SELECTED PAPERS OF PROFESSOR ZHONGJUN ZHANG

ENGLISH EDITION

张钟俊教授论文集

(英文版)

上海交通大学出版社

序

2005年是张钟俊教授90寿辰，也是他逝世十周年纪念，上海交通大学决定出版张钟俊教授生前编辑的《张钟俊教授论文集》(英文版)第二卷。钟俊先生的学生韩正之教授，出版社总编辑，请我为此集的出版作序，欣然命笔。

张钟俊教授是我国自动化的先驱，他早年获中美文化教育基金会的奖学金，由交通大学选送赴美留学，在美著名学府麻省理工学院学习，获工学硕士和科学博士学位。在麻省理工学院学习期间，曾听过控制论的创始人罗勃特·维纳的课程。张钟俊教授于1938年学成回国。当时交通大学西迁重庆，24岁的张钟俊加盟交通大学，与朱物华等在重庆九龙坡组建电机系，任系主任。1948年，张钟俊在交通大学开设“伺服机构原理”课程，这是在中国大地第一次开设自动控制理论的教程，钟俊先生也因之而被誉为中国自动控制之父。建国前后，他曾在上海市公用事业局担任总工程师，在建国后的多项建设中有突出的贡献，周恩来总理还指示要他起草中国电气化十五年规划。鉴于他在自动化领域的开创性工作，1981年张钟俊当选为中国科学院学部委员。

钟俊先生是我国著名的自动化领域教育家。建国前，他培养的硕士在该领域几乎在国内占到三分之一，我国自动化、电气工程、计算机领域的很多专家学者都出自钟俊先生门下。文革之后，他专心治学，又培养了一大批硕士和博士，在自动化的各个领域都有独特的贡献。先生在世曾亲自编辑了《张钟俊教授论文集》中文三卷，英文一卷。这卷英文版系他在世时动手，而没能竟稿的遗集，最近由韩正之、席裕庚、施颂椒、邵惠鹤等四位教授重新整理而成，文集集有在国内外期刊会议上发表的论文24篇，内容涉及鲁棒控制、非线性控制、预测控制、智能控制、大系统控制、机器人、故障诊断、平行算法等多个领域，其中不少文章到现在看来还是很有价值的。

钟俊先生长我十三岁，我与先生亦师亦友，我于1953年从哈工大研究生班毕业留校，从事电力系统的教学与研究，1958年从苏联获副博士学位回国，当时主攻电力系统技术与理论，与钟俊先生同行，以后又一起转向自动控制研究，又在同一个领域工作。长期的同行给了我很多向他学习的机会。当时，我时常去上海先生家拜访，聆听先生讲述研究心得和历史掌故。钟俊先生知识渊博、学识过人，见识精辟。每次与他谈话都得益匪浅，他的很多教导对我很有启发，极有裨益。

今天，上海交通大学出版社出版《张钟俊教授论文集》英文版第二卷，这不仅是对先生的纪念，也是在弘扬先生严谨治学、勤学不辍的作风。这是一件非常有意义的事。我们应该以张钟俊教授为榜样，爱国、爱民、爱科学，为国家的发展和民族的振兴做出自己的贡献。

冯纯伯

2005年2月

Prof.Zhongjun Zhang (Member of CAS)
(1915·9 ~ 1995·12)

CONTENTS

I

ROBUST CONTROL OF LINEAR SYSTEMS

ROBUST DESIGN FOR SYSTEMS WITH ORSERVERS

Abstract

Systems with structured uncertainties in the state matrices, the input matrices and the output matrices are considered in this paper. An observer based feedback control law is derived to make the closed-loop system robust under the constraint of pole assignment.

1 Introduction

An observer is usually constructed when the states of a system are not all available. If the system is exactly described, the difference between the states of the system and those of the observer will vanish at the infinite time. But for an uncertain system, this will usually not be true. Stability robustness of the observer has been considered in many papers and robustness design for the closed-loop system has been discussed by Petersen, where a design method was given by solving two Riccati equations. The uncertainty considered is parameter dependent.

Consider the following system with structured uncertainties:

$$\left.\begin{aligned} \dot{x}(t) &= (A+G_1\Delta_1 H_1)x(t)+(B+G_2\Delta_2 H_2)u(t) \\ y(t) &= (C+G_3\Delta_3 H_3)x(t) \end{aligned}\right\} \tag{1}$$

where $x(t)\in \mathbf{R}^n$ is the state, $u(t)\in \mathbf{R}^m$ is the control and $y(t)\in \mathbf{R}^p$ is the measuarable output. A, B, and C are the nominal state matrix, input matrix and output matrix, respectively, with the appropriate dimensions. $G_i, H_i, i=1,2,3$ are known matrices defining the structure of the uncertainties. Δ_1, Δ_2 and Δ_3 are uncertain matrices bounded by the norm:

$$\bar{\sigma}[\Delta_i] \leqslant 1, \quad i=1,2,3$$

Denote $\Delta A=G_1\Delta_1 H_1, \Delta B=G_2\Delta_2 H_2, \Delta C=C_3\Delta_3 H_3$. Roubust design methods using state feedback have been developed for the case when $C=I$, $\Delta B=0$. In the paper by Hinrichsen, no constraint is made on the eigenvalues of $A+BF$ except for stability. While in the paper by Kautsky, the eigenvalues of $A+BF$ are prescribed.

In this paper, the general case will be studied. Since C is not necessarily to be identity, a static or dynamic output feedback should be applied. Here, The problem of designing an observer based feedback control law to make the system roubust under the constraint of pole assignment will be discussed.

2 Roubust Analysis and the Performance Index

2.1 The closed-loop system

Assumptions: (A,B) is controllable and (C,A) is observable.

Consider the system (1) given in Section 1. The following observer is constructed to approximate the state:

$$\dot{z} = Az + L(Cz - y) + Bu \tag{2}$$

and the feedback $u=Fz$ is applied. Then, the closed-loop system becomes:

$$\left.\begin{aligned} \dot{x} &= (A+\Delta A)x + (B+\Delta B)Fz \\ \dot{z} &= -L(C+\Delta C)x + (A+LC+BF)z \end{aligned}\right\} \tag{3}$$

Let $e=z-x$. After state transformation, we have:

$$\begin{bmatrix} \dot{x} \\ \dot{e} \end{bmatrix} = \begin{bmatrix} A+BF+\Delta A+\Delta BF & BF+\Delta BF \\ -\Delta A-\Delta BF-L\Delta C & A+LC-\Delta BF \end{bmatrix} \begin{bmatrix} x \\ e \end{bmatrix} \tag{4}$$

Denote $\Omega = \begin{bmatrix} A+BF & BF \\ 0 & A+LC \end{bmatrix}$ and $\Delta\Omega = \begin{bmatrix} \Delta A+\Delta BF & \Delta BF \\ -\Delta A-\Delta BF-L\Delta C & -\Delta BF \end{bmatrix}$, then Ω is the nominal closed-loop state matrix and $\Delta\Omega$ is its uncertain part.

2.2 Sufficient condition for robust stability

Noting that $\Delta A=G_1\Delta_1 H_1$, $\Delta B=G_2\Delta_2 H_2$, and $\Delta C=G_3\Delta_3 H_3$, $\Delta\Omega$ may be rewriten as:

$$\Delta\Omega = \begin{bmatrix} G_1 & G_2 & 0 \\ -G_1 & -G_2 & -LG_3 \end{bmatrix} \begin{bmatrix} \Delta_1 & 0 & 0 \\ 0 & \Delta_2 & 0 \\ 0 & 0 & \Delta_3 \end{bmatrix} \begin{bmatrix} H_1 & 0 \\ H_2F & H_2F \\ H_3 & 0 \end{bmatrix}$$

The following theorem is needed for robust stability.

Theorem 1 Let

$$T(s) = \begin{bmatrix} H_1 & 0 \\ H_2F & H_2F \\ H_3 & 0 \end{bmatrix} \cdot \left(sI - \begin{bmatrix} A+BF & BF \\ 0 & A+LC \end{bmatrix} \right)^{-1} \cdot \begin{bmatrix} G_1 & G_2 & 0 \\ -G_1 & -G_2 & LG_3 \end{bmatrix}$$

Denote $G = \begin{bmatrix} G_1 & G_2 & 0 \\ -G_1 & -G_2 & -LG_3 \end{bmatrix}$, $H = \begin{bmatrix} H_1 & 0 \\ H_2F & H_2F \\ H_3 & 0 \end{bmatrix}$, then $T(s)=H(sI-\Omega)^{-1}G$.

Suppose that Ω is stable, then the closed-loop system (3) is robustly stable if:

$$\| T(s) \|_\infty < 1 \tag{5}$$

2.3 Pole assignment constraint

Given a set of self-conjugate complex numbers, $\lambda_1, \lambda_2, \cdots, \lambda_n$, and $\delta_1, \delta_2, \cdots, \delta_n$, all with negative real part. It is desired that the eigenvalues of $A+BF$ and $A+LC$ to be $\lambda_1, \lambda_2, \cdots, \lambda_n$, and $\delta_1, \delta_2, \cdots, \delta_n$, respectively. Suppose that $\lambda_1, \lambda_2, \cdots, \lambda_n$, and $\delta_1, \delta_2, \cdots, \delta_n$, are different from one another, then, the pole assignment constraint may be equivalently stated as follows:

$$V_1^{-1}(A+BF)V_1 = \mathrm{diag}[\lambda_1, \cdots, \lambda_n] := \Lambda_1 \tag{6}$$

$$T_2(A+LC)T_2^{-1} = \mathrm{diag}[\delta_1, \cdots, \delta_n] := \Lambda_2 \tag{7}$$

So, robust design under the constraint of pole assignment may be formulated as:

$$\left.\begin{aligned} &\min \ \| T(s) \|_\infty \\ &\text{s. t. } V_1^{-1}(A+BF)V_1 = \Lambda_1 \\ &\quad\ T_2(A+LC)T_2^{-1} = \Lambda_2 \end{aligned}\right\} \tag{8}$$

2.4 Performance index related to robustness

The optimization problem (8) is difficult to solve. In the following, a performance index which is closely related to $\| T(s) \|_\infty$ will be derived to obtam a sub-optimal solution of (8).

Suppose that the constraints (6) and (7) are both satisfied. Let Y be the solution of

$$A_1Y - YA_2 = V_1^{-1}BFT_2^{-1} \tag{9}$$

then

$$\begin{bmatrix} V_1^{-1} & YT_2 \\ 0 & T_2 \end{bmatrix}\begin{bmatrix} V_1 & -V_1Y \\ 0 & T_2^{-1} \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}\begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix}$$

$$= \begin{bmatrix} V_1^{-1} & YT_2 \\ 0 & T_2 \end{bmatrix}\begin{bmatrix} A+BF & BF \\ 0 & A+LC \end{bmatrix}\begin{bmatrix} V_1 & -V_1Y \\ 0 & T_2^{-1} \end{bmatrix}$$

$$\Omega = \begin{bmatrix} A+BF & BF \\ 0 & A+LC \end{bmatrix} = \begin{bmatrix} V_1 & -V_1Y \\ 0 & T_2^{-1} \end{bmatrix}\begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix}\begin{bmatrix} V_1^{-1} & YT_2 \\ 0 & T_2 \end{bmatrix}$$

$$T(s) = H(sI-\Omega)^{-1}G = H \cdot \begin{bmatrix} V_1 & -V_1Y \\ 0 & T_2^{-1} \end{bmatrix}\left(sI - \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix}\right)^{-1}\begin{bmatrix} V_1^{-1} & YT_2 \\ 0 & T_2 \end{bmatrix}G$$

Denote $\begin{bmatrix} V_1 & -V_1Y \\ 0 & T_2^{-1} \end{bmatrix} = V$, $\begin{bmatrix} V_1^{-1} & YT_2 \\ 0 & T_2 \end{bmatrix} = T$, then $TV = I$,

$$T(s) = HV\left(sI - \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix}\right)^{-1} TG \tag{10}$$

so,

$$\| T(s) \|_\infty \leqslant \sigma[HV]\,\sigma[TG]\,\sup_\omega \sigma\left[\left(j\omega I - \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix}\right)^{-1}\right]$$

$$= \sigma[HV]\,\sigma[TG]\,\max_i\left[\frac{1}{|\mathrm{Re}\lambda_i|}, \frac{1}{|\mathrm{Re}\delta_i|}\right] \tag{11}$$

In the above steps, a similarity transformation is applied to separate ω from F, V_1, L, T_2 the parameters to be determined. Notice that H, V, T, and G depend on F, V_1, L and T_2.

The inequality (11) shows that the minimization of $\sigma[HV]\,\sigma[TG]$ will tend to minimize $\| T(s) \|_\infty$.

Define $\| \cdot \|$ as the spectral norm, or the singular value of a matrix:

$$\| Q \| = \sigma[Q] = \lambda_{\max}^{\frac{1}{2}}[Q'Q]$$

Choose the performance index J as

$$J = \frac{1}{2}\| HV \|^2 \| TG \|^2$$

Replace H, V, T and G by

$$\begin{bmatrix} H_1 & 0 \\ H_2F & H_2F \\ H_3 & 0 \end{bmatrix}, \begin{bmatrix} V_1 & -V_1Y \\ 0 & T_2 \end{bmatrix}, \begin{bmatrix} V_1^{1} & YT_2 \\ 0 & T_2 \end{bmatrix} \text{ and } \begin{bmatrix} G_1 & G_2 & 0 \\ -G_1 & -G_2 & LG_3 \end{bmatrix}$$

respectively, it follows

$$J = \frac{1}{2}\left\| \begin{matrix} H_1V_1 & -H_1V_1Y \\ H_2FV_1 & -H_2FV_1Y + H_2FT_2^{-1} \\ H_3V_1 & -H_3V_1Y \end{matrix} \right\|^2 \cdot \left\| \begin{matrix} V_1^{-1}G_1 - YT_2G_1 & V_1^{-1}G_2 - YT_2G_2 & -YT_2LG_3 \\ -T_2G_1 & -T_2G_2 & -T_2LG_3 \end{matrix} \right\|^2 \tag{12}$$

Now consider the following optimization problem:

$$\left.\begin{aligned} &\min J \\ \text{s. t. } &V_1^{-1}(A+BF)V_1 = A_1 \\ &T_2(A+LC)T_2^{-1} = A_2 \\ &A_1Y - YA_2 = V_1^{-1}BFT_2^{-1} \end{aligned}\right\} \tag{13}$$

The above three constraints may be relaxed by introducing two parameters $U\in \mathbf{R}^{m\times n}$ and $W\in \mathbf{R}^{n\times p}$ and defining two functions in the following Section 3.

2.5 Real diagonal form of A_1 and A_2

Suppose that the desired eigenvalues of $A+BF$ are $\lambda_1,\cdots,\lambda_n$, where λ_i,λ_{i+1}, are complex conjugate pairs, λ_i, $\lambda_{i+1}=\alpha_i+\mathrm{j}\beta_i$. Define

$$A_1 = \operatorname{diag}\left[\lambda_1,\cdots,\begin{vmatrix}\alpha_i & \beta_i \\ -\beta_i & \alpha_i\end{vmatrix},\cdots,\lambda_n\right] \tag{14}$$

Similarly, if δ_i, δ_{i+1} are complex pairs, $\delta_i,\delta_{i+1}=\varepsilon_i\pm \mathrm{j}\gamma_i$, also define

$$A_2 = \operatorname{diag}\left[\delta_1,\cdots,\begin{vmatrix}\varepsilon_i & \gamma_i \\ -\gamma_i & \varepsilon_i\end{vmatrix},\cdots,\delta_n\right] \tag{15}$$

When A_1 and A_2 are defined in this way, the complex consideration is avoided while the pole assignment constraint is kept unchanged.

3 Relaxation of the Constraints

The constraints in (13) may be relaxed by defining function $f_1: U\in \mathbf{R}^{m\times n}\to(F,V_1)\in \mathbf{R}^{m\times n}\times\mathbf{R}^{n\times n}$ and function $f_2: W\in\mathbf{R}^{n\times p}\to(T_2,L)\in\mathbf{R}^{n\times n}\times\mathbf{R}^{n\times p}$.

Definition 1 Suppose that (A,B) is controllable, A_1 is a real diagonal matix as defined in (14). Assume that A and A_1 have no common eigenvalue, then the function f_1 may be defined as follows:

Let $U\in\mathbf{R}^{m\times n}$, solve

$$AV_1 - V_1A_1 = -BU \tag{16}$$

for V_1. If V_1 is nonsingular, define

$$E = UV_1^{-1}$$

Then (F,V_1) is the image of U under f_1, that is, $(F,V_1)=f_1(U)$. The domain of f_1, denoted by D_{f_1}, is defined as all the U that make V_1 nonsingular, and the range of f_1, denoted by R_{f_1}, is the image of D_{f_1} under f_1.

Theorem 2 The range of f_1 is the set of all the pairs (F,V_1) that satisfy $V_1^{-1}(A+BF)\cdot V_1=A_1$. And if the domain of f_1 is not empty, it must be a dense and open set in $\mathbf{R}^{m\times n}$.

Proof See our previous paper on Proceeding of MTNS, Vol. 2., Birkhaüser, 1990.

Those for f_2 may be defined in a dual manner:

Definition 2 Suppose that (C,A) is observable, A_2 is a real diagonal matrix as defined in (15). Assume that A and A_2 have no common eigenvalue, then f_2 is defined as follows:

Let $W\in\mathbf{R}^{n\times p}$, solve

$$T_2A - A_2T_2 = -WC \tag{17}$$

If T_2 is nonsingular, define

$$L = T_2^{-1}W$$

Then (T_2, L) is the image of W under f_2, that is, $(T_2, L) = f_2(W)$. The domain of f_2, denoted by D_{f_2}, is defined as all the W that make T_2 nonsingular, and the range of f_2, denoted by R_{f_2}, is defined as the image of D_{f_2} under f_2.

Theorem 3 The range of f_2 is the set of all the pairs (T_2, L) that satisfy $T_2(A+LC) \cdot T_2^{-1} = A_2$. If the domain of f_2 is not empty, it must be a dense and open set in $\mathbf{R}^{n\times p}$.

Using the above results, the first two constraints of (13) may be replaced by

$$(F, V_1) = f_1(U), \quad U \in D_{f_1}$$

and

$$(T_2, L) = f_2(W), \quad W \in D_{f_2}$$

Since A_1 and A_2 have no common eigenvalue, Y is uniquely determined by V_1, F, and T_2. So, Y is a function of U and W. Thus J is a functional of U and W.

A method to compute $\frac{\partial J}{\partial U}$ and $\frac{\partial J}{\partial W}$ will be given in the next section. Since D_{f_1} and D_{f_2} are dense and open set in $\mathbf{R}^{m\times n}$ and $\mathbf{R}^{n\times p}$, J may be minimized by the gradient method.

4 The Gradient Method

From the previous section, the optimization problem (13) may be formulated as:

$$\min\left\{ J = \frac{1}{2}\left\| \begin{matrix} V_1^{-1}G_1 - YT_2G_1 & V_1^{-1}G_2 - YT_2G_2 & -YT_2LG_3 \\ -T_2G_1 & -T_2G_2 & -T_2LG_3 \end{matrix} \right\|^2 \cdot \left\| \begin{matrix} H_1V_1 & -H_1V_1Y \\ H_2FV_1 & -H_2FV_1Y + H_2FT_2^{-1} \\ H_3V_1 & -H_3V_1Y \end{matrix} \right\|^2 \right\}$$

where

$$(F, V_1) = f_1(U): AV_1 - V_1A_1 = -BU, \quad F = UV_1^{-1}, U \in D_{f_1}$$

$$(T_2, L) = f_2(W): T_2A - A_2T_2 = -WC, \quad L = T_2^{-1}W, W \in D_{f2}$$

and

$$A_1Y - YA_2 = V_1^{-1}BFT_2^{-1} \tag{18}$$

V_1 and F are determined by U, T_2 and L are determined by W, and Y is jointly determined by U and W. Therefore J is a fuctional of U and W, and U and W are free parameters.

Let

$$J_1 = \frac{1}{2}\left\| \begin{matrix} V_1^{-1}G_1 - YT_2G_1 & V_1^{-1}G_2 - YT_2G_2 & -YT_2LG_3 \\ -T_2G_1 & -T_2G_2 & -T_2LG_3 \end{matrix} \right\|^2 \tag{19}$$

$$J_2 = \frac{1}{2}\left\| \begin{matrix} H_1V_1 & -H_1V_1Y \\ H_2FV_1 & -H_2FV_1Y + H_2FT_2^{-1} \\ H_3V_1 & -H_3V_1Y \end{matrix} \right\|^2 \tag{20}$$

then $J = 2J_1J_2$,

$$\frac{\partial J}{\partial U} = 2J_1\left(\frac{\partial J_2}{\partial U}\right) + 2J_2\left(\frac{\partial J_1}{\partial U}\right)$$

$$\frac{\partial J}{\partial W}=2J_1\left(\frac{\partial J_2}{\partial W}\right)+2J_2\left(\frac{\partial J_1}{\partial W}\right)$$

(1) Firstly, find $\frac{\partial J_1}{\partial U}$ and $\frac{\partial J_1}{\partial W}$.

Let

$$P_1=V_1^{-1}G_1-YT_2G_1,\quad P_2=V_1^{-1}G_2-YT_2G_2$$
$$P_3=YT_2LG_3,\quad P_4=T_2G_1$$
$$P_5=T_2G_2,\quad P_6=T_2LG_3$$

Denote $P=\begin{bmatrix} P_1 & P_2 & -P_3 \\ -P_4 & -P_5 & -P_6 \end{bmatrix}$, then

$$J_1=\frac{1}{2}\parallel P\parallel^2=\frac{1}{2}\lambda_{\max}[P' \quad P] \tag{21}$$

Let w be the right eigenvetor of $P'P$ corresponding to $\lambda_{\max}[P'P]$ and $w'w=1$. Part $ww'P'$ as,

$$\begin{bmatrix} S'_1 & -S'_4 \\ S'_2 & -S'_5 \\ -S'_3 & -S'_6 \end{bmatrix}$$

where S_i has the same size with P_i. It is derived in the appendix that

$$\frac{\partial J_1}{\partial U}=(ZB+X_2B+V_1^{-1}T_2^{-1}X_1V_1^{-1}B)' \tag{22}$$

where Z, X_1 and X_2 satisfy

$$ZA-A_1Z=-Q_1,\quad X_1A_1-A_2X_1=Q_2$$
$$X_2A-A_1X_2=V_1^{-1}BFT_2^{-1}X_1V_1^{-1}+V_1^{-1}T_2^{-1}X_1V_1^{-1}BF$$

and

$$Q_1=-V_1^{-1}(G_1S'_1+G_2S'_2)V_1^{-1}$$
$$Q_2=-T_2G_1S'_1-T_2G_2S'_2+WG_3S'_3$$

$$\frac{\partial J_1}{\partial W}=(Q_3+CZ+CX_2)'$$

where Z, X_1 and X_2 satisfy

$$AZ-ZA_2=-Q_4$$
$$X_1A_1-A_2X_1=Q_5(=Q_2)$$
$$AX_2-X_2A_2=T_2^{-1}X_1V_1^{-1}BFT_2^{-1}$$

and

$$Q_3=G_3S'_3Y+G_3S'$$
$$Q_4=-G_1S'_1Y-G_2S'_2Y+G_1S'_4+G_2S'_5$$
$$Q_5=Q_2$$

(2) Secondly, find $\frac{\partial J_2}{\partial U}$ and $\frac{\partial J_2}{\partial W}$.

Let

$$N_1=H_1V_1, N_2=H_2FV_1,\quad N_3=H_3V_1, N_4=H_1V_1Y$$
$$N_5=H_2FV_1Y-H_2FT_2^{-1},\quad N_6=H_3V_1Y$$

Denote $N=\begin{bmatrix} N_1 & -N_4 \\ N_2 & -N_5 \\ N_3 & -N_6 \end{bmatrix}$, then $J_2=\frac{1}{2}\parallel N\parallel^2=\frac{1}{2}\lambda_{\max}[N'N]$.

Let v be the right eigenvector of $N'N$ corresponding to $\lambda_{\max}[N'N]$ and $v'v=1$. Partition $vv'N'$ as $\begin{bmatrix} R'_1 & R'_2 & R'_3 \\ -R'_4 & -R'_5 & -R'_6 \end{bmatrix}$, where R_i has the same size with N_i, hence

$$\frac{\partial J_2}{\partial U} = ZB + Q_6 + X_2B + V_1^{-1}T_2^{-1}X_1V_1^{-1}B \tag{23}$$

where Z, X_1 and X_2 satisfy

$$ZA - A_1Z = -Q_7, \quad X_1A_1 - A_2X_1 = Q_8$$
$$X_2A - A_1X_2 = V_1^{-1}BFT_2^{-1}X_1V_1^{-1} + V_1^{-1}T_2^{-1}X_1V_1^{-1}BF$$

and

$$Q_6 = R'_2H_2 + YR'_5H_2 - V_1^{-1}T_2^{-1}R'_5H_2$$
$$Q_7 = R'_1H_1 + R'_8H_8 + YR'_4H_1 + V_1^{-1}T_2^{-1}R'_5H_2F + YR'_6H_3$$
$$Q_8 = R'_4H_1V_1 + R'_5H_2U + R'_6H_8V_1$$

$$\frac{\partial J_2}{\partial W} = (CZ + CX_2)' \tag{24}$$

where Z, X_1 and X_2 satisfy

$$AZ - ZA_2 = -Q_8$$
$$X_1A_1 - A_2X_1 = Q_{10} (= Q_8)$$
$$AX_2 - X_2A_2 = T_2^{-1}X_1V_1^{-1}BFT_2^{-1}$$

and

$$Q_9 = T_2^{-1}R'_5H_2FT_2^{-1}$$
$$Q_{10} = Q_8$$

(3) Combining $\frac{\partial J_1}{\partial U}$ and $\frac{\partial J_2}{\partial U}$ and $\frac{\partial J_1}{\partial W}$ and $\frac{\partial J_2}{\partial W}$. One will be found

$$\frac{\partial J}{\partial U} = 2(V_1^{-1}T_2^{-1}X_1V_1^{-1}B + J_1Q_6 + ZB + X_2B) \tag{25}$$

where Z, X_1 and X_2 satisfy

$$\frac{\partial J}{\partial W} = 2(J_2Q_8 + CZ + CX_2)' \tag{26}$$

where Z, X_1 and X_2 satisfy

$$AZ - ZA_2 = -J_2Q_4 - J_1Q_9$$
$$X_1A_1 - A_2X_1 = J_2Q_5 + J_1Q_{10} = J_2Q_2 + J_1Q_8$$
$$AX_2 - X_2A_2 = T_2^{-1}X_1V_1^{-1}BFT_2^{-1}$$

Thus the calculation of $\frac{\partial J}{\partial U}$ and $\frac{\partial J}{\partial W}$ requires solving five Sylvester equations. Since A_1 and A_2 are diagonal matrices, each Sylvester equation is equivalent to an n-ordered linear algebraic equations.

The computation is to obtain $Q_1, Q_2, \cdots, Q_{10}$, which involves many matrix production and addition.

If J is not convex, it may have many local infimums. But the optimization is smooth, since for any $U \in D_{f_1}$ and $W \in D_{f_2}$, $\frac{\partial J}{\partial U}$ and $\frac{\partial J}{\partial W}$ are uniquely determined. So, the gradient method will guarantee J to decrease until a local infimum is reached. For the purpose of robust design, several infimums should be found and choose the smallest among them. To ascertain robust stability of the closed-loop system, the condition $\| T(s) \|_\infty < 1$ should be checked. If this condition does not satisfied, a search for smaller infimums of J should be made or adjust the prescribed eigenvalues, λ_i and δ_i.

5 Concluding Remarks

The main contribution of this paper includes obtaining performance indexes related to robustness, relaxing the pole assignment constraint and deriving formulas to calculate the gradients by using some properties of Kroneker product.

Similar method may be applied for solving robustness design problems, with the constraint of pole assignment. For example, for state feedback control, J may be chosen as, $J=\frac{1}{2}\|A+BF\|^2, \frac{1}{2}\|HV\|^2$, $\|VG^{-1}\|^2, \frac{1}{2}\|V^{-1}(A_a+B_a)V\|^2$, or $\|P\|^2$, etc, where V is the eigenvector matrix of $A+BF$ and P is the solution to $(A+BF)'P+P(A+BF)=-Q$.

If there is no constraint on the eigenvalues of $A+BF$ and $A+LC$, except for stability. J may also be optimized by gradient method, since the computation of $\frac{\partial J}{\partial \lambda_i}$ is not a difficult task.

6 Appendix—Derivation of Formulas to Calculate the Gradients

The derivation of the gradients is based on some properties of Kroneker product. A preliminary knowledge of knonecker product etc is given at first.

Preliminary knowledge:

Definition 3 Let $X \in \mathbf{R}^{m\times n}, Y \in \mathbf{R}^{p\times q}$, the Kroneker product of X and $Y, X\otimes Y$ is defined as:

$$X \otimes Y := \begin{bmatrix} x_{11}Y & x_{12}Y & \cdots & x_{1n}Y \\ x_{21}Y & x_{22}Y & & \vdots \\ \vdots & \vdots & & \vdots \\ x_{m1}Y & \cdots & \cdots & x_{mn}Y \end{bmatrix} \in \mathbf{R}^{mp\times nq}$$

Definition 4 Let $X \in \mathbf{R}^{m\times n}$, the row vector form of X, rs$X$, is defined as:

$$\mathrm{rs}X := [x_{11}, x_{12}, \cdots, x_{1n}, x_{21}, x_{22}, \cdots, x_{2n}, \cdots, x_{mn}]$$

and the collum vector form of X, cs$X$, is defined as:

$$\mathrm{cs}X := [x_{11}, x_{21}, \cdots, x_{m1}, x_{12}, x_{22}, \cdots, x_{m2}, \cdots, x_{mn}]'$$

Definition 5 Let $X \in \mathbf{R}^{n\times n}$, then the trace of X, tr$X$, is defined as:

$$\mathrm{tr}X := \sum x_{ii}$$

Some properties of $\otimes$, cs, rs and tr are given as follows:

Fact 1 $\mathrm{tr}XY=\mathrm{tr}YX$

Fact 2 $\mathrm{tr}XY=\mathrm{rs}X \cdot \mathrm{cs}Y=\mathrm{rs}Y \cdot \mathrm{cs}X$

Fact 3 $\mathrm{cs}(XYZ)=(Z'\otimes X) \cdot \mathrm{cs}Y$

$\mathrm{rs}(XYZ)=\mathrm{rs}Y(X'\otimes Z)$

Fact 4 For $X \in \mathbf{R}^{m\times n}$ and $Y \in \mathbf{R}^{n\times p}$,

$\mathrm{cs}(XY)=(I_p\otimes X) \cdot \mathrm{cs}Y=(Y'\otimes I_m) \cdot \mathrm{cs}X$

$\mathrm{rs}(XY)=\mathrm{rs}Y \cdot (X'\otimes I_p)=\mathrm{rs}X \cdot (I_m\otimes Y)$

Proofs for the above facts can be found in Graham's book. The following theorem is a direct consequence from Fact 4.

Theorem 4 For $A \in \mathbf{R}^{n\times n}$, $B \in \mathbf{R}^{n\times m}$ and $C \in \mathbf{R}^{p\times n}$,

(1) $AV_1 - V_1A = -BU$ is eqivalent to:

$$(I_n \otimes A - A'_1 \otimes I_n) \cdot \mathrm{cs}V_1 = -(I_m \otimes B) \cdot \mathrm{cs}\, U$$

(2) $T_2A - A_2T_2 = -WC$ is equvalent to:

$$(A' \otimes I_n - I_n \otimes A_2) \cdot \mathrm{cs}\, T_2 = -(C' \otimes I_p) \cdot \mathrm{cs}W$$

(3) $A_1Y - YA_2 = -V_1^{-1}BFT_2^{-1}$ is equivalent to:

$$(I_n \otimes A_1 - A'_2 \otimes I_n) \cdot \mathrm{cs}Y = \mathrm{cs}(V_1^{-1}BFT_2^{-1})$$

Now to calculate $\frac{\partial J_1}{\partial U}$:

$$J_1 = \frac{1}{2} \| P \|^2 = \frac{1}{2}\lambda_{\max}[P'P]$$

Let $\lambda = \lambda_{\max}[P'P]$, then $P'Pw = \lambda w$,

$$\left(\frac{\partial P'}{\partial u_{ij}}P + P'\frac{\partial P}{\partial u_{ij}}\right)w + P'P\frac{\partial w}{\partial u_{ij}} = \frac{\partial \lambda}{\partial u_{ij}}w + \lambda\frac{\partial w}{\partial u_{ij}}$$

Multiply both side with w', notice $w'P'P = \lambda w'$, $w'w = 1$, one gets

So $$\frac{\partial \lambda}{\partial u_{ij}} = w'\left(\frac{\partial P^j}{\partial u_{ij}} + P + P'\frac{\partial P}{\partial u_{ij}}\right)w = 2w'P'\frac{\partial P}{\partial u_{ij}}w$$

$$\frac{\partial J_1}{\partial u_{ij}} = w'P'\frac{\partial P}{\partial u_{ij}}w = \mathrm{tr}\left[ww'P'\frac{\partial P}{\partial u_{ij}}\right]$$

$$= \mathrm{tr}\begin{bmatrix} S'_1 & -S'_4 \\ S'_2 & -S'_5 \\ -S'_3 & -S'_6 \end{bmatrix}\begin{bmatrix} \frac{\partial P_1}{\partial u_{ij}} & \frac{\partial P_2}{\partial u_{ij}} & -\frac{\partial P_3}{\partial u_{ij}} \\ -\frac{\partial P_4}{\partial u_{ij}} & -\frac{\partial P_5}{\partial u_{ij}} & -\frac{\partial P_6}{\partial u_{ij}} \end{bmatrix}$$

$$= \mathrm{tr}\sum_{i=1}^{6} S'_i \frac{\partial P_i}{\partial u_{ij}}$$

$$= \mathrm{tr}\left[S'_1\left(\frac{\partial V_1^{-1}}{\partial u_{ij}}G_1 - \frac{\partial Y}{\partial u_{ij}}T_2G_1\right) + S'_2\left(\frac{\partial V_1^{-1}}{\partial u_{ij}}G_2 - \frac{\partial Y}{\partial u_{ij}}T_2G_2\right) + S'_3\frac{\partial Y}{\partial u_{ij}}WG_3\right]$$

$$= \left(\text{since } \frac{\partial T_2}{\partial u_{ij}} = 0, T_2L = W, \frac{\partial L}{\partial u_{ij}} = 0, \text{from (18)}\right)$$

$$= \mathrm{tr}\left[S'_1\left(-V_1^{-1}\frac{\partial V_1}{\partial u_{ij}}V_1^{-1}G_1 - \frac{\partial Y}{\partial u_{ij}}T_2G_1\right) + S'_2\left(-V_1^{-1}\frac{\partial V_1}{\partial u_{ij}}V_1^{-1}G_2 - \frac{\partial Y}{\partial u_{ij}}T_2G_2\right) + S'_3\frac{\partial Y}{\partial u_{ij}}WG_3\right]$$

$$= \left(\frac{\partial V_1^{-1}}{\partial u_{ij}} = -V_1^{-1}\frac{\partial V_1}{\partial u_{ij}}V_1^{-1}\right)$$

$$= \mathrm{tr}\left[-V_1^{-1}(G_1S'_1V_1^{-1} + G_2S'_2V_1^{-1})\frac{\partial V_1}{\partial u_{ij}} + (-T_2G_1S'_1 - T_2G_2S'_2 + WG_3S'_3)\frac{\partial Y}{\partial u_{ij}}\right]$$

(since $\mathrm{tr}MN = \mathrm{tr}NM$)

Let

$$Q_1 = -V_1^{-1}(G_1S'_1V_1^{-1} + G_2S'_2V_1^{-1})$$
$$Q_2 = -T_2G_1S'_1 - T_2G_2S'_2 + WG_3S'_3$$

Then

$$\frac{\partial J_1}{\partial u_{ij}} = \mathrm{tr}\left[Q_1 \frac{\partial V_1}{\partial u_{ij}} + Q_2 \frac{\partial Y}{\partial u_{ij}}\right]$$

$$= \mathrm{rs}Q_1 \frac{\partial \mathrm{cs}V_1}{\partial u_{ij}} + \mathrm{rs}Q_2 \frac{\partial \mathrm{cs}Y}{\partial u_{ij}} \text{(fact 2)}$$

$$= - \mathrm{rs}Q_1[(I_n \otimes A - A'_1 \otimes I_n)^{-1}(I_m \otimes B)] \frac{\partial \mathrm{cs}U}{\partial u_{ij}}$$

$$+ \mathrm{rs}Q_2[(I_n \otimes A_1 - A'_1 \otimes I_n)^{-1}] \frac{\partial \mathrm{cs}(V_1^{-1}BFT_2^{-1})}{\partial u_{ij}} \text{(from theorem 4)}$$

Let Z be the solution of $ZA - A_1Z = -Q_1$ and X_1 be the solution of $X_1A_1 - A_2X_1 = Q_2$, from Fact 4, it is easy to show that

$$\underbrace{-\mathrm{rs}Q_1[(I_n \otimes A - A_1 \otimes I_n)^{-1}]}_{\mathrm{rs}\,Z}(I_m \otimes B) = \mathrm{rs}(ZB)$$

and

$$\mathrm{rs}Q_2[(I_n \otimes A_1 - A'_2 \otimes I_n)^{-1}] = \mathrm{rs}X_1$$

So,

$$\frac{\partial J_1}{\partial u_{ij}} = \mathrm{rs}(ZB) \frac{\partial \mathrm{cs}U}{\partial u_{ij}} + \mathrm{rs}X_1 \frac{\partial \mathrm{cs}(V_1^{-1}BFT_2^{-1})}{\partial u_{ij}} \quad (F = UV_1^{-1})$$

$$\frac{\partial \mathrm{cs}(V_1^{-1}BFT_2^{-1})}{\partial u_{ij}} = \mathrm{cs}\left(-V_1^{-1} \frac{\partial V_1}{\partial u_{ij}} V_1^{-1}BFT_2^{-1} + V_1^{-1}B \frac{\partial U}{\partial u_{ij}} V_1^{-1}T_2^{-1} - V_1^{-1}BUV_1^{-1} \frac{\partial V_1}{\partial u_{ij}} V_1^{-1}T_2^{-1}\right)$$

$$= -[(V_1^{-1}BFT_2^{-1})' \otimes V_1^{-1} + (V_1^{-1}T_2^{-1})' \otimes (V_1^{-1}BF)] \frac{\partial \mathrm{cs}V_1}{\partial u_{ij}}$$

$$+ (V_1^{-1}T_2^{-1})' \otimes (V_1^{-1}B) \frac{\partial \mathrm{cs}U}{\partial u_{ij}}$$

$$(\mathrm{cs}(XYZ) = (Z' \otimes X) \cdot \mathrm{cs}Y)$$

$$\mathrm{rs}X_1 \frac{\partial \mathrm{cs}(V_1^{-1}BFT_2^{-1})}{\partial u_{ij}} = -\mathrm{rs}[V_1^{-1}BFT_2^{-1}X_1V_1^{-1} + V_1^{-1}T_2^{-1}X_1V_1^{-1}BF] \frac{\partial \mathrm{cs}V_1}{\partial u_{ij}}$$

$$+ \mathrm{rs}(V_1^{-1}T_2^{-1}X_1V_1^{-1}B) \frac{\partial \mathrm{cs}U}{\partial u_{ij}}$$

Since

$$\mathrm{rs}(XYZ) = \mathrm{rs}Y \cdot (X' \otimes Z))$$

$$= [\mathrm{rs}(V_1^{-1}BFT_2^{-1}X_1 V_1^{-1} + V_1^{-1}T_2^{-1}X_1V_1^{-1}BF)(I_n \otimes A - A'_1 \otimes I_n)^{-1}(I_m \otimes B)$$

$$+ \mathrm{rs}(V_1^{-1}T_2^{-1}X_1V_1^{-1}B)] \frac{\partial \mathrm{cs}U}{\partial u_{ij}}$$

Let X_2 be the solution of

$$X_2A - A_1X_2 = V_1^{-1}BFT_2^{-1}X_1V_1^{-1} + V_1^{-1}T_2^{-1}X_1V_1^{-1}BF$$

then $\mathrm{rs}(V_1^{-1}BFT_2^{-1}X_1V_1^{-1} + V_1^{-1}T_2^{-1}X_1V_1^{-1}BF)(I_n \otimes A - A'_1 \otimes I_n)^{-1}(I_m \otimes B) = \mathrm{rs}(X_2B)$

Hence

$$\frac{\partial J_1}{\partial u_{ij}} = \mathrm{rs}(ZB + X_2B + V_1^{-1}T_2^{-1}X_1V_1^{-1}B) \cdot \frac{\partial U}{\partial u_{ij}}$$

It follows

$$\frac{\partial J_1}{\partial U} = (ZB + X_2B + V_1^{-1}T_2^{-1}X_1V_1^{-1}B)'$$

where Z, X_1 and X_2 satisfy

$$ZA - A_1Z = -Q_1, \quad X_1A_1 - A_2X_1 = Q_2$$

$$X_2A - A_2X_2 = V_1^{-1}BFT_2^{-1}X_1V_1^{-1} + V_1^{-1}T_2^{-1}X_1V_1^{-1}BF$$

and

$$Q_1 = -V_1^{-1}(G_1S'_1V_1^{-1} + G_2S'_2V_1^{-1})$$