Chapter 1

Stable controller coefficient perturbation in floating point implementation

JUN WU
National Key Laboratory of Industrial Control Technology
Institute of Advanced Process Control
Zhejiang University
Hangzhou 310027
P. R. China
jwu@iipc.zju.edu.cn

SHENG CHEN
Department of Electronics and Computer Science
University of Southampton, Highfield
Southampton SO17 1BJ
United Kingdom
sqc@ecs.soton.ac.uk

1.1 Description of the problem

For real matrix $\mathbf{X} = [x_{ij}]$, denote

$$\|\mathbf{X}\|_{\max} = \max_{i,j} |x_{ij}|. \tag{1.1}$$

For real matrices $\mathbf{X} = [x_{ij}]$ and $\mathbf{Y} = [y_{ij}]$ of the same dimension, denote the Hadamard product of $\mathbf{X}$ and $\mathbf{Y}$ as

$$\mathbf{X} \circ \mathbf{Y} = [x_{ij} y_{ij}]. \tag{1.2}$$

1

A square real matrix is said to be stable if its eigenvalues are all in the interior of the unit disc.

Consider a stable discrete-time closed-loop control system, consisting of a linear time invariant plant $P(z)$ and a digital controller $C(z)$. The plant model $P(z)$ is assumed to be strictly proper with a state-space description

$$\begin{cases} \mathbf{x}_P(k+1) = \mathbf{A}_P\mathbf{x}_P(k) + \mathbf{B}_P\mathbf{u}(k) \\ \mathbf{y}(k) = \mathbf{C}_P\mathbf{x}_P(k) \end{cases} \tag{1.3}$$

where $\mathbf{A}_P \in \mathcal{R}^{m \times m}$, $\mathbf{B}_P \in \mathcal{R}^{m \times l}$ and $\mathbf{C}_P \in \mathcal{R}^{q \times m}$. The controller $C(z)$ is described by

$$\begin{cases} \mathbf{x}_C(k+1) = \mathbf{A}_C\mathbf{x}_C(k) + \mathbf{B}_C\mathbf{y}(k) \\ \mathbf{u}(k) = \mathbf{C}_C\mathbf{x}_C(k) + \mathbf{D}_C\mathbf{y}(k) \end{cases} \tag{1.4}$$

where $\mathbf{A}_C \in \mathcal{R}^{n \times n}$, $\mathbf{B}_C \in \mathcal{R}^{n \times q}$, $\mathbf{C}_C \in \mathcal{R}^{l \times n}$ and $\mathbf{D}_C \in \mathcal{R}^{l \times q}$. It can be shown easily that the transition matrix of the closed-loop system is

$$\mathbf{A} = \left[\begin{array}{cc} \mathbf{A}_P + \mathbf{B}_P\mathbf{D}_C\mathbf{C}_P & \mathbf{B}_P\mathbf{C}_C \\ \mathbf{B}_C\mathbf{C}_P & \mathbf{A}_C \end{array} \right] \in \mathcal{R}^{(m+n) \times (m+n)}. \tag{1.5}$$

It is well known that a discrete-time closed-loop system is stable if and only if its transition matrix is stable. Since the closed-loop system, consisting of (1.3) and (1.4), is designed to be stable, $\mathbf{A}$ is stable. Let

$$\mathbf{B} = \left[\begin{array}{cc} \mathbf{B}_P & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{array} \right] \in \mathcal{R}^{(m+n) \times (l+n)}, \tag{1.6}$$

$$\mathbf{C} = \left[\begin{array}{cc} \mathbf{C}_P & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{array} \right] \in \mathcal{R}^{(q+n) \times (m+n)}, \tag{1.7}$$

$$\mathbf{W} = \left[\begin{array}{cc} \mathbf{D}_C & \mathbf{C}_C \\ \mathbf{B}_C & \mathbf{A}_C \end{array} \right] \in \mathcal{R}^{(l+n) \times (q+n)}, \tag{1.8}$$

where $\mathbf{0}$ and $\mathbf{I}$ denote the zero and identity matrices of appropriate dimensions, respectively. Define the set

$$\mathcal{S} = \{\boldsymbol{\Delta} : \boldsymbol{\Delta} \in \mathcal{R}^{(l+n) \times (q+n)}, \mathbf{A} + \mathbf{B}(\mathbf{W} \circ \boldsymbol{\Delta})\mathbf{C} \text{ is stable}\} \tag{1.9}$$

and further define

$$\upsilon = \inf\{\|\boldsymbol{\Delta}\|_{\max} : \boldsymbol{\Delta} \in \mathcal{R}^{(l+n) \times (q+n)}, \boldsymbol{\Delta} \notin \mathcal{S}\}. \tag{1.10}$$

The open problem is: calculate the value of υ.

1.2   Motivation of the problem

The classical digital controller design methodology often assumes that the controller is implemented exactly, even though in reality a control law can only be realized with a digital processor of finite word length (FWL). It may seem

that the uncertainty resulting from finite-precision computing of the digital controller is so small, compared to the uncertainty within the plant, such that this controller "uncertainty" can simply be ignored. Increasingly, however, researchers have realized that this is not necessarily the case. Due to the FWL effect, a casual controller implementation may degrade the designed closed-loop performance or even destabilize the designed stable closed-loop system, if the controller implementation structure is not carefully chosen [1, 2].

With decreasing in price and increasing in availability, the use of floating-point processors in controller implementations has increased dramatically. When a real number x is implemented in a floating-point format, it is perturbed to $x(1 + \delta)$ with $|\delta| < \eta$, where η is the maximum round-off error of the floating-point representation [3]. It can be seen that the perturbation resulting from finite-precision floating-point arithmetic is multiplicative.

For the closed-loop system described in section 1.1, when $C(z)$ is implemented in finite-precision floating-point format, the controller realization $\mathbf{W}$ is perturbed to $\mathbf{W} + \mathbf{W} \circ \boldsymbol{\Delta}$. Each element of $\boldsymbol{\Delta}$ is bounded by $\pm\eta$, that is,

$$\|\boldsymbol{\Delta}\|_{\max} < \eta. \tag{1.11}$$

With the perturbation $\boldsymbol{\Delta}$, the transition matrix of the closed-loop system becomes $\mathbf{A} + \mathbf{B}(\mathbf{W} \circ \boldsymbol{\Delta})\mathbf{C}$. If an eigenvalue of $\mathbf{A} + \mathbf{B}(\mathbf{W} \circ \boldsymbol{\Delta})\mathbf{C}$ is outside the open unit disc, the closed-loop system, designed to be stable, becomes unstable with the FWL floating-point implemented $\mathbf{W}$.

It is therefore critical to know the ability of the closed-loop stability to tolerate the coefficient perturbation $\boldsymbol{\Delta}$ in $\mathbf{W}$ resulted from finite-precision implementation. This means that we would like to know the largest "cube" in the perturbation space, within which the closed-loop system remains stable. The measure v defined in (1.10) gives the exact size of the largest "stable perturbation cube" for $\mathbf{W}$. If the value of v can be computed, it becomes a simple matter to check whether $\mathbf{W}$ is "robust" to FWL errors, because $\mathbf{A} + \mathbf{B}(\mathbf{W} \circ \boldsymbol{\Delta})\mathbf{C}$ remains stable when $v > \eta$.

Furthermore, $\mathbf{W}$ or $(\mathbf{A}_C, \mathbf{B}_C, \mathbf{C}_C, \mathbf{D}_C)$ is a realization of the controller $C(z)$. The realizations of $C(z)$ are not unique. Different realizations are all equivalent if they are implemented in infinite precision. In fact, if $(\mathbf{A}_C^0, \mathbf{B}_C^0, \mathbf{C}_C^0, \mathbf{D}_C^0)$ is a realization of $C(z)$, all the realizations of $C(z)$ form a set

$$\mathcal{S}_C = \left\{ \mathbf{W} : \mathbf{W} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{T}^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{D}_C^0 & \mathbf{C}_C^0 \\ \mathbf{B}_C^0 & \mathbf{A}_C^0 \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{T} \end{bmatrix} \right\} \tag{1.12}$$

where the transformation matrix $\mathbf{T} \in \mathcal{R}^{n \times n}$ is an arbitrary non-singular matrix. A useful observation is that different $\mathbf{W}$ have different values of v. Provided that the value of v is computationally tractable, an optimal realization of $C(z)$, which has a maximum tolerance to FWL errors, can be obtained via optimization.

The open problem defined in section 1.1 was first seen in [3]. At present, there exists no available result. An approach to bypass the difficulty in computing v is to define some approximate upper bound of v using a first-order approximation, which is computationally tractable (see [3]).

One of the thorny items in the open problem is the Hadamard product $\mathbf{W} \circ \boldsymbol{\Delta}$. The form of structured perturbation, which was adopted in μ-analysis methods [4], may be used to deal with this Hadamard product: $\boldsymbol{\Delta}$ can be transformed into a generalized perturbation $\tilde{\boldsymbol{\Delta}}$ which has certain structure such as block-diagonal. The fixed matrices $\tilde{\mathbf{A}}$, $\tilde{\mathbf{B}}$ and $\tilde{\mathbf{C}}$ may be obtained such that the stability of $\tilde{\mathbf{A}} + \tilde{\mathbf{B}}\tilde{\boldsymbol{\Delta}}\tilde{\mathbf{C}}$ is equivalent to that of $\mathbf{A} + \mathbf{B}(\mathbf{W} \circ \boldsymbol{\Delta})\mathbf{C}$. Although the stability of $\tilde{\mathbf{A}} + \tilde{\mathbf{B}}\tilde{\boldsymbol{\Delta}}\tilde{\mathbf{C}}$ can be treated satisfactorily by μ-analysis methods, the open problem cannot be solved successfully by μ-analysis methods. This is because μ-analysis methods are concerned about the maximal singular value $\overline{\sigma}(\tilde{\boldsymbol{\Delta}})$ of $\tilde{\boldsymbol{\Delta}}$. In fact, the distance between $\overline{\sigma}(\tilde{\boldsymbol{\Delta}})$ and $\|\boldsymbol{\Delta}\|_{\max}$ can be quite large, and $\|\boldsymbol{\Delta}\|_{\max}$ is the other thorny item which makes the open problem difficult.

Acknowledgements

Bibliography

[1] M. Gevers and G. Li, *Parameterizations in Control, Estimation and Filtering Problems: Accuracy Aspects*. London: Springer Verlag, 1993.

[2] R.S.H. Istepanian and J.F. Whidborne, Eds., *Digital Controller Implementation and Fragility: A Modern Perspective*. London: Springer Verlag, 2001.

[3] J. Wu, S. Chen, J.F. Whidborne and J. Chu, "Optimal floating-point realizations of finite-precision digital controllers," In *Proc. 41st IEEE Conference on Decision and Control* (Las Vegas, USA), Dec.10-13, 2002, pp.2570–2575.

[4] J. Doyle, "Analysis of feedback systems with structured uncertainties," *Proc. IEE*, Vol.129-D, No.6, pp.242–250, 1982.