.1 H_∞ 滤波器设计	213
10.2 H_2/H_∞ 滤波器设计	219
第 11 章 大系统的分散控制	223
11.1 时滞系统的分散稳定化控制	223
11.2 离散关联系统的分散保性能控制	229
11.2.1 保性能分析	229
11.2.1 分散保性能控制器设计	234
附录 A LMI 工具箱介绍	241
A.1 线性矩阵不等式及相关术语	241
A.2 线性矩阵不等式的确定	242
A.3 信息提取	249
A.4 线性矩阵不等式求解器	250
A.5 结果验证	258
A.6 修改一个线性矩阵不等式系统	259
A.7 一些进一步的功能	261
A.8 系统模型描述	267
参考文献	270

第 1 章 引言

自 20 世纪 50 年代末现代控制理论诞生以来, 控制理论得到了飞速的发展, 并在 20 世纪 60 年代的航天领域中得到成功的应用。但是, 现代控制理论在随后的工业应用中却遇到了很大的困难。我们知道, 现代控制理论的许多结果都是基于对象的一个数学模型, 根据系统的性能要求, 通过对被控对象的数学模型进行分析来设计系统的控制律, 进而将所得到的控制律应用于被控对象来保证闭环系统具有所期望的性能。显然, 当对象模型不能精确地描述被控对象或在系统运行过程中模型和实际对象产生偏离时, 基于这样的模型设计的控制系统很难保证具有所期望的性能要求。

实际上, 对于复杂物理系统的模型, 存在以下两个问题:

1. 描述物理系统的解析模型很难, 甚至不可能精确地刻画, 因此为了便于处理, 不得不简化模型;

2. 一个模型, 无论多么详细, 都不可能是物理系统的一个精确表示。因此, 模型存在本质的不精确性。

建模中的以上两个方面称为模型的不确定性。

对于一个复杂系统, 为了得到一个较为简单的模型, 一种处理方法是将其分解成线性部分和非线性部分的组合, 进而用一个更容易处理和分析的对象来替代这个非线性部分, 达到简化原来复杂系统模型的目的。

考虑由以下非线性微分方程描述的复杂动态系统:

$$\begin{aligned}\dot{\mathbf{x}} &= \mathbf{f}(\mathbf{x}, \mathbf{u}) \\ \mathbf{y} &= \mathbf{h}(\mathbf{x}, \mathbf{u})\end{aligned}\tag{1.1}$$

初始条件是 $\mathbf{x}(0)$, $\mathbf{x}(t)$ 、 $\mathbf{y}(t)$ 和 $\mathbf{u}(t)$ 是向量值函数, \mathbf{f} 和 \mathbf{h} 是光滑的向量值函数。在一个特殊的运行点附近, 可以将系统 (1.1) 分解成一个线性部分和一个非线性部分的组合。特别地, 可以在原点 $(\mathbf{x}, \mathbf{u}) = (\mathbf{0}, \mathbf{0})$ 处进行这样的分解。定义系统:

$$\begin{aligned}\dot{\mathbf{x}} &= \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u} + \mathbf{g}(\mathbf{x}, \mathbf{u}) \\ \mathbf{y} &= \mathbf{C}\mathbf{x} + \mathbf{D}\mathbf{u} + \mathbf{r}(\mathbf{x}, \mathbf{u})\end{aligned}\tag{1.2}$$

其中: \mathbf{A} 、 \mathbf{B} 、 \mathbf{C} 和 \mathbf{D} 是系统 (1.1) 的一个线性化近似,

$$\begin{aligned}\mathbf{g}(\mathbf{x}, \mathbf{u}) &= \mathbf{f}(\mathbf{x}, \mathbf{u}) - \mathbf{A}\mathbf{x} - \mathbf{B}\mathbf{u} \\ \mathbf{r}(\mathbf{x}, \mathbf{u}) &= \mathbf{h}(\mathbf{x}, \mathbf{u}) - \mathbf{C}\mathbf{x} - \mathbf{D}\mathbf{u}\end{aligned}$$

显然, 这样定义的系统 (1.2) 和系统 (1.1) 是等价的。因此, 它们之间存在一一对应的关系。得到这样的等价系统的一种方式是将函数 \mathbf{f} 和 \mathbf{h} 在原点处线性化, 可得:

$$\mathbf{A} = \left. \frac{\partial \mathbf{f}}{\partial \mathbf{x}} \right|_{(\mathbf{x}, \mathbf{u})=(\mathbf{0}, \mathbf{0})}, \quad \mathbf{B} = \left. \frac{\partial \mathbf{f}}{\partial \mathbf{u}} \right|_{(\mathbf{x}, \mathbf{u})=(\mathbf{0}, \mathbf{0})}, \quad \mathbf{C} = \left. \frac{\partial \mathbf{h}}{\partial \mathbf{x}} \right|_{(\mathbf{x}, \mathbf{u})=(\mathbf{0}, \mathbf{0})}, \quad \mathbf{D} = \left. \frac{\partial \mathbf{h}}{\partial \mathbf{u}} \right|_{(\mathbf{x}, \mathbf{u})=(\mathbf{0}, \mathbf{0})}$$

进一步可以将方程 (1.2) 写成以下等价的形式:

$$\dot{x} = Ax + Bu + w_1 \quad (1.3)$$

$$y = Cx + Du + w_2 \quad (1.4)$$

$$(w_1, w_2) = (g(x, u), r(x, u)) \quad (1.5)$$

设 G 是由 (1.3) ~ (1.4) 式确定的映射: 对给定的初始条件 $x(0)$, $(w_1, w_2, u) \mapsto (x, u, y)$ 。 Q 是由 (1.5) 式确定的映射: $(x, u) \mapsto (w_1, w_2)$ 。因此, (1.3) ~ (1.5) 式描述的系统可以用图 1.1 来表示。

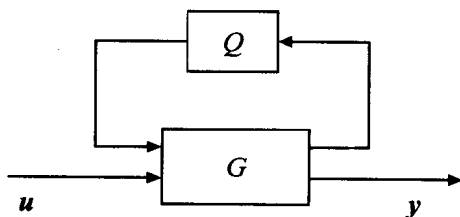


图 1.1 系统分解

容易看到, G 是系统的线性部分, Q 是静态的非线性映射。这样就将系统的非线性部分分离出来, 归入到映射 Q 中, 非线性部分和线性部分通过反馈关联联系起来。

更一般地, 我们用这样的方法不仅可以处理系统的非线性特性, 而且也可以处理系统的某些动态特性。考虑由以下方程组描述的系统:

$$\begin{cases} \dot{x}_1 \\ \dot{x}_2 \end{cases} = \begin{bmatrix} f_1(x_1, x_2, u) \\ f_2(x_1, x_2, u) \end{bmatrix} \quad (1.6) \\ y = h(x_1, x_2, u)$$

采用前面系统分解的思想, 对系统 (1.6) 中的方程 $\dot{x}_1 = f_1(x_1, x_2, u)$, $y = h(x_1, x_2, u)$ 进行分解, 并得到:

$$\begin{aligned} \dot{x}_1 &= A_1 x_1 + B_1 u + g_1(x_1, x_2, u) \\ \dot{x}_2 &= f_2(x_1, x_2, u) \\ y &= C_1 x_1 + D_1 u + r(x_1, x_2, u) \end{aligned} \quad (1.7)$$

进一步, 系统 (1.7) 中的方程等价于以下的线性方程:

$$\begin{aligned} \dot{x}_1 &= A_1 x_1 + B_1 u + w_1 \\ y &= C_1 x_1 + D_1 u + w_2 \end{aligned} \quad (1.8)$$

其中:

$$\begin{aligned} (w_1, w_2) &= (g_1(x_1, x_2, u), r(x_1, x_2, u)) \\ \dot{x}_2 &= f_2(x_1, x_2, u) \end{aligned} \quad (1.9)$$

设 G 是由方程 (1.8) 描述的线性系统: $(w_1, w_2, u) \mapsto (x_1, u, y)$, Q 是由 (1.9) 式描述的系统: $(x_1, u) \mapsto (w_1, w_2)$ 。对这样定义的 G 和 Q , 图 1.1 也同样描述了系统 (1.6)。

注意到在这个结构中，系统的部分动态行为和所有的非线性行为被从原来的系统中分离出来，包含在映射 Q 中。

在上面介绍的系统分解中，将一个复杂的非线性系统分解成一个线性部分 G 和一个非线性部分 Q 的关联。其中的 Q 不仅是非线性的，还可能包含动态特性，因此它往往是比较复杂的。如果能找到一种更容易处理的对象来替代映射 Q ，则这样的分解会使得整个系统的分析变得更加方便和容易，其意义也才会显现出来。以下介绍具有这样性质的替代映射 Q 的一个映射集合 Δ 。

集合 Δ 是具有以下性质的一个映射集合：如果对某个输入 p ， $q = Q(p)$ ，则存在一个映射 $\Delta \in \Delta$ ，使得 $q = \Delta(p)$ 。如果集合 Δ 具有这样的性质，则：

$$\{(p, q) : q = Q(p)\} \subset \{(p, q) : \text{存在 } \Delta \in \Delta, \text{ 使得 } q = \Delta(p)\}$$

显然，上式左边的集合定义了一个函数，而右边由 Δ 产生的输入输出对一般来说只是一种关系。这两个集合相差的程度确定了由 Δ 来替代 Q 后所产生的保守程度。

之所以引进集合 Δ 是因为一般来说集合 Δ 中的映射比映射 Q 要简单得多。例如，当 Q 是一个非线性映射时，可以用一组线性映射 Δ 来产生所有可能的满足 $q = Q(p)$ 的输入输出对 (p, q) 。因此，我们可以用简单的线性映射集合 Δ 来替代这个非线性映射 Q 。

现在利用集合 Δ 来分析图 1.1 描述的系统。设 $S(G, Q)$ 表示图 1.1 中的映射： $u \mapsto y$ ，如果用映射集 $S(G, \Delta) = \{S(G, \Delta) : \Delta \in \Delta\}$ 来替代映射 $S(G, Q)$ ，则可以看到：如果所有映射 $S(G, \Delta)$ 所描述的系统具有一个给定的性质，则映射 $S(G, Q)$ 所描述的系统也必定具有这样的性质。然而，反之却未必成立。因此，应用这样的方法来分析系统的性质具有一定的保守性。

例 1.1 考虑系统 (1.1) 的分解。为了简化，假定 x 、 y 和 u 都是标量值函数。假设定义映射 Q 的函数 r 和 g 满足以下的不等式：

$$\begin{aligned} |w_1(t)| &\leq k_{11}|x(t)| + k_{12}|u(t)| \\ |w_2(t)| &\leq k_{21}|x(t)| + k_{22}|u(t)| \end{aligned} \quad (1.10)$$

其中： k_{ij} 是已知的正常数。则对某个给定的 $(w_1, w_2) = Q(x, u)$ ，存在标量函数 $\delta_{11}(t)$ 、 $\delta_{12}(t)$ 、 $\delta_{21}(t)$ 和 $\delta_{22}(t)$ ，其中的每个 $\delta_{ij}(t) \in [-k_{ij}, k_{ij}]$ ，使得：

$$\begin{aligned} w_1(t) &= \delta_{11}x(t) + \delta_{12}u(t) \\ w_2(t) &= \delta_{21}x(t) + \delta_{22}u(t) \end{aligned} \quad (1.11)$$

定义集合 Δ ：

$$\Delta = \left\{ \Delta = \begin{bmatrix} \delta_{11}(t) & \delta_{12}(t) \\ \delta_{21}(t) & \delta_{22}(t) \end{bmatrix} : |\delta_{ij}(t)| \leq k_{ij} \right\}$$

则根据前面的分析，对满足 $(w_1, w_2) = Q(x, u)$ 的任意输入输出，存在 $\Delta \in \Delta$ ，使得关系式 (1.11) 成立。

例 1.2 考虑系统 (1.6) 的分解，注意到在这种情况下， Q 所表示的可能是一个动态系统。假定 x_1 和 u 是标量函数，并且假定如果 $(w_1, w_2) = Q(x_1, u)$ ，则以下的能量不等式

成立:

$$\begin{aligned} \int_0^{\infty} |w_1(t)|^2 dt &\leq k_1 \left(\int_0^{\infty} |x_1(t)|^2 dt + \int_0^{\infty} |u(t)|^2 dt \right) \\ \int_0^{\infty} |w_2(t)|^2 dt &\leq k_2 \left(\int_0^{\infty} |x_1(t)|^2 dt + \int_0^{\infty} |u(t)|^2 dt \right) \end{aligned} \quad (1.12)$$

其中: (1.12) 式右边的积分假定是有限的。定义映射集合 Δ 是所有满足不等式 (1.12) 的线性映射 $\Delta: (x_1, u) \mapsto (w_1, w_2)$ 的全体, 则对某个有界能量函数 x_1 和 u , 如果 $(w_1, w_2) = Q(x_1, u)$, 则存在一个映射 $\Delta \in \Delta$, 使得 $\Delta(x_1, u) = (w_1, w_2)$ 。因此, 集合 Δ 中的所有映射将覆盖映射 Q 所产生的输入输出对。

在系统分析中, 用映射集合 Δ 来替代映射 Q 除了以上目的以外, 还有一个目的就是: 任何一个物理系统都不可能用一个数学模型精确表示。由于对系统某些特性或环节缺乏足够的了解 (即难以建模的部分)、系统环境的变化、元器件的老化、某些物理参数的漂移或随时间的未知变化等因素所带来的系统行为的变化都有可能产生模型不确定性的产生。因此, 集合 Δ 的引进提供了一种解释不能精确模型化的那部分系统行为。

总结起来, 集合 Δ 的引进可以为以下两个相关但不同的目的服务:

- 提供一种简化模型的方法
- 用于模型和解释不确定的动态行为

对用这一类模型描述的系统的分析和综合问题称为参数不确定系统的鲁棒性能分析和综合问题, 它是近 20 年来国际自动控制界最活跃的研究领域之一, 吸引了大量研究人员对其进行深入研究, 提出了一系列诸如 H_{∞} 控制、 μ 方法等新的研究结果和方法, 开拓了许多新的研究分支, 一些研究结果在许多工业控制领域得到了成功的应用。

在时间域中研究参数不确定系统的鲁棒分析和综合问题的主要理论基础是 Lyapunov 稳定性理论, 早期的一种主要方法是 Riccati 方程处理方法。它是通过将系统的鲁棒分析和综合问题转化成 Riccati 型矩阵方程的可解性问题, 进而应用求解 Riccati 方程的方法给出系统具有给定鲁棒性能的条件和鲁棒控制器的设计方法。尽管 Riccati 方程处理方法可以给出控制器的结构形式, 便于进行一些理论分析, 但是在实施这一方法之前, 往往需要设计者事先确定一些待定参数, 这些参数的选择不仅影响到结论的好坏, 而且还会影响到问题的可解性。但在现有的 Riccati 方程处理方法中, 还缺乏寻找这些参数最佳值的方法, 参数的这种人为确定方法给分析和综合结果带来了很大的保守性。另一方面, Riccati 型矩阵方程本身的求解也还存在一定的问题。目前存在很多求解 Riccati 型矩阵方程的方法, 但多为迭代方法, 这些方法的收敛性并不能得到保证。

20 世纪 90 年代初, 随着求解凸优化问题的内点法的提出, 线性矩阵不等式再一次受到控制界的关注, 并被应用到系统和控制的各个领域。许多控制问题可以转化为一个线性矩阵不等式系统的可行性问题, 或者是一个具有线性矩阵不等式约束的凸优化问题。由于有了求解凸优化问题的内点法, 使得这些问题可以得到有效的解决。1995 年, MATLAB 推出了求解线性矩阵不等式问题的 LMI 工具箱, 从而使得人们能够更加方便和有效地来处理、求解线性矩阵不等式系统, 进一步推动了线性矩阵不等式方法在系统和控制领域中的应用。

线性矩阵不等式处理方法可以克服 Riccati 方程处理方法中存在的许多不足。线性矩阵不等式方法给出了问题可解的一个凸约束条件, 因此, 可以应用求解凸优化问题的有效方法来进行求解。正是这种凸约束条件, 使得在控制器设计时, 得到的不仅仅是一个满足设计要求的控制器, 而是从凸约束条件的任意一个可行解都可以得到一个控制器, 即可以得到满足设计要求的一组控制器。这一性能在求解系统的多目标控制问题时是特别有用的。

本书将结合作者的研究工作, 较为系统地介绍线性矩阵不等式的概念、求解方法、LMI 工具箱、鲁棒控制的线性矩阵不等式处理方法及一些相关的研究成果。

第 2 章 线性矩阵不等式

近 10 年来, 线性矩阵不等式被广泛用来解决系统与控制中的一些问题, 随着解决线性矩阵不等式的内点法的提出、MATLAB 软件中 LMI 工具箱的推出, 线性矩阵不等式这一工具越来越受到人们的注意和重视, 应用线性矩阵不等式来解决系统与控制问题已成为这些领域中的一大研究热点。

本章主要介绍线性矩阵不等式的一些基本概念、求解线性矩阵不等式的主要算法以及应用线性矩阵不等式来解决系统与控制问题时要用到的一些基本结论。

2.1 线性矩阵不等式的表示

2.1.1 线性矩阵不等式的一般表示

一个线性矩阵不等式就是具有形式

$$F(x) = F_0 + x_1 F_1 + \cdots + x_m F_m < 0 \quad (2.1.1)$$

的一个表达式。其中 x_1, \dots, x_m 是 m 个实数变量, 称为线性矩阵不等式 (2.1.1) 的决策变量, $x = (x_1, \dots, x_m)^T \in \mathbf{R}^m$ 是由决策变量构成的向量, 称为决策向量, $F_i = F_i^T \in \mathbf{R}^{n \times n}$, $i = 0, 1, \dots, m$ 是一组给定的实对称矩阵, (2.1.1) 式中的不等号 “<” 指的是矩阵 $F(x)$ 是负定的, 即对所有非零的向量 $v \in \mathbf{R}^n$, $v^T F(x)v < 0$, 或者 $F(x)$ 的最大特征值小于零。

如果把 $F(x)$ 看成是从 \mathbf{R}^m 到实对称矩阵集 $S^n = \{M: M = M^T \in \mathbf{R}^{n \times n}\}$ 的一个映射, 则可以看出 $F(x)$ 并不是一个线性函数, 而只是一个仿射函数。因此, 更确切地说, 不等式 (2.1.1) 应该称为一个仿射矩阵不等式。但由于历史原因, 目前线性矩阵不等式这一名称已被广泛接受和使用。

在许多系统与控制问题中, 问题的变量是以矩阵的形式出现的。例如 Lyapunov 矩阵不等式:

$$F(X) = A^T X + XA + Q < 0 \quad (2.1.2)$$

其中: $A, Q \in \mathbf{R}^{n \times n}$ 是给定的常数矩阵, 且 Q 是对称的, $X \in \mathbf{R}^{n \times n}$ 是对称的未知矩阵变量, 因此该矩阵不等式中的变量是一个矩阵。设 E_1, E_2, \dots, E_M 是 S^n 中的一组基, 则对任意对称矩阵 $X \in \mathbf{R}^{n \times n}$, 存在 x_1, x_2, \dots, x_M , 使得 $X = \sum_{i=1}^M x_i E_i$ 。因此,

$$\begin{aligned}
 F(\mathbf{X}) &= F\left(\sum_{i=1}^M x_i \mathbf{E}_i\right) = \mathbf{A}^T \left(\sum_{i=1}^M x_i \mathbf{E}_i\right) + \left(\sum_{i=1}^M x_i \mathbf{E}_i\right) \mathbf{A} + \mathbf{Q} \\
 &= \mathbf{Q} + x_1 (\mathbf{A}^T \mathbf{E}_1 + \mathbf{E}_1 \mathbf{A}) + \cdots + x_M (\mathbf{A}^T \mathbf{E}_M + \mathbf{E}_M \mathbf{A}) \\
 &< \mathbf{0}
 \end{aligned}$$

即 Lyapunov 矩阵不等式 (2.1.2) 写成了线性矩阵不等式的一般形式 (2.1.1)。

如果在 (2.1.1) 式中用“ \leq ”代替“ $<$ ”，则相应的矩阵不等式称为非严格的线性矩阵不等式。对 $\mathbf{R}^m \rightarrow S^n$ 的任意仿射函数 $F(\mathbf{x})$ 和 $G(\mathbf{x})$ ， $F(\mathbf{x}) > \mathbf{0}$ ， $F(\mathbf{x}) < G(\mathbf{x})$ 也是线性矩阵不等式，因为它们可以等价地写成 $-F(\mathbf{x}) < \mathbf{0}$ ， $F(\mathbf{x}) - G(\mathbf{x}) < \mathbf{0}$ 。

所有满足线性矩阵不等式 (2.1.1) 的 \mathbf{x} 的全体构成一个凸集，这就是以下的引理 2.1.1。

引理 2.1.1 $\Phi = \{\mathbf{x} : F(\mathbf{x}) < \mathbf{0}\}$ 是一个凸集。

证明 对任意的 $\mathbf{x}_1, \mathbf{x}_2 \in \Phi$ 和任意的 $\alpha \in (0, 1)$ ，由于 $F(\mathbf{x}_1) < \mathbf{0}$ ， $F(\mathbf{x}_2) < \mathbf{0}$ 以及 $F(\mathbf{x})$ 是一个仿射函数，故

$$F(\alpha \mathbf{x}_1 + (1-\alpha)\mathbf{x}_2) = \alpha F(\mathbf{x}_1) + (1-\alpha)F(\mathbf{x}_2) < \mathbf{0}$$

所以 $\alpha \mathbf{x}_1 + (1-\alpha)\mathbf{x}_2 \in \Phi$ ，即 Φ 是凸的。从而引理 2.1.1 得证。

引理 2.1.1 说明了线性矩阵不等式 (2.1.1) 这个约束条件定义了自变量空间中的一个凸集，因此是自变量的一个凸约束。正是线性矩阵不等式的这一性质使得可以应用解决凸优化问题的有效方法来求解相关的线性矩阵不等式问题。

2.1.2 可转化成线性矩阵不等式表示的问题

系统与控制中的许多问题初看起来不是一个线性矩阵不等式问题，或不具有 (2.1.1) 式的形式，但可以通过适当的处理将问题转换成具有 (2.1.1) 式形式的一个线性矩阵不等式问题。下面给出了这方面的一些典型例子。

1. 多个线性矩阵不等式

$$F_1(\mathbf{x}) < \mathbf{0}, \dots, F_k(\mathbf{x}) < \mathbf{0}$$

称为一个线性矩阵不等式系统。引进 $F(\mathbf{x}) = \text{diag}\{F_1(\mathbf{x}), \dots, F_k(\mathbf{x})\}$ ，则 $F_1(\mathbf{x}) < \mathbf{0}, \dots, F_k(\mathbf{x}) < \mathbf{0}$ 同时成立当且仅当 $F(\mathbf{x}) < \mathbf{0}$ 。因此，一个线性矩阵不等式系统也可以用一个单一的线性矩阵不等式来表示。

2. 考虑问题

$$\begin{cases} F(\mathbf{x}) < \mathbf{0} \\ \mathbf{A}\mathbf{x} = \mathbf{b} \end{cases}$$

其中的 $F: \mathbf{R}^m \rightarrow S^n$ 是一个仿射函数， $\mathbf{A} \in \mathbf{R}^{n \times m}$ 和 $\mathbf{b} \in \mathbf{R}^n$ 是给定的常数矩阵和向量。由于 $\mathbf{A}\mathbf{x} = \mathbf{b}$ 的解向量的全体构成了 \mathbf{R}^m 中的一个线性子空间，因此可以考虑更一般的问题：

$$\begin{cases} F(x) < 0 \\ x \in M \end{cases} \quad (2.1.3)$$

其中的 M 是 \mathbf{R}^m 中的一个仿射集, 即

$$M = x_0 + M_0 = \{x_0 + m \mid m \in M_0\}$$

上式中的 $x_0 \in \mathbf{R}^m$, M_0 是 \mathbf{R}^m 中的一个线性子空间。以下证明这样一个多约束问题可以转化为一个单一的线性矩阵不等式约束。

设 $e_1, \dots, e_k \in \mathbf{R}^m$ 是线性空间 M_0 的一组基, 而仿射函数 $F(x)$ 可以分解成 $F(x) = F_0 + T(x)$, 其中 $T(x)$ 是一个线性函数。由于对任意的 $x \in M$, x 可以表示成 $x = x_0 + \sum_{i=1}^k x_i e_i$ 。因此, 问题 (2.1.3) 成立当且仅当

$$\begin{aligned} 0 > F(x) &= F_0 + T(x_0 + \sum_{i=1}^k x_i e_i) \\ &= F_0 + T(x_0) + \sum_{i=1}^k x_i T(e_i) \\ &= \tilde{F}_0 + x_1 \tilde{F}_1 + \dots + x_k \tilde{F}_k \\ &= \tilde{F}(\tilde{x}) \end{aligned}$$

其中: $\tilde{F}_0 = F_0 + T(x_0)$, $\tilde{F}_i = T(e_i)$, $\tilde{x} = [x_1, \dots, x_k]^T$ 。注意, \tilde{x} 的维数要小于 x 的维数。

3. 在许多将一些非线性矩阵不等式转化成线性矩阵不等式的问题中, 我们常常用到矩阵的 Schur 补性质。考虑一个矩阵 $S \in \mathbf{R}^{n \times n}$, 并将 S 进行分块:

$$S = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix}$$

其中的 S_{11} 是 $r \times r$ 维的。假定 S_{11} 是非奇异的, 则 $S_{22} - S_{21} S_{11}^{-1} S_{12}$ 称为 S_{11} 在 S 中的 Schur 补。以下引理给出了矩阵的 Schur 补性质。

引理 2.1.2 对给定的对称矩阵 $S = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix}$, 其中 S_{11} 是 $r \times r$ 维的。以下三个条件

是等价的:

- (i) $S < 0$;
- (ii) $S_{11} < 0$, $S_{22} - S_{12}^T S_{11}^{-1} S_{12} < 0$;
- (iii) $S_{22} < 0$, $S_{11} - S_{12} S_{22}^{-1} S_{12}^T < 0$ 。

证明 (i) \Leftrightarrow (ii) 由于 S 是对称的, 故有 $S_{11} = S_{11}^T$, $S_{22} = S_{22}^T$, $S_{21} = S_{12}^T$ 。应用矩阵块的初等运算, 可以得到

$$\begin{bmatrix} I & 0 \\ -S_{21} S_{11}^{-1} & I \end{bmatrix} \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} \begin{bmatrix} I & 0 \\ -S_{21} S_{11}^{-1} & I \end{bmatrix}^T = \begin{bmatrix} S_{11} & 0 \\ 0 & S_{22} - S_{21} S_{11}^{-1} S_{12} \end{bmatrix}$$

因此,

$$\begin{aligned} S < 0 &\Leftrightarrow \begin{bmatrix} I & \mathbf{0} \\ -S_{21}S_{11}^{-1} & I \end{bmatrix} \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} \begin{bmatrix} I & \mathbf{0} \\ -S_{21}S_{11}^{-1} & I \end{bmatrix}^T < 0 \\ &\Leftrightarrow \begin{bmatrix} S_{11} & \mathbf{0} \\ \mathbf{0} & S_{22} - S_{21}S_{11}^{-1}S_{12} \end{bmatrix} < 0 \\ &\Leftrightarrow \text{(ii)} \end{aligned}$$

这就证明了结论 (i) 和结论 (ii) 是等价的。

(i) \Leftrightarrow (iii) 注意到

$$\begin{bmatrix} I & -S_{12}S_{22}^{-1} \\ \mathbf{0} & I \end{bmatrix} \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} \begin{bmatrix} I & -S_{12}S_{22}^{-1} \\ \mathbf{0} & I \end{bmatrix}^T = \begin{bmatrix} S_{11} - S_{12}S_{22}^{-1}S_{12}^T & \mathbf{0} \\ \mathbf{0} & S_{22} \end{bmatrix}$$

类似于前面的证明即可以得到这一部分的结论。

综合以上两部分的证明, 可得引理的结论。

对线性矩阵不等式 $F(x) < 0$, 其中 $F(x) = \begin{bmatrix} F_{11}(x) & F_{12}(x) \\ F_{21}(x) & F_{22}(x) \end{bmatrix}$, $F_{11}(x)$ 是方阵。则应用矩阵的 Schur 补性质可以得到: $F(x) < 0$ 当且仅当

$$\begin{cases} F_{11}(x) < 0 \\ F_{22}(x) - F_{12}^T(x)F_{11}^{-1}(x)F_{12}(x) < 0 \end{cases} \quad (2.1.4)$$

或

$$\begin{cases} F_{22}(x) < 0 \\ F_{11}(x) - F_{12}(x)F_{22}^{-1}(x)F_{12}^T(x) < 0 \end{cases} \quad (2.1.5)$$

注意到 (2.1.4) 式或 (2.1.5) 式中的第二个不等式是一个非线性矩阵不等式, 因此以上的等价关系也说明了应用矩阵的 Schur 补性质, 一些非线性矩阵不等式可以转化成线性矩阵不等式。另一方面, 这一等价关系也说明了 (2.1.4) 式或 (2.1.5) 式中的非线性矩阵不等式也定义了一个关于变量 x 的凸约束。

在一些控制问题中, 经常遇到二次型矩阵不等式:

$$A^T P + PA + PBR^{-1}B^T P + Q < 0 \quad (2.1.6)$$

其中: $A, B, Q = Q^T > 0, R = R^T > 0$ 是给定的适当维数的常数矩阵, P 是对称矩阵变量, 则应用引理 2.1.2, 可以将矩阵不等式 (2.1.6) 的可行性问题转化成一个等价的矩阵不等式

$$\begin{bmatrix} A^T P + PA + Q & PB \\ B^T P & -R \end{bmatrix} < 0 \quad (2.1.7)$$

的可行性问题, 而后者是一个关于矩阵变量 P 的线性矩阵不等式。

2.1.3 复线性矩阵不等式的处理

前面讨论的线性矩阵不等式问题和 LMI 工具箱中的线性矩阵不等式求解器只能处理实的线性矩阵不等式。为了处理复线性矩阵不等式，我们需要将它们转化成实线性矩阵不等式的形式。注意到映射

$$a + jb \mapsto \begin{bmatrix} a & -b \\ -b & a \end{bmatrix}$$

已经建立起了复数空间 \mathbf{C} 和实矩阵空间 $\mathbf{R}^{2 \times 2}$ 之间的一个同构关系。因此，一个复矩阵 $\mathbf{M} = \mathbf{A} + j\mathbf{B}$ 可以等价地用一个增维的实矩阵 $\begin{bmatrix} \mathbf{A} & -\mathbf{B} \\ -\mathbf{B} & \mathbf{A} \end{bmatrix}$ 来表示。对一些复矩阵的关系式，

只要用复矩阵的等价实矩阵表示来替代其中的复矩阵，就可以得到相应等价的实矩阵表示式。例如：两个复矩阵 $\mathbf{M} = \mathbf{A} + j\mathbf{B}$ 和 $\mathbf{N} = \mathbf{C} + j\mathbf{D}$ 的乘积 $\mathbf{P} = \mathbf{X} + j\mathbf{Y}$ 可以通过以下实矩阵之间的乘法运算得到，即

$$\begin{bmatrix} \mathbf{X} & -\mathbf{Y} \\ -\mathbf{Y} & \mathbf{X} \end{bmatrix} = \begin{bmatrix} \mathbf{A} & -\mathbf{B} \\ -\mathbf{B} & \mathbf{A} \end{bmatrix} \begin{bmatrix} \mathbf{C} & -\mathbf{D} \\ -\mathbf{D} & \mathbf{C} \end{bmatrix}$$

埃尔米特矩阵 $\mathbf{P} = \mathbf{X} + j\mathbf{Y}$ 是正定的当且仅当 $\begin{bmatrix} \mathbf{X} & -\mathbf{Y} \\ -\mathbf{Y} & \mathbf{X} \end{bmatrix} > \mathbf{0}$ 。

例 2.1.1 求复值仿射矩阵函数 $\mathbf{M}(\mathbf{x})$ 的最大奇异值。

解 这个问题可以通过求解以下的优化问题得到：

$$\begin{aligned} & \min_{\mathbf{x}, \gamma} \gamma \\ & \text{s.t.} \quad \begin{bmatrix} -\gamma \mathbf{I} & \mathbf{M}^H(\mathbf{x}) \\ \mathbf{M}(\mathbf{x}) & -\gamma \mathbf{I} \end{bmatrix} \leq \mathbf{0} \end{aligned}$$

显然这是一个具有复线性矩阵不等式约束的优化问题。应用前面介绍的将复线性矩阵不等式转化成实线性矩阵不等式的方法，并记 $\mathbf{M}_R = \text{Re}(\mathbf{M})$, $\mathbf{M}_I = \text{Im}(\mathbf{M})$ ，可得等价的具有实线性矩阵不等式约束的优化问题：

$$\begin{aligned} & \min_{\mathbf{x}, \gamma} \gamma \\ & \text{s.t.} \quad \begin{bmatrix} -\gamma \mathbf{I} & \mathbf{0} & \mathbf{M}_R^T(\mathbf{x}) & -\mathbf{M}_I^T(\mathbf{x}) \\ \mathbf{0} & -\gamma \mathbf{I} & \mathbf{M}_I^T(\mathbf{x}) & \mathbf{M}_R^T(\mathbf{x}) \\ \mathbf{M}_R(\mathbf{x}) & \mathbf{M}_I(\mathbf{x}) & -\gamma \mathbf{I} & \mathbf{0} \\ -\mathbf{M}_I(\mathbf{x}) & \mathbf{M}_R(\mathbf{x}) & \mathbf{0} & -\gamma \mathbf{I} \end{bmatrix} \leq \mathbf{0} \end{aligned}$$

该问题可以应用 LMI 工具箱中标准的线性矩阵不等式求解器直接求解。