ISSN 1562 - 2703

INTERNATIONAL JOURNAL AUTOMATION AUSTRIA

Heft 1 im Jahrgang 20 (2012)

Inhalt *Seite*

SCOPE

"International Journal Automation Austria" (IJAA) publishes top quality, peer reviewed papers in all areas of automatic control concerning continuous and discrete processes and production systems.

Only original papers will be considered. No paper published previously in another journal, transaction or book will be accepted. Material published in workshops or symposia proceedings will be considered. In such a case the author is responsible for obtaining the necessary copyright releases. In all cases, the author must obtain the necessary copyright releases before preparing a submission. Papers are solicited in both theory and applications

Before preparing submissions, please visit our instructions to authors (see back cover) or web page.

International Journal Automation Austria also is the official executive authority for publications of IFAC-Beirat Österreich.

Imprint: Propagation of Automatic Control in Theory and Practice.

Frequency: Aperiodically, usually twice a year.
Publisher: IFAC-Beirat Österreich, Peter Kopacek, Alexander Weinmann
Editors in Chief: Alexander Weinmann, Peter Kopacek

Address: 1) Institut für Automatisierungs- und Regelungstechnik (E376), TU Wien,
A-1040 Wien, Gußhausstrasse 27-29, Austria
Phone: +43 1 58801 37677, FAX: +43 1 58801 37699
email: danzinger@acin.tuwien.ac.at
Homepage: http://www.acin.tuwien.ac.at/publikationen/ijaa/

2) Intelligente Handhabungs- und Robotertechnik (E325/A6), TU-Wien,
A-1040 Wien, Favoritenstrasse 9-11, Austria
email: e318@ihrt.tuwien.ac.at

Layout: Rainer Danzinger
Printing: Grafisches Zentrum an der TU Wien

PERFORMANCE GRADIENTS IN AUTOMATIC CONTROL ANALYSIS
A Catalogue of Correspondences
Second Edition

Alexander Weinmann, OVE, Senior Member IEEE
Vienna University of Technology, Institute of Automatic Control
Gusshausstrasse 27-29/376, A-1040 Vienna / Austria
Phone: +43 1 58801 37611, Fax: +43 1 58801 37699
email: weinmann@acin.tuwien.ac.at

Manuscript received March 26, 2012

Abstract: *A comprehensive collection of relevant formulas and correspondencies is given for differential quotients of expressions in norm, trace, determinant and eigenvalue with respect to vectors and matrices as often used in automatic control.*

The results are provided for several indices of performance: Stability radius, entirety of eigenvalues, integral criteria of squared structures as well as classical control setups and dynamic interaction of multiple-input multiple-output systems.

The controllers are investigated in many representations: Transfer functions, state space, polynomial matrices, three-dimensional arrays and s-power-oriented coefficient matrices.

Both continuous-time and discrete-time representations are in use.

Keywords: Stability margin, transfer matrices, polynomial matrices, power-oriented matrices, three-dimensional arrays, square energy performance, per-unit square index of performance, interaction energy gradient, perturbed systems, robustness, robustification.

Contents

Index

1 Introduction

Several properties of control systems are given by scalars, e.g., eigenvalues, singular values, norms, traces, determinants, condition numbers, indices of performance etc.

The paper addresses performance criteria as used in the field of automatic control. In most cases they are utilized to determine gradients of the controller implementations if optimal design of the controller is intended.

The results are provided for several indices of performance: Stability radius, entirety of eigenvalues, integral criteria of squared structures as well as classical control setups and dynamic interaction of multiple-input multiple-output systems.

The controllers are investigated in many representations: Transfer functions, state space, polynomial matrices, three-dimensional arrays and s-power-oriented coefficient matrices. Both continuous-time and discrete-time representations are in use.

There are several disadvantages for gradients if they are considered from a purely mathematically abstract viewpoint, such as selection of a specific performance index, unknown stepsize, vague stopping condition, the local minimum problem.

There is also a plenty of advantages if the problems are considered based on a good physical, chemical or economic insight, selection of various performance indices and a combination of their gradients, the free and process-oriented selection of step size and accuracy-oriented stopping condition. Process uncertainties and robustness considerations can be included without difficulty.

In a multitude of control system references, the stability border of a control system is investigated. By way of contrast, this paper addresses the gradients starting from an arbitrary nonoptimum controller assumption which might arise on the occasion of some refurbishment of an automatic control system.

Intentionally, throughout the paper several details and some redundant remarks are included to avoid any unclearness, e.g., as far as matrix dimensions, transpositions etc. are concerned. The index is based on the equation number Eq.$(\cdot)$, not on the journal page. It was intended to present a comprehensive index with deep itemizing. Even synonyms were included to guarantee the access to all the details included in the collection, and to keep the orientation in rather complex subjects.

The formulas are collected from numerous publications. Several are selected from recent publications of the author. This second edition offers a 70 percent increase in volume versus a former version (*Weinmann, A., 2008c*). A few examples were included and some derivations extended to give more illustration.

1.1 Notations

Many correspondences in this paper are influenced by the intention of straightforward optimal control system design. optimization. Among these expressions are the parameter vector $\mathbf{p}$, the matrix $\mathbf{K}$ of the controller in state space or a general matrix $\mathbf{M}$.

The symbols p and $\mathbf{p}$ are used for general scalar and vector-valued parameters, respectively. For a general matrix the term $\mathbf{M}$ or $\mathbf{X}$ is chosen; the symbol $\mathbf{P}$ is avoided since it is commonly used as the Lyapunov matrix. If in practice an expression is only used associated with controller matrix then $\mathbf{M}$ is replaced by $\mathbf{K}$.

The following notations are adopted. The set of real and complex matrices is $\mathcal{R}^{m\times n}$ and $\mathcal{C}^{f\times z}$, respectively, where the superscript denotes the dimension. Alternatively, the dimension is often denoted by $\mathbf{e}^{[r]}$, $\mathbf{A}^{[n\times m]}$ or $\mathbf{A}^{(n\times m)}$ or $\in \mathcal{R}^{n\times m}$. $\Re e$ and $\Im m$ assign the real and imaginary part of a complex quantity. The identity matrix of dimension $n \times n$ is $\mathbf{I}_n$. The meaning of $i = \{1, 4\}$ is all the integers between 1 and 4. The entire continuum between 1 and r is termed $k = [1, r]$. Hence, an interval polynomial is, e.g., $s^3 + [1, 2]s^2 + 5s + 3$ (*Barmish, B.R., 1994*). A set of a polynomial family is termed $\mathcal{P} = \{p(s, q) : q \in Q\}$. The vector $\mathbf{s}^H \overset{\triangle}{=} (\mathbf{s}^*)^T$ is the "Hermite" conjugate transpose.

Another notation example: For the stable open-loop system F_o and unity feedback, the Nyquist stability criterion can be written as $F_o(j\omega) \notin (-\infty, -1]\forall\omega \in \mathcal{R}$.

The jth column (row) of a matrix $\mathbf{M}$ is $M_{\cdot j}$ ($M_{i\cdot}$), corresponding to MATLAB $\mathtt{M}(:,\mathtt{j})$ and $\mathtt{M}(\mathtt{i},:)$, respectively.

The meaning of $\overset{(61)}{=}$ is that for evolving the current execution, the result of Eq.(61) was utilized. For the sake of abbreviation $[\mathbf{M};\ \mathbf{N}] \overset{\triangle}{=} \begin{bmatrix} \mathbf{M} \\ \mathbf{N} \end{bmatrix}$ is used for columnizing. "col" is an operator concatenating all the column partitions of a matrix; "loc" is the opposite operation.

2 Algebraic and Combinatoric Matrix Correspondences

In general, the elements of all vectors $\mathbf{p}$ or matrices $\mathbf{M}$, $\mathbf{K}$ are considered as independent variables with no interelement dependencies. Considering a derivative of a function with respect to the vector $\mathbf{p}$ or the matrix $\mathbf{M}$, vector or matrix interrelations in $\mathbf{p}$ or $\mathbf{M}$, respectively, must not exist, in general.

Matrix element interrelations (dependencies) are addressed in a specific Section 32.

For the sake of abbreviation,

$$(\mathbf{A}^T)^{-1} \equiv \mathbf{A}^{-T} \tag{1}$$

is used. The superscript ^{H} („Hermite“) means conjugate transpose

$$\mathbf{A}^H = (\mathbf{A}^*)^T \ . \tag{2}$$

$$\text{Kronecker delta} \quad \delta_{ij} = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases} \ . \tag{3}$$

$$\text{Unit vector} \quad \mathbf{e}_k = (0, 0, \ldots 1, \ldots, 0, 0)^T = \mathbf{e}_k^{(m\times 1)} \quad (1 \text{ at position } k \text{ only}) \tag{4}$$

$$\text{Kronecker matrix} \quad \mathbf{E}_{ik}^{(n\times m)} = \mathbf{e}_i \mathbf{e}_k^T \in \mathcal{R}^{n\times m} \quad (1 \text{ at position } (i,k) \text{ only}). \tag{5}$$

Selection via unit vector and Kronecker matrix: The unit vector supports a selection of a row/column or element of $\mathbf{A}$, respectively,

$$\mathbf{e}_i^T \mathbf{A} = A_{i\cdot}, \quad \mathbf{A}\mathbf{e}_k = A_{\cdot k} \ , \ \mathbf{e}_j^T \mathbf{A} \mathbf{e}_i = A_{ji} \ . \tag{6}$$

Selection of a single matrix element

$$\mathrm{tr}[\mathbf{E}_{ij}\mathbf{A}] = A_{ji} \ . \tag{7}$$

Exchange of rows i and k in the matrix $\mathbf{R}^{[n\times k]}$ (or in a vector) by premultiplication with $\mathbf{V}_T$. Select

$$\mathbf{V}_T = \mathbf{I}_n \ \text{ but } \mathbf{V}_T(i,i) = 0; \ \mathbf{V}_T(j,j) = 0; \ 1 = \mathbf{V}_T(i,j) = \mathbf{V}_T(j,i); \quad \mathbf{R} \mapsto \mathbf{V}_T \ \mathbf{R} \ . \tag{8}$$

Correspondingly, matrix columns can be exchanged by postmultiplication.

Kronecker matrix provides a useful selection. Premultiplication of a matrix $\mathbf{A}$ by $\mathbf{E}_{ij}$ yields the matrix $\mathbf{V}$

$$\mathbf{V} = \mathbf{E}_{ij}\mathbf{A} = \begin{cases} \text{null matrix } \mathbf{0} \\ \text{except the } i\text{th row } V_{i\cdot} = A_{j\cdot} \end{cases} , \tag{9}$$

i.e., a matrix of zeros where the jth row of $\mathbf{A}$ is assigned to the ith row.

Postmultiplication of a matrix $\mathbf{A}$ by $\mathbf{E}_{ij}$ yields the matrix $\mathbf{W}$

$$\mathbf{W} = \mathbf{A}\mathbf{E}_{ij} = \begin{cases} \text{null matrix } \mathbf{0} \\ \text{except the } j\text{th column } W_{\cdot j} = A_{\cdot i} \end{cases} \tag{10}$$

i.e., a matrix of zeros where the ith column of $\mathbf{A}$ is displaced to the jth column.

For $\mathbf{M} \in \mathcal{R}^{n\times n}$:

$$\det(-\mathbf{M}) = (-1)^n \det(\mathbf{M}) \tag{11}$$

$$\mathbf{adj}(-\mathbf{M}) = (-1)^{n+1}\mathbf{adj}(\mathbf{M}) \tag{12}$$

$$(-\mathbf{M})^{-1} = -\mathbf{M}^{-1} \quad \text{for arbitrary } n. \tag{13}$$

Given a matrix $\mathbf{A} = \mathbf{matrix}[A_{ij}]$, striking out the i-th row and the j-column and evaluating the determinant, one gets the minors, multiplying with $(-1)^{i+j}$ one gets the cofactors q_{ij}, then the adjoint of $\mathbf{A}$ is $\mathbf{adj}\ [\mathbf{A}] = (\mathbf{matrix}[q_{ij}])^T$, and finally the inverse is $\mathbf{A}^{-1} = \mathbf{adj}\ [\mathbf{A}]/\det \mathbf{A}$. Only for $\mathbf{A} \in \mathcal{C}^{2\times 2}$ one has the simple expression $\mathbf{adj}\ [\mathbf{A}] = \begin{pmatrix} A_{22} & -A_{12} \\ -A_{21} & A_{11} \end{pmatrix}$.

Changing the order in a product of symmetric matrices: For $\mathbf{V} = \mathbf{V}^T$ and $\mathbf{W} = \mathbf{W}^T$ the product $\mathbf{VW}$ is not symmetric $\mathbf{VW} \neq (\mathbf{VW})^T = \mathbf{WV}$.

For $\mathbf{A},\ \mathbf{B}$ commutative, one has for the trace tr (sum of main-diagonal elements of a square matrix)

$$\mathrm{tr}[\mathbf{ab}^T] = \mathbf{a}^T\mathbf{b}\ , \quad \mathrm{tr}[\mathbf{AB}] = \mathrm{tr}[\mathbf{BA}] = \mathrm{tr}[(\mathbf{AB})^T] \tag{14}$$

$$\sum_{k=1}^{n} \mathbf{E}_{kk}\mathbf{A}^{(n\times n)} = \mathbf{A} \tag{15}$$

$$\mathrm{tr}[\mathbf{E}_{ij}] = I_{ji} = \delta_{ij} \quad (= 1 \ \text{ for } i = j; \text{ and } = 0 \text{ for } i \neq j) \tag{16}$$

$$\mathrm{tr}[\mathbf{ME}_{ij}] = M_{ji} = (\mathbf{M}^T)_{ij}\ . \tag{17}$$

Note that $e^{\mathbf{A}(t+\tau)} = e^{\mathbf{A}t}e^{\mathbf{A}\tau}$ but $e^{\mathbf{A}+\mathbf{BK}} \neq e^{\mathbf{A}}e^{\mathbf{BK}}$ even for matrices small in norm sense. The matrix product is not commutative, the order of matrix addtion is arbitrary

$$e^{\mathbf{A}}e^{\mathbf{BK}} \neq e^{\mathbf{BK}}e^{\mathbf{A}} \quad \text{but} \quad \mathbf{A} + \mathbf{BK} = \mathbf{BK} + \mathbf{A}\ , \tag{18}$$

hence $e^{\mathbf{A}+\mathbf{BK}} \neq e^{\mathbf{A}}e^{\mathbf{BK}}$. Nevertheless, for the scalars t_1 and t_2

$$e^{\mathbf{A}(t_1+t_2)} = e^{\mathbf{A}t_1+\mathbf{A}t_2} = e^{\mathbf{A}t_1}e^{\mathbf{A}t_2} = e^{\mathbf{A}t_2}e^{\mathbf{A}t_1} \quad \text{because} \quad (\mathbf{A}t_1)(\mathbf{A}t_2) = (\mathbf{A}t_2)(\mathbf{A}t_1)\ . \tag{19}$$

The product of matrices $\mathbf{A}$ and $\mathbf{B}$ is commutative if $\mathbf{B}$ is a function of $\mathbf{A}$, e.g., $\mathbf{B} = (\mathbf{I} - \mathbf{A})^{-1}$ or $\mathbf{B} = e^{\mathbf{A}}$.

Matrix exponential and matrix (natural) logarithm[1]

$$e^{\mathbf{A}} = \mathbf{I} + \mathbf{A} + \frac{1}{2!}\mathbf{A}^2 + \ \ldots \frac{1}{\nu!}\mathbf{A}^\nu + \ldots \tag{20}$$

$$\log \mathbf{A} = \mathbf{A} - \mathbf{I} - \frac{1}{2}(\mathbf{A}-\mathbf{I})^2 + \frac{1}{3}(\mathbf{A}-\mathbf{I})^3 - \ldots + \frac{(-1)^{\nu+1}}{\nu}(\mathbf{A}-\mathbf{I})^\nu \quad \text{for } \|(\mathbf{A}-\mathbf{I})\|_F < 1 \,. \tag{21}$$

Hence, $e^{\log \mathbf{A}} = \mathbf{A}, \quad \log e^{\mathbf{A}} = \mathbf{A}$.

Relation between determinant, matrix exponential and trace

$$\det e^{\mathbf{A}} = e^{\mathrm{tr}[\mathbf{A}]} \,. \tag{22}$$

The Kronecker product $\otimes$ of two matrices (or direct product, tensor product) is defined by a partitioned matrix whose (i,j)-partition is $A_{ij}\mathbf{B}$

$$\mathbf{A}\otimes\mathbf{B} \triangleq \begin{pmatrix} A_{11}\mathbf{B} & A_{12}\mathbf{B} & \ldots & A_{1n}\mathbf{B} \\ A_{21}\mathbf{B} & A_{22}\mathbf{B} & \ldots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ A_{m1}\mathbf{B} & A_{m2}\mathbf{B} & \ldots & A_{mn}\mathbf{B} \end{pmatrix} = \mathbf{matrix}[A_{ij}\mathbf{B}] \,, \qquad \begin{array}{l} \mathbf{A} \in \mathcal{C}^{n\times m} \\ \mathbf{B} \in \mathcal{C}^{r\times s} \\ \mathbf{A}\otimes\mathbf{B} \in \mathcal{C}^{nr\times ms} \,. \end{array} \tag{23}$$

The Kronecker sum $\oplus$ is defined by (*Brewer, J.W., 1978*)

$$\mathbf{N}\oplus\mathbf{M} = \mathbf{N}\otimes\mathbf{I}_m + \mathbf{I}_n\otimes\mathbf{M} \,; \qquad \mathbf{N}\in\mathcal{C}^{n\times n}, \ \ \mathbf{M}\in\mathcal{C}^{m\times m}, \ \ \mathbf{N}\oplus\mathbf{M}\in\mathcal{C}^{mn\times mn} \,. \tag{24}$$

The matrix $\mathbf{U}_{r,m}$ is the permutation matrix in Kronecker matrix sense; it is always square and consists of entries zero except one solitary digit one in *each* row and column. $\mathbf{E}_{ij}$ is the Kronecker matrix of Eq.(5) with entries zero except a *solitary* digit one at position i,j

$$\mathbf{U}_{k,l} \triangleq \mathbf{U}_{k,l}^{(kl\times kl)} \triangleq \sum_{i=1}^{k}\sum_{j=1}^{l} \mathbf{E}_{ij}^{(k\times l)} \otimes \mathbf{E}_{ji}^{(l\times k)} = \sum_{i=1}^{k}\sum_{j=1}^{l} \mathbf{E}_{ij}^{(k\times l)} \otimes \left(\mathbf{E}_{ij}^{(k\times l)}\right)^T. \tag{25}$$

E.g.,

$$\mathbf{E}_{1,3}^{(2\times 4)} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \tag{26}$$

$$\mathbf{U}_{2,3}^{(6\times 6)} = \begin{pmatrix} 1&0&0&0&0&0 \\ 0&0&1&0&0&0 \\ 0&0&0&0&1&0 \\ 0&1&0&0&0&0 \\ 0&0&0&1&0&0 \\ 0&0&0&0&0&1 \end{pmatrix}, \quad \mathbf{U}_{3,2}^{(6\times 6)} = \begin{pmatrix} 1&0&0&0&0&0 \\ 0&0&0&1&0&0 \\ 0&1&0&0&0&0 \\ 0&0&0&0&1&0 \\ 0&0&1&0&0&0 \\ 0&0&0&0&0&1 \end{pmatrix}. \tag{27}$$

[1] Remember MATLAB matrix notation `expm` and `logm` for matrices.

Note $\mathbf{U}_{1,n} = \mathbf{U}_{n,1} = \mathbf{I}_n$, and the orthogonality property $\mathbf{U}_{k,l}^{-1} = \mathbf{U}_{k,l}^T = \mathbf{U}_{l,k}$,

$$\mathbf{U}_{k,l}^T \mathbf{U}_{k,l} = \mathbf{I}_{(kl)}^{(kl \times kl)} \ , \tag{28}$$

and

$$(\mathbf{I}_r \otimes \mathbf{U}_{k,l})^{-1} = \mathbf{I}_r \otimes \mathbf{U}_{k,l}^T = \mathbf{I}_r \otimes \mathbf{U}_{l,k} \ . \tag{29}$$

In addition, one has

$$\frac{\partial \mathbf{M}^T}{\partial \mathbf{M}} = \mathbf{U}_{k,l} \qquad \mathbf{M} \in \mathcal{R}^{k \times l} \tag{30}$$

and

$$\text{col } \mathbf{M}^T = \mathbf{U}_{k,l} \text{ col}\mathbf{M} \qquad \mathbf{M} \in \mathcal{R}^{k \times l} \ . \tag{31}$$

As a special case of the permutation matrix, the rotation or reflection matrix is defined as the unity matrix with unity entries in the secondary diagonal, only,

$$\mathbf{I}^{\swarrow} \triangleq \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} . \tag{32}$$

For comparison, the unity matrix reads as

$$\mathbf{I} = \mathbf{I}^{\searrow} = \mathbf{diag}\{1, 1, \ldots, 1\} \ . \tag{33}$$

The symplectic matrix and its properties are

$$\mathbf{I}^{\oslash} \triangleq \begin{pmatrix} \mathbf{0} & -\mathbf{I}_m \\ \mathbf{I}_m & \mathbf{0} \end{pmatrix} \in \mathcal{C}^{2m \times 2m} \qquad \mathbf{I}^{\oslash -1} = \mathbf{I}^{\oslash T} = -\mathbf{I}^{\oslash} \ , \tag{34}$$

and, surprisingly, $(\mathbf{I}^{\oslash})^2 = -\mathbf{I}_{2m}$, i.e., the same feature as the square of the scalar complex unit $(j)^2 = (\sqrt{-1})^2 = -1$.

An unitary matrix $\mathbf{V}$ has the property $\mathbf{V}^H = \mathbf{V}^{-1}$. An involutory matrix is both Hermitian and unitary, that is, $\mathbf{G}^H = \mathbf{G}$ and $\mathbf{G}^H = \mathbf{G}^{-1}$,

$$\mathbf{G} = \mathbf{G}^{-1} \quad \text{and} \quad \mathbf{G}^2 = \mathbf{I} \ . \tag{35}$$

A matrix is denoted half involutory if it is both skew-Hermitian and unitary, i.e.,

$$\mathbf{G}^2 = -\mathbf{I} \ . \tag{36}$$

The Riccati coefficient matrix $\mathbf{M}_R$ (or associated Hamiltonian), based on the Riccati matrix $\mathbf{P}(t)$ in the Riccati differential equation

$$-\dot{\mathbf{P}}(t) = \mathbf{A}^T\mathbf{P} + \mathbf{P}\mathbf{A} + \mathbf{Q} - \mathbf{P}\mathbf{B}\mathbf{R}^{-1}\mathbf{B}^T\mathbf{P} \qquad \mathbf{P}(0) = \mathbf{P}_o \ . \tag{37}$$

is

$$\mathbf{M}_R = \begin{pmatrix} \mathbf{A} & \mathbf{B}\mathbf{R}^{-1}\mathbf{B}^T \\ \mathbf{Q} & -\mathbf{A}^T \end{pmatrix} \qquad \mathbf{M}_R \in \mathcal{R}^{2n\times 2n} \quad , \tag{38}$$

see also Eq.(822). Defining a square $(2n, 2n)$-matrix

$$\mathbf{I}^{\oslash} \triangleq \begin{pmatrix} \mathbf{0} & -\mathbf{I}_n \\ \mathbf{I}_n & \mathbf{0} \end{pmatrix} \qquad \mathbf{I}^{\oslash -1} = -\mathbf{I}^{\oslash} \ , \tag{39}$$

the Riccati coefficient matrix $\mathbf{M}_R$ obeys the Hamiltonian property

$$\mathbf{I}^{\oslash}\ \mathbf{M}_R^T\ \mathbf{I}^{\oslash} = \mathbf{M}_R \ \text{ or } \ -(\mathbf{I}^{\oslash})\ \mathbf{M}_R^T\ (\mathbf{I}^{\oslash}) = -\mathbf{M}_R \ , \tag{40}$$

that is, $\mathbf{M}_R$ is similar to $-\mathbf{M}_R^T$. Decomposing the matrix $\mathbf{P}(t)$ to eigenvectors of $\mathbf{M}_R$ of Eq.(38) leads to the sensitivity with respect to an arbitrary matrix (*Brewer, J.W., 1977a*).

The self-derivative matrix (rectangular matrix $\partial\mathbf{M}/\partial\mathbf{M}$, only square if $k = l$,) is given by

$$\frac{\partial \mathbf{M}^{(k\times l)}}{\partial \mathbf{M}} = \bar{\mathbf{U}}_{k,l} = \bar{\mathbf{U}}_{k,l}^{(k^2\times l^2)} \triangleq \sum_{i=1}^{k}\sum_{j=1}^{l} \mathbf{E}_{ij}^{(k\times l)} \otimes \mathbf{E}_{ij}^{(k\times l)} \ , \tag{41}$$

e.g.,

$$\bar{\mathbf{U}}_{2,1}^{(4\times 1)} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} , \quad \bar{\mathbf{U}}_{2,3}^{(4\times 9)} = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{pmatrix} . \tag{42}$$

Mixed product rule (only applicable if $\mathbf{A}, \mathbf{D}$ and $\mathbf{B}, \mathbf{G}$ are conformable)

$$(\mathbf{A}\otimes\mathbf{B})\ (\mathbf{D}\otimes\mathbf{G}) = (\mathbf{A}\mathbf{D})\otimes(\mathbf{B}\mathbf{G}) \ . \tag{43}$$

$$(\mathbf{N}\otimes\mathbf{M})^{-1} = \mathbf{N}^{-1}\otimes\mathbf{M}^{-1} \ . \tag{44}$$

$$\mathbf{G}^{(k\times l)}\otimes\mathbf{H}^{(n\times m)} \equiv \mathbf{U}_{k,n}\ (\mathbf{H}\otimes\mathbf{G})\ \mathbf{U}_{m,l} \tag{45}$$

The Frobenius norm of the Kronecker product is

$$||\mathbf{I}_m\otimes\mathbf{M}||_F = \sqrt{m}||\mathbf{M}||_F \ , \tag{46}$$

moreover

$$||(\text{col } \mathbf{I}_m)\otimes\mathbf{I}_n||_F = \sqrt{mn} \tag{47}$$

$$||\mathbf{x}_0\mathbf{x}_0^T||_F = ||\text{col}(\mathbf{x}_0\mathbf{x}_0^T)||_F = ||\mathbf{x}_0||_F^2 \tag{48}$$

$$\text{tr}[\mathbf{M}^{(n\times n)}] \equiv (\text{col } \mathbf{I}_n)^T \text{col } \mathbf{M} \tag{49}$$

Only for $\mathbf{M} \in \mathcal{R}^{2\times 2}$:

$$\mathrm{tr}[\mathbf{adj}\ \mathbf{M}] = \mathrm{tr}[\mathbf{M}]; \quad \|\mathbf{adj}\ \mathbf{M}\|_F = \|\mathbf{M}\|_F; \quad \|\mathbf{M}^{-1}\|_F = \|\mathbf{M}\|_F / |\ \det\ \mathbf{M}\ | \ . \tag{50}$$

$$\mathrm{col}(\mathbf{HFG}) \equiv (\mathbf{G}^T \otimes \mathbf{H})\mathrm{col}\mathbf{F} \tag{51}$$

e.g., $\mathrm{col}(\mathbf{ABCD}) = [(\mathbf{CD})^T \otimes \mathbf{A}]\mathrm{col}\ \mathbf{B} = [\mathbf{D}^T \otimes (\mathbf{AB})]\mathrm{col}\ \mathbf{C} = [(\mathbf{BCD})^T \otimes \mathbf{I}]\mathrm{col}\ \mathbf{A}$ (52)

$$\mathrm{col}[\mathbf{AM} + \mathbf{MB}] = (\mathbf{B}^T \oplus \mathbf{A})\ \mathrm{col}\,\mathbf{M}. \tag{53}$$

Here the abbreviation "col" (or "vec") is the column operator, concatenating all the columns of a matrix to an entire column vector, also written as $\mathbf{A}^\flat \equiv \mathrm{col}[\mathbf{A}]$.

$$\mathrm{col}(\mathbf{x}\mathbf{y}^T) = \mathbf{y} \otimes \mathbf{x} \tag{54}$$

$$\mathrm{col}(\mathbf{x}^T\mathbf{M}) = \mathbf{Mx}\ , \qquad \mathrm{col}(\mathbf{x}^T) = \mathbf{x} \tag{55}$$

$$\mathrm{col}(\mathbf{x}^T\mathbf{G}^T\mathbf{G}) = (\mathbf{I}_n \otimes (\mathbf{x}^T\mathbf{G}^T\mathbf{G})\mathrm{col}\,\mathbf{I}_n \stackrel{(51)}{=} \mathbf{G}^T\mathbf{G}\cdot\mathbf{x}\ . \tag{56}$$

$$\mathrm{col}\ \mathbf{K}^T = \mathbf{U}_{m,r}\mathrm{col}\ \mathbf{K}\ , \qquad \mathbf{K} \in \mathcal{R}^{m\times r} \tag{57}$$

$$\mathrm{col}\ \mathbf{I}_n \equiv \sum_{i=1}^{i=n} \mathbf{e}_i \otimes \mathbf{e}_i \tag{58}$$

$$\mathrm{col}(\mathbf{c}\ \mathbf{c}^T) = \mathrm{col}(\mathbf{c}\cdot 1\cdot \mathbf{c}^T) = \mathbf{c}\otimes\mathbf{c}, \quad \mathrm{loc}_n(\mathbf{x}_0 \otimes \mathbf{x}_0) = \mathbf{x}_0\mathbf{x}_0^T\ . \tag{59}$$

Row manipulation

$$\mathrm{row}\mathbf{A} = (\mathrm{col}\mathbf{A}^T)^T\ . \tag{60}$$

$$(\mathbf{I}_n \otimes c)\mathbf{Q}^{[n\times m]} = (c\mathbf{I}_n)\mathbf{Q} = c\mathbf{Q} = c\mathbf{Q}\mathbf{I}_m = \mathbf{Q}\mathbf{I}_m c = \mathbf{Q}(\mathbf{I}_m \otimes c) \tag{61}$$

Unexpected, and surprisingly simple,

$$(\mathbf{p}^T \otimes \mathbf{I}_n)\mathrm{col}\ \mathbf{I}_n \quad = \quad \mathbf{p} \tag{62}$$

$$(\mathbf{I}_n \otimes \mathbf{p}^T)\mathrm{col}\ \mathbf{I}_n \quad = \quad \mathbf{p}\ . \tag{63}$$

An algorithm for analytical columnizing as follows: Columnizing $\mathbf{K}_y \in \mathcal{R}^{m\times r}$

$$\mathrm{col}[\mathbf{K}_y] = \sum_{i=1}^{r} \mathbf{e}_i^{(r\times 1)} \otimes [\mathbf{K}_y \cdot (\mathbf{1}^{(1\times r)} \otimes \mathbf{e}_i^{(r\times 1)}) \cdot \mathbf{e}_i^{(r\times 1)}]\ \in \mathcal{R}^{mr}\ , \tag{64}$$

where $\mathbf{1} \stackrel{\triangle}{=} (1\ 1\ 1\ \ldots\ 1) \in \mathcal{R}^{1\times r}$. Alternatively,

$$\mathrm{col}\ [\mathbf{K}_y] = \left[\sum_{i=1}^{r} (\mathbf{e}_i^{[r]} \otimes \mathbf{I}_m)\mathbf{K}_y(\mathbf{e}_i^{[r]} \otimes \mathbf{1})\right]\mathbf{e}_1^{[r]}\ . \tag{65}$$

Due to the structure of Eq.(65)

$$\text{col}\,[\mathbf{K}_y] = \mathbf{h}_1^T \mathbf{K}_y \mathbf{l}_1 + \mathbf{h}_2^T \mathbf{K}_y \mathbf{l}_2 + \ldots , \tag{66}$$

arithmetic calculations are feasible, "col" itself is not suitable for differentiation.

A concise formula for columnization is

$$\text{col}\,\mathbf{K}_y^{[m\times r]} = \sum_{i=1}^{r}\sum_{j=1}^{m} \left[(\mathbf{e}_j^{[m]})^T\ \mathbf{K}_y\ \mathbf{e}_i^{[r]}\right]\ \mathbf{e}_{j+(i-1)m}^{[mr]} \triangleq \mathbf{h}^{[mr\times 1]}\ . \tag{67}$$

Differentiation requires

$$\frac{\partial \text{col}\,\mathbf{K}_y}{\partial p} \overset{(65)}{=} \sum_{i=1}^{r} \mathbf{h}_i^T \frac{\partial \mathbf{K}_y}{\partial p}\mathbf{l}_i = \left[\sum_{i=1}^{r}(\mathbf{e}_i^{[r]} \otimes \mathbf{I}_m)\frac{\partial \mathbf{K}_y}{\partial p}(\mathbf{e}_i^{[r]} \otimes \mathbf{1})\right]\mathbf{e}_1^{[r]}\ . \tag{68}$$

Decolumnizing $\mathbf{h} \in \mathcal{R}^{mr\times 1}$

$$\text{loc}_m[\mathbf{h}] = \sum_{i=1}^{r}(\mathbf{e}_i^{(r\times 1)})^T \otimes \{[(\mathbf{e}_i^{(r\times 1)})^T \otimes \mathbf{I}_m]\cdot \mathbf{h}\}\ . \tag{69}$$

Alternative decolumnizing

$$[\text{loc}_r\mathbf{h}]^T = [\mathbf{I}_m \otimes (\mathbf{h}^{(mr\times 1)})^T][(\text{col}\,\mathbf{I}_m)\otimes \mathbf{I}_r]\ , \tag{70}$$

$$\text{loc}_m[\mathbf{h}^{(mr\times 1)}] = \sum_{i=1}^{r}[(\mathbf{e}_i^r)^T \otimes \mathbf{I}_m][\mathbf{E}_{ii}^{rr}\otimes \mathbf{I}_m]\cdot \mathbf{h}\cdot(\mathbf{e}_i^r)^T = \sum_{i=1}^{r}[(\mathbf{e}_i^r)^T\otimes \mathbf{I}_m]\cdot \mathbf{h}\cdot(\mathbf{e}_i^r)^T \in \mathcal{R}^{m\times r}\ . \tag{71}$$

A concise version for decolumnizing is

$$\text{loc}_m[\mathbf{h}^{[mr\times 1]}] = \sum_{j=1}^{r}\sum_{i=1}^{m}(\mathbf{e}_{i+(j-1)m}^{[mr\times 1]})^T\ \mathbf{h}\ \mathbf{E}_{ij}^{[m\times r]} = \mathbf{K}_y^{[m\times r]}\ . \tag{72}$$

The dyadic product $\mathbf{a}\mathbf{b}^T$ is characterized both by

$$\det(\mathbf{a}\mathbf{b}^T) = 0 \quad \text{and} \quad \mathbf{adj}[\mathbf{a}\mathbf{b}^T] = \mathbf{0}\ . \tag{73}$$

For the dyadic-like product

$$\det(\mathbf{A}\mathbf{B}^T) = 0,\ \ \det[\mathbf{C}\otimes(\mathbf{A}\mathbf{B}^T)] = 0,\ \ \mathbf{A},\ \mathbf{B}\ \in \mathcal{R}^{n\times m},\ \ m < n,\ \mathbf{C}\ \text{arbitrary square.} \tag{74}$$

Linear combination of p separate matrices $\mathbf{M}_i^{[m\times n]}$, $n \geq m$

$$\mathbf{L} \ = \ \sum_{i=1}^{p}\beta_i\mathbf{M}_i = (\boldsymbol{\beta}\otimes \mathbf{I}_m)^T\ \ \text{blcol}[\mathbf{M}_i] \quad \in \mathcal{R}^{m\times n}\ , \tag{75}$$

$$\text{where}\quad \boldsymbol{\beta} \triangleq \text{col}\ \beta_i \in \mathcal{R}^p, \qquad \text{blcol}[\mathbf{M}_i] \triangleq [\mathbf{M}_1\ ;\ \mathbf{M}_2;\ \vdots\ ;\ \mathbf{M}_p] \quad \in \mathcal{R}^{mp\times n} \tag{76}$$

Sum of rows/columns: Given a matrix $\mathbf{M} \in \mathcal{R}^{n \times m}$, defining $\mathbf{h} = \mathbf{e}_i + \mathbf{e}_k$ (a null row except unitary ones at position i and k), premultiplying yields a row

$$\mathbf{h}^T\mathbf{M} \;=\; M_{i.} + M_{k.} \tag{77}$$

which is the sum of the ith and the kth row of $\mathbf{M}$. Postmultiplying $\mathbf{Mh}$ yields a sum of columns, correspondingly.

In order to delete a superfluous row μ and column μ in a matrix $\mathbf{M}$, utilize

$$\mathbf{I}_{n,wr\mu}\mathbf{M}\;\mathbf{I}_{n,wr\mu}^T \;, \tag{78}$$

where $\mathbf{I}_{n,wr\mu}$ is the identity matrix $\mathbf{I}_n$ without row μ.

Interrelation between determinant and trace for $\mathbf{M} \in \mathcal{C}^{2\times 2}$, only:

$$\det\mathbf{M} = 0.5\{(\mathrm{tr}[\mathbf{M}])^2 - \mathrm{tr}[\mathbf{M}^2]\} \tag{79}$$

(*Friedland, S., 1983*).

Interrelation of eigenvalues for the Kronecker matrix product:

$$\boldsymbol{\lambda}[\mathbf{A}\otimes\mathbf{B}] = \mathrm{col}\{\boldsymbol{\lambda}[\mathbf{A}]\;\boldsymbol{\lambda}^T[\mathbf{B}]\}\;, \tag{80}$$

where $\boldsymbol{\lambda}[\cdot]$ denotes the vector of eigenvalues, i.e., the spectrum.

The product of two specific Kronecker matrices, Eq.(5),

$$\mathbf{E}_{ik}\mathbf{E}_{ik}^T = \mathbf{e}_i\mathbf{e}_k^T\mathbf{e}_k\mathbf{e}_i^T = \mathbf{e}_i\mathbf{e}_i^T \tag{81}$$

is a $[k \times k]$ null matrix except a single unit at the ith diagonal position.

$$\frac{\partial\mathrm{col}\mathbf{A}^{[n\times m]}}{\partial\mathrm{col}\mathbf{A}} = \{\mathrm{col}\mathbf{I}_{n\cdot m}\}^{[(n\cdot m)^2\times 1]}\;. \tag{82}$$

Hadamard product rules for $\mathbf{A},\;\mathbf{B} \in \mathcal{R}^{n\times m}$ are

$$\mathbf{A}\cdot *\mathbf{B} = \mathbf{B}\cdot *\;\mathbf{A} = \mathrm{matrix}_{ij}[A_{ij}B_{ij}], \qquad (\mathbf{A}\cdot *\mathbf{B})^T = \mathbf{A}^T\cdot *\;\mathbf{B}^T = \mathbf{B}^T\cdot *\;\mathbf{A}^T\;. \tag{83}$$

(These rules differ from the regular matrix product

$$\mathbf{A}^{[n\times m]} = \mathbf{B}^{[n\times p]}\mathbf{C}^{[p\times m]} \qquad A_{ij} = \sum_{\nu=1}^{p} B_{i\nu}C_{\nu j}\;.) \tag{84}$$

Further rules are

$$\frac{\partial}{\partial B_{ij}}[\mathbf{A}\cdot *\;\mathbf{B}] = \mathbf{A}\cdot *\;\frac{\partial\mathbf{B}}{\partial B_{ij}} = \mathbf{A}\cdot *\;\mathbf{E}_{ij} = A_{ij}\mathbf{E}_{ij} \tag{85}$$

$$\mathbf{A} \cdot * \mathbf{B} = \mathbf{C}, \quad \sum_{ij} C_{ij} = (\mathrm{col}\mathbf{A})^T \mathrm{col}\mathbf{B} \ . \tag{86}$$

For weighted matrices

$$\frac{\partial}{\partial \mathbf{M}} \|\mathbf{W} \ \cdot * \ \mathbf{M}\|_F^2 = 2 \ \mathbf{W} \ \cdot * \ \mathbf{W} \ \cdot * \ \mathbf{M} \tag{87}$$

$$\frac{\partial \det(\mathbf{W} \ \cdot * \ \mathbf{M}}{\partial \mathbf{M}} \overset{(188)}{=} \mathbf{W} \ \cdot * \ \mathbf{adj}^T(\mathbf{W} \ \cdot * \ \mathbf{M}) \ . \tag{88}$$

2.1 Formulas Concerning the Inverse, Partitioned Matrices, and State Space

Matrix equal to its inverse:

$$\begin{pmatrix} \mathbf{I} & \mathbf{B} \\ \mathbf{0} & -\mathbf{I} \end{pmatrix}^{-1} = \begin{pmatrix} \mathbf{I} & \mathbf{B} \\ \mathbf{0} & -\mathbf{I} \end{pmatrix} \ . \tag{89}$$

Sherman-Morrison-Woodbury formula, or matrix inversion lemma (see Fig. 1), if all the matrices involved are nonsingular:

$$(\mathbf{A} + \mathbf{BDC})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}(\mathbf{D}^{-1} + \mathbf{CA}^{-1}\mathbf{B})^{-1}\mathbf{CA}^{-1} \tag{90}$$

$$(\mathbf{A} + \mathbf{UV}^T)^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{U}(\mathbf{I} + \mathbf{V}^T\mathbf{A}^{-1}\mathbf{U})^{-1}\mathbf{V}^T\mathbf{A}^{-1}. \tag{91}$$

For $\mathbf{A} \in \mathcal{R}^{n\times n}$, $\mathbf{B} \in \mathcal{R}^{n\times r}$, $\mathbf{C} \in \mathcal{R}^{r\times n}$, $\mathbf{D} \in \mathcal{R}^{r\times r}$, note the economizing computational effect, especially for $r \ll n$: If $\mathbf{A}^{-1}$ is given, then due to Eq.(90) the inverse of $(\mathbf{A} + \mathbf{BDC})$ requires the inverse of an (r, r)-matrix, only.

$$(\mathbf{A} + \mathbf{D})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}(\mathbf{D}^{-1} + \mathbf{A}^{-1})^{-1}\mathbf{A}^{-1} \tag{92}$$

or, for $\Delta\mathbf{D}$ small in norm sense,

$$(\mathbf{A} + \Delta\mathbf{D})^{-1} \doteq \mathbf{A}^{-1} - \mathbf{A}^{-1} \cdot \Delta\mathbf{D} \cdot \mathbf{A}^{-1} \tag{93}$$

$$(\mathbf{U} + \mathbf{B} \cdot \Delta\mathbf{D} \cdot \mathbf{C})^{-1} \doteq \mathbf{U}^{-1} - \mathbf{U}^{-1}\mathbf{B} \cdot \Delta\mathbf{D} \cdot \mathbf{CU}^{-1}. \tag{94}$$

Avoiding the inverse of a complex matrix, where only the inverse of a real matrix needed

$$(\mathbf{A} \pm j\mathbf{B})^{-1} = (\mathbf{B}^{-1}\mathbf{A} \mp j\mathbf{I})(\mathbf{AB}^{-1}\mathbf{A} + \mathbf{B})^{-1}. \tag{95}$$

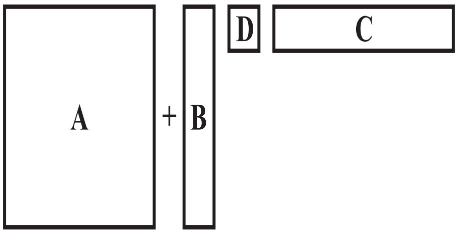

Figure 1: Matrix dimensions where Sherman-Morrison-Woodbury formula is advantageous from the numerical viewpoint

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix}^{-1} = \begin{pmatrix} \mathbf{A}^{-1} + \mathbf{A}^{-1}\mathbf{BECA}^{-1} & -\mathbf{A}^{-1}\mathbf{BE} \\ -\mathbf{ECA}^{-1} & \mathbf{E} \end{pmatrix} \tag{96}$$

$$= \begin{pmatrix} \mathbf{F} & -\mathbf{A}^{-1}\mathbf{BE} \\ -\mathbf{D}^{-1}\mathbf{CF} & \mathbf{E} \end{pmatrix} \tag{97}$$

$$\tag{98}$$

where $\mathbf{E} \triangleq (\mathbf{D} - \mathbf{CA}^{-1}\mathbf{B})^{-1}$, $\mathbf{F} \triangleq (\mathbf{A} - \mathbf{BD}^{-1}\mathbf{C})^{-1}$.

Alternatively, if $\mathbf{D}^{-1}$ exists, and for $\mathbf{F} \triangleq (\mathbf{A} + \mathbf{BD}^{-1}\mathbf{C})^{-1}$

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix}^{-1} = \begin{pmatrix} \mathbf{F} & -\mathbf{FBD}^{-1} \\ -\mathbf{D}^{-1}\mathbf{CF} & -\mathbf{D}^{-1} + \mathbf{D}^{-1}\mathbf{CFBD}^{-1} \end{pmatrix} . \tag{99}$$

For block-triangle matrices, e.g., $\mathbf{C} = \mathbf{0}$, there are reasonable simplifications.

$$(\mathbf{A} + \mathbf{ab}^T)^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\frac{\mathbf{ab}^T}{1 + \mathbf{b}^T\mathbf{A}^{-1}\mathbf{a}}\mathbf{A}^{-1} \tag{100}$$

$$(\mathbf{I} - \mathbf{ab}^T)^{-1} = \mathbf{I} - \frac{1}{\mathbf{a}^T\mathbf{b} - 1}\mathbf{ab}^T , \tag{101}$$

Determinant of a partitioned matrix

$$\det \begin{pmatrix} \mathbf{N} & \mathbf{P} \\ \mathbf{Q} & \mathbf{R} \end{pmatrix} = \det \mathbf{N} \times \det(-\mathbf{QN}^{-1}\mathbf{P} + \mathbf{R}) = \det \mathbf{R} \times \det(-\mathbf{PR}^{-1}\mathbf{Q} + \mathbf{N}) . \tag{102}$$

For state-space applications and matrices of appropriate dimension

$$\mathbf{G}(s) \triangleq \mathbf{C}(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{B} + \mathbf{D} \triangleq \left[\begin{array}{c|c} \mathbf{A} & \mathbf{B} \\ \hline \mathbf{C} & \mathbf{D} \end{array}\right] , \tag{103}$$

with a similarity transformation $\mathbf{T}$

$$\left[\begin{array}{c|c} \mathbf{T}^{-1}\mathbf{AT} & \mathbf{T}^{-1}\mathbf{B} \\ \hline \mathbf{CT} & \mathbf{D} \end{array}\right] \equiv \left[\begin{array}{c|c} \mathbf{A} & \mathbf{B} \\ \hline \mathbf{C} & \mathbf{D} \end{array}\right] , \tag{104}$$

transposing and replacing s by $-s$

$$\left(\left[\begin{array}{c|c} \mathbf{A} & \mathbf{B} \\ \hline \mathbf{C} & \mathbf{D} \end{array}\right]\right)^T \Big|_{s\,:=\,-s} = \left[\begin{array}{c|c} -\mathbf{A}^T & -\mathbf{C}^T \\ \hline \mathbf{B}^T & \mathbf{D}^T \end{array}\right] . \tag{105}$$

Replacing $\mathbf{A},\ \mathbf{B},\ \mathbf{C},\ \mathbf{D}$ by $\mathbf{D},\ \mathbf{C},\ \mathbf{B},\ (s\mathbf{I}-\mathbf{A})^{-1}$, respectively, in Eq.(90)

$$\left(\left[\begin{array}{c|c}\mathbf{A} & \mathbf{B}\\\hline \mathbf{C} & \mathbf{D}\end{array}\right]\right)^{-1} = \left[\begin{array}{c|c}\mathbf{A}-\mathbf{B}\mathbf{D}^{-1}\mathbf{C} & \mathbf{B}\mathbf{D}^{-1}\\\hline -\mathbf{D}^{-1}\mathbf{C} & \mathbf{D}^{-1}\end{array}\right]\ . \tag{106}$$

Series interconnection of two elements with transfer matrices $\mathbf{G}_1(s)$ and $\mathbf{G}_2(s)$ in convenient order

$$\left[\begin{array}{c|c}\mathbf{A}_1 & \mathbf{B}_1\\\hline \mathbf{C}_1 & \mathbf{D}_1\end{array}\right]\times\left[\begin{array}{c|c}\mathbf{A}_2 & \mathbf{B}_2\\\hline \mathbf{C}_2 & \mathbf{D}_2\end{array}\right] = \left[\begin{array}{cc|c}\mathbf{A}_2 & \mathbf{0} & \mathbf{B}_2\\ \mathbf{B}_1\mathbf{C}_2 & \mathbf{A}_1 & \mathbf{B}_1\mathbf{D}_2\\\hline \mathbf{D}_1\mathbf{C}_2 & \mathbf{C}_1 & \mathbf{D}_1\mathbf{D}_2\end{array}\right] = \left[\begin{array}{cc|c}\mathbf{A}_1 & \mathbf{B}_1\mathbf{C}_2 & \mathbf{B}_1\mathbf{D}_2\\ \mathbf{0} & \mathbf{A}_2 & \mathbf{B}_2\\\hline \mathbf{C}_1 & \mathbf{D}_1\mathbf{C}_2 & \mathbf{D}_1\mathbf{D}_2\end{array}\right]\ . \tag{107}$$

3 Eigenvalue and Eigenvector Properties

$$\det \mathbf{A}^{[n\times n]} = \prod_1^n \lambda_i[\mathbf{A}]\ ; \qquad \mathrm{tr}[\mathbf{A}^{[n\times n]}] = \sum_1^n A_{ii} = \sum_1^n \lambda_i[\mathbf{A}]\ . \tag{108}$$

$$\det(\mathbf{I}_m + \mathbf{A}^{[m\times n]}\mathbf{B}^{[n\times m]}) = \det(\mathbf{I}_n + \mathbf{B}\mathbf{A}). \tag{109}$$

$$\text{Shifting with } \mu \text{ or } s:\quad \lambda[\mu\mathbf{I}+\mathbf{A}] = \mu+\lambda[\mathbf{A}]\ ; \qquad \lambda_i[s\mathbf{I}_n-\mathbf{A}] = s-\lambda_i[\mathbf{A}]\ . \tag{110}$$

$$\lambda_i[-\mathbf{A}] = -\lambda_i[\mathbf{A}]\ ; \qquad \det(-\mathbf{A}) = (-1)^n \det\mathbf{A} = (-1)^n\prod_{i=1}^n \lambda_i[\mathbf{A}]. \tag{111}$$

$$\det(s\mathbf{I}_n-\mathbf{A}) = s^n + s^{n-1}\sum_{i=1}^n\{-\lambda_i[\mathbf{A}]\} + \ \dots\ + \prod_{i=1}^n\{-\lambda_i[\mathbf{A}]\} = 0 \text{ (Vieta's rule)}. \tag{112}$$

$$\lambda[\mathbf{A}\mathbf{B}] = \lambda[\mathbf{B}\mathbf{A}] \qquad \text{for } \mathbf{A} \text{ and } \mathbf{B} \text{ square}\ . \tag{113}$$

The regular eigenvalues and right-eigenvectors obey

$$\det(\lambda\mathbf{I}-\mathbf{A}) = 0; \quad \mathbf{P}\mathbf{v} = \mathbf{v}\lambda[\mathbf{A}] \quad \text{or} \quad \mathbf{A}\mathbf{V} = \mathbf{V}\ \mathrm{diag}\{\lambda[\mathbf{A}]\}\ , \tag{114}$$

where $\mathbf{v}$ is a column of $\mathbf{V}$ and D= diag$\{\lambda[\mathbf{A}]\}$. Via MATLAB: [V,D]=eig(A).

The left-eigenvectors result from MATLAB [L,D]=eig(A.'). Then $\mathbf{L}^T\mathbf{V} = \mathbf{I}_n$ (because of the normalization and the orthogonal properties of the left- and right-eigenvectors) and $\mathbf{L}^T\mathbf{A} = \mathrm{diag}\{\lambda[\mathbf{A}]\}\ \mathbf{L}^T$. In the matrix $\mathbf{L}$ the eigenvectors have to be concatenated in appropriate order to obtain $\mathbf{L}^T\mathbf{V} = \mathbf{I}_n$. Numerical conditions should be kept in consideration.

4 Generalized Eigenvalues and Eigenvectors

The generalized eigenvector obeys

$$\det(\lambda[\mathbf{A}]\mathbf{Q}-\mathbf{A})=0 \qquad \mathbf{AV}=\mathbf{QV}\ \mathrm{diag}\{\lambda_{Q,i}[\mathbf{A}]\} \quad \text{or} \quad \mathbf{A}\mathbf{v}_i=\mathbf{Q}\mathbf{v}_i\lambda_{Q,i}[\mathbf{A}]\ . \tag{115}$$

The generalized eigenvector $\mathbf{v}_i$ of $\mathbf{A}$ is one of the columns of $\mathbf{V}$ and corresponds to the eigenvalue $\lambda_{Q,i}[\mathbf{A}]$ The generalizing matrix $\mathbf{Q}$ is used as an additional subscript of λ. Via MATLAB: [V,D]=eig(A,Q). However, notice that the eigenvalues and $\mathbf{V}$, $\mathbf{v}$ in Eq.(115) are different from those in Eq.(114).

For the sake of completeness: Exchanging $\mathbf{A}$ and $\mathbf{Q}$ one finds

$$\det(\lambda[\mathbf{Q}]\mathbf{A}-\mathbf{Q})=0 \qquad \mathbf{QW}=\mathbf{AW}\ \mathrm{diag}\{\lambda_{A}[\mathbf{Q}]\} \quad \text{or} \quad \mathbf{Q}\mathbf{w}_i=\mathbf{A}\mathbf{w}_i\lambda_{A,i}[\mathbf{Q}]\ . \tag{116}$$

The generalized eigenvector $\mathbf{w}_i$ of $\mathbf{Q}$ is one of the columns of $\mathbf{W}$ and corresponds to the eigenvalue $\lambda_{A,i}[\mathbf{Q}]$ The generalizing matrix $\mathbf{A}$ is used as an additional subscript of λ. Via MATLAB: [W,E]=eig(Q,A), where E=diag$\{\lambda_A[\mathbf{Q}]\}$. Note the relations: The eigenvectors remain unchanged

$$\mathbf{w}_i=\mathbf{v}_i \qquad \lambda_{A,i}[\mathbf{Q}]=\frac{1}{\lambda_{Q,i}[\mathbf{A}]}\ . \tag{117}$$

The relation $\mathbf{D}^{-1}=\mathbf{E}$ (or $\mathbf{D}^{-1}=\mathbf{E}^*$ for complex arguments) requires the elements in corresponding order. Note that [L,D]=eig(A,Q); [L1,D1]=eig(A',Q); and [L2,D2]=eig(A',Q') yield D1 $\neq$ D but D2=D.

5 Important Norms

The Frobenius norm (or Euclidian or Schur or Euler norm) is defined as

$$\|\mathbf{x}\|_F \ \triangleq\ \sqrt{|x_1|^2+|x_2|^2+\ldots|x_n|^2}=\sqrt{\mathbf{x}^H\mathbf{x}}\ . \tag{118}$$

The Frobenius matrix norm is

$$\|\mathbf{G}\|_F \ \triangleq\ +\sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n}\ |G_{ij}|^2}=+\sqrt{(\mathrm{col}\mathbf{G})^H\ \mathrm{col}\mathbf{G}} \tag{119}$$

$$=\ +\sqrt{(\mathrm{col}\mathbf{G})^{*T}\ \mathrm{col}\mathbf{G}}=+\sqrt{\mathrm{tr}\ \mathbf{G}^H\mathbf{G}}=+\sqrt{\sum_i\lambda_i[\mathbf{G}^H\mathbf{G}]}\ . \tag{120}$$

Spectral (induced) norm or Hilbert Norm

$$\|\mathbf{G}\|_s \ \triangleq\ \sup_{\mathbf{x}}\frac{\|\mathbf{Gx}\|_F}{\|\mathbf{x}\|_F} \quad \text{where } \mathbf{x}\neq\mathbf{0}\ . \tag{121}$$

Singular value

$$\sigma[\mathbf{G}] \overset{\triangle}{=} +\sqrt{\lambda[\mathbf{G}^H\mathbf{G}]} \tag{122}$$

$$\sigma_{\max} = +\sqrt{\lambda_{\max}[\mathbf{G}^H\mathbf{G}]} \equiv \|\mathbf{G}\|_s \ . \tag{123}$$

The maximum singular value is a generalization of the absolute value.

Referring to the spectral norm and the maximum singular value, one has $\|\mathbf{G}(s)\|_s = \sigma_{\max}[\mathbf{G}(s)] \leq \|\mathbf{G}(s)\|_F$, and the H_∞ norm

$$\|\mathbf{G}(s)\|_\infty \overset{\triangle}{=} \sup_\omega \sigma_{\max}[\mathbf{G}(j\omega)] \ , \tag{124}$$

the H_∞ norm can be replaced or limited by the maximum of the Frobenius norm versus frequency $\|\mathbf{G}(s)\|_\infty < \sup_\omega \|\mathbf{G}(s=j\omega)\|_F$. The replacement often simplifies the derivation but only yields a sufficient result.

According to $\frac{1}{\sqrt{n}}\|\mathbf{G}\|_F \leq \sigma_{\max}[\mathbf{G}] \leq \|\mathbf{G}\|_F$ where $\mathbf{G} \in \mathcal{C}^{n\times n}$, the extent of overbounding is assessed as $\|\mathbf{G}\|_F \leq \sqrt{n}\ \sigma_{\max}[\mathbf{G}]$. For $n = 3$ e.g., overbounding is less than 73 per cent.

For the H_∞-Norm of a partitioned matrix, one finds

$$\|\begin{pmatrix}\mathbf{A}\\\mathbf{B}\end{pmatrix}\|_\infty = \sup_\omega \sigma_{\max}[\begin{pmatrix}\mathbf{A}\\\mathbf{B}\end{pmatrix}] = \sup_\omega \sqrt{\lambda_{\max}[\mathbf{A}^H\mathbf{A}+\mathbf{B}^H\mathbf{B}]} \leq \sup_\omega \sqrt{\|\mathbf{A}\|_s^2+\|\mathbf{B}\|_s^2} = \|\begin{pmatrix}\|\mathbf{A}\|_\infty\\\|\mathbf{B}\|_\infty\end{pmatrix}\|_F \ , \tag{125}$$

$$\|\mathbf{A}\|_\infty \leq \|\begin{pmatrix}\mathbf{A}\\\mathbf{B}\end{pmatrix}\|_\infty \ , \tag{126}$$

$$\|\begin{pmatrix}\|\mathbf{A}\|_\infty\\\|\mathbf{B}\|_\infty\end{pmatrix}\|_F \frac{1}{\|\begin{pmatrix}\mathbf{A}\\\mathbf{B}\end{pmatrix}\|_\infty} \leq \|\begin{pmatrix}\|\begin{pmatrix}\mathbf{A}\\\mathbf{B}\end{pmatrix}\|_\infty\\\|\begin{pmatrix}\mathbf{A}\\\mathbf{B}\end{pmatrix}\|_\infty\end{pmatrix}\|_F \frac{1}{\|\begin{pmatrix}\mathbf{A}\\\mathbf{B}\end{pmatrix}\|_\infty} = \|\begin{pmatrix}1\\1\end{pmatrix}\|_F = \sqrt{2} \ , \tag{127}$$

$$\|\begin{pmatrix}\mathbf{A}\\\mathbf{B}\end{pmatrix}\|_\infty \leq \|\begin{pmatrix}\|\mathbf{A}\|_\infty\\\|\mathbf{B}\|_\infty\end{pmatrix}\|_F \leq \sqrt{2}\ \|\begin{pmatrix}\mathbf{A}\\\mathbf{B}\end{pmatrix}\|_\infty \ , \tag{128}$$

$$\|\begin{pmatrix}\mathbf{A}&\mathbf{B}\\\mathbf{C}&\mathbf{D}\end{pmatrix}\|_\infty \leq \|\begin{pmatrix}\|\mathbf{A}\|_\infty&\|\mathbf{B}\|_\infty\\\|\mathbf{C}\|_\infty&\|\mathbf{D}\|_\infty\end{pmatrix}\|_F \leq 2\|\begin{pmatrix}\mathbf{A}&\mathbf{B}\\\mathbf{C}&\mathbf{D}\end{pmatrix}\|_\infty \ . \tag{129}$$

6 Derivative of a Scalar-Valued Function with Respect to a Scalar

Consider the scalar parameter p (*Vetter, W.J., et al., 1972*)

$$\frac{\partial \det \mathbf{A}}{\partial p} = \mathrm{tr}\ [\frac{\partial \mathbf{A}}{\partial p}\ \mathbf{adj}\ \mathbf{A}] = \mathrm{tr}[\frac{\partial \mathbf{A}}{\partial p}\mathbf{A}^{-1}] \cdot \det \mathbf{A} \ . \tag{130}$$

$$\frac{\partial \det \mathbf{A}}{\partial A_{ij}} = \mathrm{tr}\ [\frac{\partial \mathbf{A}}{\partial A_{ij}}\ \mathbf{adj}\ \mathbf{A}] = \mathrm{tr}\ [\mathbf{E}_{ij}\mathbf{adj}\ \mathbf{A}] = (\mathbf{adj}\ \mathbf{A}^T)_{ij} = A_{ji} \ . \tag{131}$$

Derivation of $\det(p\mathbf{I}_n - \mathbf{A}),\ \mathbf{A} \in \mathcal{C}^{n\times n},\ p$ scalar, yields

$$\frac{\partial}{\partial p}\det\,(p\mathbf{I}_n - \mathbf{A}) = \sum_{i=1}^{n}\det\,(p\mathbf{I}_{n-1} - \mathbf{A}_{red\ ii}), \tag{132}$$

where $\mathbf{A}_{red\ ii}$ is a square matrix evolved from $\mathbf{A}$ when deleting the row i and the column i .

For the matrix product one has

$$\frac{\partial}{\partial p}\,\mathrm{tr}[\mathbf{A}^T(p)\mathbf{A}(p)] = \frac{\partial}{\partial p}\|\mathbf{A}(p)\|_F^2 = 2\ \mathrm{tr}[\mathbf{A}^T(p)\frac{\partial \mathbf{A}(p)}{\partial p}]\ . \tag{133}$$

Eventually, for an inverse expression

$$\frac{\partial}{\partial p}\,\mathrm{tr}\ [(p\mathbf{I}+\mathbf{C})^{-2}\mathbf{D}] = -2\ \mathrm{tr}\ [(p\mathbf{I}+\mathbf{C})^{-3}\mathbf{D}] \tag{134}$$

and for the second-order adjoint

$$\mathrm{tr}[\mathbf{A}] = \mathrm{tr}[\mathbf{adj}(\mathbf{A})] \quad \text{only for } \mathbf{A} \in \mathcal{R}^{2\times 2}\ . \tag{135}$$

Derivative of the trace of the inverse:

$$\mathrm{tr}[\mathbf{A}^{-1}] = \sum_{k=1}^{n}\mathbf{e}_k^T\mathbf{A}^{-1}\mathbf{e}_k \tag{136}$$

$$\frac{\partial \mathrm{tr}[\mathbf{A}^{-1}]}{\partial f} \overset{(313)}{=} -\sum_{k=1}^{n}\mathbf{e}_k^T\mathbf{A}^{-1}\frac{\partial \mathbf{A}}{\partial f}\mathbf{A}^{-1}\mathbf{e}_k \tag{137}$$

$$\frac{\partial \mathrm{tr}[\mathbf{A}^{-1}]}{\partial f} = -\sum_{k=1}^{n}\mathrm{tr}[\mathbf{e}_k^T\mathbf{A}^{-1}\frac{\partial \mathbf{A}}{\partial f}\mathbf{A}^{-1}\mathbf{e}_k] \tag{138}$$

$$\frac{\partial \mathrm{tr}[\mathbf{A}^{-1}]}{\partial f} \overset{(14)}{=} -\sum_{k=1}^{n}\mathrm{tr}[\mathbf{A}^{-1}\mathbf{e}_k\mathbf{e}_k^T\mathbf{A}^{-1}\frac{\partial \mathbf{A}}{\partial f}] \tag{139}$$

$$\frac{\partial \mathrm{tr}[\mathbf{A}^{-1}]}{\partial f} \overset{(5)}{=} -\sum_{k=1}^{n}\mathrm{tr}[\mathbf{A}^{-1}\mathbf{E}_{kk}\mathbf{A}^{-1}\frac{\partial \mathbf{A}}{\partial f}] \tag{140}$$

$$\frac{\partial \mathrm{tr}[\mathbf{A}^{-1}]}{\partial f} = -\mathrm{tr}[(\sum_{k=1}^{n}\mathbf{A}^{-1}\mathbf{E}_{kk}\mathbf{A}^{-1})\frac{\partial \mathbf{A}}{\partial f}] \tag{141}$$

$$\frac{\partial \mathrm{tr}[\mathbf{A}^{-1}]}{\partial f} \overset{(15)}{=} -\mathrm{tr}[\mathbf{A}^{-2}\frac{\partial \mathbf{A}}{\partial f}] \tag{142}$$

$$\frac{\partial \mathrm{tr}[\mathbf{A}^{-1}]}{\partial A_{ij}} = -\mathrm{tr}[\mathbf{A}^{-2}\mathbf{E}_{ij}] \overset{(17)}{=} -(\mathbf{A}^{-2})_{ji}\ . \tag{143}$$

The Taylor expansion (power-series expansion) of a scalar f for $\mathbf{x} \in \mathcal{R}^2$ and $\mathbf{x}$ close to $\mathbf{x}_o$ is

$$f(\mathbf{x}) = f(\mathbf{x}_o) \quad + \quad \sum_{i=1}^{2}\frac{\partial f}{\partial x_i}\Big|_{x_1=x_{1o},\ x_2=x_{2o}}(x_i - x_{io})$$

$$
\begin{aligned}
&+ \quad 0.5 \sum_{i=1}^{2} \frac{\partial^2 f}{\partial x_i^2}\Big|_{x_1=x_{1o},\ x_2=x_{2o}} (x_i - x_{io})^2 \\
&+ \quad \frac{\partial^2 f}{\partial x_1 \partial x_2}\Big|_{x_1=x_{1o},\ x_2=x_{2o}} (x_1 - x_{1o})(x_2 - x_{2o}) \ + \ldots \qquad (144)
\end{aligned}
$$

The Taylor expansion (close to $\mathbf{p}_o$) of a matrix $\mathbf{A} \in \mathcal{R}^{n\times m}$, which depends on a vector $\mathbf{p} \in \mathcal{R}^r$, is

$$
\mathbf{A}(\mathbf{p}) \doteq \mathbf{A}(\mathbf{p}_o) + \sum_{i=1} \frac{1}{i!}[(\mathbf{p}-\mathbf{p}_o)^{T,[[i]]} \otimes \mathbf{I}_n]\Big[\frac{\partial^i}{\partial \mathbf{p}^i}\mathbf{A}(\mathbf{p})\Big]_{\mathbf{p}=\mathbf{p}_o} , \qquad (145)
$$

where the superscript $^{[[i]]}$ terms the ith Kronecker power

$$
(\mathbf{p}-\mathbf{p}_o)^{[[i]]} = (\mathbf{p}-\mathbf{p}_o) \otimes (\mathbf{p}-\mathbf{p}_o) \otimes \ldots \otimes (\mathbf{p}-\mathbf{p}_o) \in \mathcal{R}^{ri} \qquad (146)
$$

(*Vetter, W.J., 1970*).

7 Derivative of a Scalar-Valued Function with Respect to a Vector

The derivative of a function with respect to a vector or a matrix is a definition of a scheme, only. This abbreviation is suitable when cascades of calculations are taken into consideration.

The vector derivative operator (or gradient or nabla operator) with respect to the vector $\mathbf{p} = \mathbf{vector}[p_i]$ is termed

$$
\frac{\partial}{\partial \mathbf{p}} \triangleq \begin{pmatrix} \frac{\partial}{\partial p_1} \\ \frac{\partial}{\partial p_2} \\ \vdots \\ \frac{\partial}{\partial p_n} \end{pmatrix} \triangleq \mathbf{grad}_{\mathbf{p}} = \nabla_p = \Big(\frac{\partial}{\partial p_1} \ \ \frac{\partial}{\partial p_2} \ \ \ldots \ \ \frac{\partial}{\partial p_n}\Big)^T . \qquad (147)
$$

$$
\text{Divergence :} \quad \Big(\frac{\partial}{\partial \mathbf{p}}\Big)^T \mathbf{g} \triangleq \nabla_{\mathbf{p}} \, . \, \mathbf{g} = \sum_{i=1}^{n_p} \frac{\partial g_i}{\partial p_i} \triangleq \text{div } \mathbf{g} \, . \qquad (148)
$$

$$
\text{If } \mathbf{g} = \frac{\partial I}{\partial \mathbf{p}} \quad \text{then} \quad \text{div } \mathbf{g} = \Big(\frac{\partial}{\partial \mathbf{p}}\Big)^T \frac{\partial I}{\partial \mathbf{p}} = \nabla^2_{\mathbf{p}} I = \triangle I = \sum_{i=1}^{n_p} \frac{\partial^2 I}{\partial p_i^2} \, . \qquad (149)
$$

Derivative of an inner vector product

$$
\frac{\partial}{\partial \mathbf{p}} \ \mathbf{a}^T(\mathbf{p}) \ \mathbf{b}(\mathbf{p}) = \mathbf{vector}_i \ [\Big(\frac{\partial \mathbf{a}}{\partial p_i}\Big)^T \mathbf{b} + \Big(\frac{\partial \mathbf{b}}{\partial p_i}\Big)^T \mathbf{a}] = \Big(\frac{\partial \mathbf{a}^T}{\partial \mathbf{p}}\Big)\mathbf{b} + \Big(\frac{\partial \mathbf{b}^T}{\partial \mathbf{p}}\Big)\mathbf{a} \, . \qquad (150)
$$

For a square term in $\mathbf{p}$, we have

$$\frac{\partial}{\partial \mathbf{p}} \mathbf{p}^T \mathbf{A} \mathbf{p} = (\mathbf{A} + \mathbf{A}^T)\mathbf{p} \qquad \text{if } \mathbf{A} \text{ independent of } \mathbf{p} \ . \tag{151}$$

$$\text{If } \mathbf{Q} = \mathbf{Q}^T \text{independent of } \mathbf{p} \ , \quad \frac{\partial}{\partial \mathbf{p}} \mathbf{a}^T(\mathbf{p}) \ \mathbf{Q} \ \mathbf{a}(\mathbf{p}) = \frac{\partial}{\partial \mathbf{p}} \text{tr } \mathbf{Q}\mathbf{a}\mathbf{a}^T = 2\frac{\partial \mathbf{a}^T}{\partial \mathbf{p}} \mathbf{Q}\mathbf{a} \ . \tag{152}$$

From above, if $\mathbf{a} = \mathbf{R}\mathbf{p}$, $\mathbf{R}$ constant

$$\frac{\partial}{\partial \mathbf{p}} \mathbf{p}^T \mathbf{R}^T \mathbf{Q} \mathbf{R} \mathbf{p} = 2\mathbf{R}^T \mathbf{Q} \mathbf{R} \mathbf{p} \ . \tag{153}$$

7.1 Minimizing a Square Expression

$$\begin{aligned} \text{Resolve } \ I_r(\mathbf{x}) &= 0.5\mathbf{x}^T \mathbf{Q} \mathbf{x} + \mathbf{b}^T \mathbf{x} + a \quad \to \quad \min_{\mathbf{x}} && (154) \\ \text{gradient } \ \frac{\partial I_r}{\partial \mathbf{x}} &= \mathbf{Q}\mathbf{x} + \mathbf{b} && (155) \\ \text{Hesse matrix } \ \frac{\partial^2 I_r}{\partial \mathbf{x}^2} \triangleq \frac{\partial^2 I_r}{\partial \mathbf{x} \partial \mathbf{x}^T} &= \mathbf{Q} > 0 \ , && (156) \end{aligned}$$

i.e., a positive definite Hesse matrix is required for the minimum, see also Eq.(410).

Define $\mathbf{Q}_i \in \mathcal{R}^{i \times i}$, $i = \{1, \ldots, n\}$, all the successive northwest subdeterminants d_i (main principle minors) of the matrix $\mathbf{Q}$. Then, irrespective if n is odd or even, the Sylvester conditions for positive (negative) definiteness of $\mathbf{Q}$ are

$$\mathbf{Q} > 0 \ : \ \det \mathbf{Q}_i > 0; \qquad \mathbf{Q} < 0 \ : \ (-1)^i \det \mathbf{Q}_i > 0 \ \ \forall i \ . \tag{157}$$

Positive definiteness: The main (leading) principal minors d_i of the matrix $\mathbf{Q} \in \mathcal{R}^{n \times n}$ (where only symmetric matrices make sense in square functions $\mathbf{x}^T \mathbf{Q} \mathbf{x}$) are

$$d_i \triangleq \det \begin{pmatrix} Q_{11} & \ldots & Q_{1i} \\ \vdots & \ddots & \vdots \\ Q_{i1} & \ldots & Q_{ii} \end{pmatrix} \ , \quad 1 \le i \le n \ , \quad Q_{i1} = Q_{1i} \ , \tag{158}$$

$d_n = \det \mathbf{Q}$. For $\mathbf{Q} = \mathbf{Q}^T$ all $\lambda[\mathbf{A}]$ are real and $\mathbf{Q} > 0$ ($\mathbf{Q}$ is positive definite, i.e., $\mathbf{x}^T \mathbf{Q} \mathbf{x} > 0, \ \forall \mathbf{x} \neq \mathbf{0}$) iff $d_i > 0 \quad \forall 1 \le i \le n$.

7.2 Minimizing a Square Expression Subject to a Condition

Find the $\min_{\mathbf{a}} \mathbf{a}^T \mathbf{Q} \mathbf{a}$ s.t. the equality condition $\mathbf{V}\mathbf{a} - \mathbf{b} = \mathbf{0}$, where $\mathbf{V} \in \mathcal{R}^{m \times n}$, $n > m$ and $\mathbf{Q}^T = \mathbf{Q} > 0$. Using the vector Lagrange multipier $\boldsymbol{\lambda}$, the optimum results are $\boldsymbol{\lambda}^\star = -2(\mathbf{V}\mathbf{Q}^{-1}\mathbf{V}^T)^{-1}\mathbf{b}$, $\mathbf{a}^\star = \mathbf{Q}^{-1}\mathbf{V}^T(\mathbf{V}\mathbf{Q}^{-1}\mathbf{V}^T)^{-1}\mathbf{b}$ and the minimum of the square expression is $\mathbf{b}^T(\mathbf{V}\mathbf{Q}^{-1}\mathbf{V}^T)^{-1}\mathbf{b}$.

7.3 Optimal Reference Model

Find the controller $\mathbf{K}$ to approximate the model $\mathbf{A}_{cl,ref} \in \mathcal{R}^{n\times n}$ s.t. $k_{WK} = \|\mathbf{K}\mathbf{W}_K\|_F$ (*Weinmann, A., 2006d*). Including the weighting matrices $\mathbf{W}_A$ and $\mathbf{W}_K$,

$$\|\mathbf{W}_A(\mathbf{A}_{cl,ref} - \mathbf{A} - \mathbf{B}\mathbf{K})\|_F^2 + \lambda\|\mathbf{K}\mathbf{W}_K\|_F^2 \to \min_{\mathbf{K}} \qquad \mathbf{B},\ \mathbf{W}_K \in \mathcal{R}^{n\times m} \ . \quad (159)$$

$$\text{The result is} \quad \mathbf{K}^\star = \mathbf{L}^{-1}\mathbf{F} \ , \quad \text{where } \mathbf{L} \triangleq \lambda\mathbf{W}_K\mathbf{W}_K^T + \mathbf{B}^T\mathbf{W}_A^T\mathbf{W}_A\mathbf{B} \quad (160)$$

$$\mathbf{F} \triangleq \mathbf{B}^T\mathbf{W}_A^T\mathbf{W}_A(\mathbf{A}_{cl,ref} - \mathbf{A}) \quad (161)$$

$$k_{WK}^2 = \mathrm{tr}\{\mathbf{W}_K^T\mathbf{F}\mathbf{L}^{-T}\mathbf{L}^{-1}\mathbf{F}\mathbf{W}_K\} \rightsquigarrow \lambda \ . \quad (162)$$

8 Derivative of a Scalar-Valued Function with Respect to a Matrix

$$\frac{\partial c(\mathbf{M})}{\partial \mathbf{M}^T} = \Big(\frac{\partial c(\mathbf{M})}{\partial \mathbf{M}}\Big)^T \quad (163)$$

Product rule (since $\mathbf{H}\otimes a = a\otimes\mathbf{H} = a\mathbf{H}$)

$$\frac{\partial a(\mathbf{M})b(\mathbf{M})}{\partial \mathbf{M}} \overset{(351)}{=} a(\mathbf{M})\frac{\partial b(\mathbf{M})}{\partial \mathbf{M}} + b(\mathbf{M})\frac{\partial a(\mathbf{M})}{\partial \mathbf{M}} \ . \quad (164)$$

8.1 Eigenvalue Differential Quotients

The eigenvalues $\lambda_i[\mathbf{A}]$ and $\lambda_i[\mathbf{A}^{-1}]$ are reciprocals, the eigenvectors of the matrix and its inverse are shared. Moreover, $\lambda_i[s\mathbf{I}+\mathbf{A}] = s + \lambda_i[\mathbf{A}]$.

The complex eigenvalues $\lambda_i[\mathbf{A}]$ for real matrices $\mathbf{A} \in \mathcal{R}^{n\times n}$ always occur in conjugate pairs.

The eigenvalue derivative with respect to its generating matrix $\mathbf{A}$ (sensitivity of the eigenvalue) is (for distinct eigenvalues) (*Vetter, W.J., 1971; Brewer, J.W., 1978*)

$$\frac{\partial \lambda_i[\mathbf{A}]}{\partial \mathbf{A}} = \frac{\partial(\mathbf{a}_i^{\triangleleft H}\mathbf{A}\mathbf{a}_i)}{\partial \mathbf{A}} = \frac{\partial\ \mathrm{tr}\ [\mathbf{A}\mathbf{a}_i\mathbf{a}_i^{\triangleleft H}]}{\partial \mathbf{A}} = (\mathbf{a}_i\mathbf{a}_i^{\triangleleft H})^T = \mathbf{a}_i^{\triangleleft *}\mathbf{a}_i^T \ , \quad (165)$$

where $\mathbf{a}_i$ is the right eigenvector of $\mathbf{A}$ associated with $\lambda_i[\mathbf{A}]$ and $\mathbf{a}_i^{\triangleleft}$ is the left eigenvector. The derivation is based on simply left-multiplying $\lambda_i[\mathbf{A}]\ \mathbf{a}_i = \mathbf{A}\mathbf{a}_i$ by $\mathbf{a}_i^{\triangleleft H}$ and $\mathbf{a}_i^{\triangleleft H}\mathbf{a}_i = 1$. Although Eq.(205) requires "$\mathbf{A}$, $\mathbf{B}$ independent of $\mathbf{M}$", the influence of the eigenvector dependence on $\mathbf{A}$ is zero (*Franklin, J.N., 1968*). Alternatively,

$$\det(\lambda_i\mathbf{I} - \mathbf{A}) = \lambda_i - \det\mathbf{A} = 0 \quad (166)$$

$$\Delta\det(\lambda_i\mathbf{I} - \mathbf{A}) \overset{(476)}{=} \mathrm{tr}[\{\mathbf{adj}(\lambda_i\mathbf{I} - \mathbf{A})\}\Delta(\lambda_i\mathbf{I} - \mathbf{A})] = 0 \quad (167)$$

$$-\mathrm{tr}[\mathbf{adj}(\lambda_i \mathbf{I}-\mathbf{A})\Delta\mathbf{A}]+\mathrm{tr}[\{\mathbf{adj}(\lambda_i \mathbf{I}-\mathbf{A})\}\Delta\lambda_i] \overset{(166)}{=} 0 \tag{168}$$

$$\frac{\partial\lambda_i[\mathbf{A}]}{\partial\mathbf{A}} \overset{(450)}{=} \frac{\mathbf{adj}^T(\lambda_i \mathbf{I}_n-\mathbf{A})}{\mathrm{tr}[\mathbf{adj}\ (\lambda_i \mathbf{I}_n-\mathbf{A})]}\ . \tag{169}$$

The sensitivity of the closed-loop eigenvalues $\lambda_i[\mathbf{A}]$ with respect to any matrix $\mathbf{K}$ is

$$\frac{\partial\lambda_i[\mathbf{A}]}{\partial\mathbf{K}} = (\mathbf{I}_m \otimes \mathbf{a}_i^{\triangleleft *T})\frac{\partial\mathbf{A}}{\partial\mathbf{K}}(\mathbf{I}_n \otimes \mathbf{a}_i)\ . \tag{170}$$

For an eigenvalue at the origin $\lambda_i = 0$

$$\frac{\mathbf{adj}^T\mathbf{M}}{\mathrm{tr}\ \mathbf{adj}\mathbf{M}} = \frac{\mathbf{M}^{-T}\det\mathbf{M}}{\mathrm{tr}[\mathbf{M}^{-1}]\det\mathbf{M}} = \frac{\mathbf{M}^{-T}}{\mathrm{tr}[\mathbf{M}^{-1}]}\ . \tag{171}$$

8.2 Eigenvalue Differential Sensitivity

The coefficient matrix of a system, especially of the closed-loop control system $\mathbf{A}_{cl}$, is termed $\mathbf{F}$ and its eigenvalues are assumed distinct. By definition, the generalized right-eigenvectors of $\mathbf{F}$ are $\mathbf{f}_i$, the generalized left-eigenvectors of $\mathbf{F}$ are $\mathbf{f}_i^{\triangleleft}$. Taking the derivative of both sides of $\mathbf{F}\mathbf{f}_i = \mathbf{H}\mathbf{f}_i\lambda_i[\mathbf{F}]$ with respect to the (m,n)-matrix $\mathbf{K}$ by using the matrix product rule Eq.(351) yields

$$\frac{\partial\mathbf{F}}{\partial\mathbf{K}}(\mathbf{I}_n \otimes \mathbf{f}_i) + (\mathbf{I}_m \otimes \mathbf{F})\frac{\partial\mathbf{f}_i}{\partial\mathbf{K}} = \frac{\partial\mathbf{H}\mathbf{f}_i}{\partial\mathbf{K}}(\mathbf{I}_n \otimes \lambda_i[\mathbf{F}]) + [\mathbf{I}_m \otimes (\mathbf{H}\mathbf{f}_i)]\frac{\partial\lambda_i[\mathbf{F}]}{\partial\mathbf{K}}\ . \tag{172}$$

Taking the conjugate transpose of both sides, it results (*Feliachi, A., 1986*)

$$(\mathbf{I}_n \otimes \mathbf{f}_i^H)(\frac{\partial\mathbf{F}}{\partial\mathbf{K}})^T + \underbrace{(\frac{\partial\mathbf{f}_i^*}{\partial\mathbf{K}})^T(\mathbf{I}_m \otimes \mathbf{F}^T)}_{\mathbf{L}} = \underbrace{(\mathbf{I}_n \otimes \lambda_i^*)(\frac{\partial\mathbf{f}_i}{\partial\mathbf{K}})^T}_{\mathbf{R}} + \frac{\partial\lambda_i^*}{\partial\mathbf{K}^T}(\mathbf{I}_m \otimes \mathbf{f}_i^H)\Big| \times (\mathbf{I}_m \otimes \mathbf{f}_i^{\triangleleft})\ . \tag{173}$$

It can be shown that the terms $\mathbf{L}$ and $\mathbf{R}$ can be cancelled:

$$\begin{aligned}
\mathbf{L}(\mathbf{I}_m \otimes \mathbf{f}_i^{\triangleleft}) &= (\frac{\partial\mathbf{f}_i}{\partial\mathbf{K}})^T(\mathbf{I}_m \otimes \mathbf{F}^T)(\mathbf{I}_m \otimes \mathbf{f}_i^{\triangleleft}) = (\frac{\partial\mathbf{f}_i}{\partial\mathbf{K}})^T(\mathbf{I}_m \otimes \mathbf{F}^T\mathbf{f}_i^{\triangleleft}) \\
&= (\frac{\partial\mathbf{f}_i}{\partial\mathbf{K}})^T[\mathbf{I}_m \otimes (\mathbf{f}_i^{\triangleleft}\lambda_i^*)] = (\frac{\partial\mathbf{f}_i}{\partial\mathbf{K}})^T(\mathbf{I}_m \otimes \mathbf{f}_i^{\triangleleft})(\mathbf{I}_m \otimes \lambda_i^*) &(174)\\
&\overset{(61)}{=} (\mathbf{I}_n \otimes \lambda_i^*)(\frac{\partial\mathbf{f}_i}{\partial\mathbf{K}})^T(\mathbf{I}_m \otimes \mathbf{f}_i^{\triangleleft}) = \mathbf{R}(\mathbf{I}_m \otimes \mathbf{f}_i^{\triangleleft})\ . &(175)
\end{aligned}$$

Advantageously, the sensitivities of the eigenvectors do not influence the result, only the eigenvectors themselves. From Eq.(173) follows

$$(\mathbf{I}_n \otimes \mathbf{f}_i^{*T})(\frac{\partial\mathbf{F}}{\partial\mathbf{K}})^T(\mathbf{I}_m \otimes \mathbf{f}_i^{\triangleleft}) = \frac{\partial\lambda_i^*}{\partial\mathbf{K}^T}(\mathbf{I}_m \otimes \mathbf{f}_i^H\mathbf{H}^T)(\mathbf{I}_m \otimes \mathbf{f}_i^{\triangleleft}) = (\mathbf{I}_m \otimes \mathbf{f}^H\mathbf{H}\mathbf{f})\frac{\partial\lambda_i^*}{\partial\mathbf{K}^T} \tag{176}$$

$$\frac{\partial\lambda_i[\mathbf{F}]}{\partial\mathbf{K}} = (\mathbf{f}^H\mathbf{H}\mathbf{f})^{-1}(\mathbf{I}_m \otimes \mathbf{f}_i^{\triangleleft H})\frac{\partial\mathbf{F}}{\partial\mathbf{K}}(\mathbf{I}_n \otimes \mathbf{f}_i)\ . \tag{177}$$

8.3 Derivative of the Singular Values

Referring to Eqs.(203), (165) and the definitions in Section 8.4,

$$\frac{\partial \sigma_i}{\partial \mathbf{M}} = \frac{\partial \mathbf{u}_i^H \mathbf{M}^H \mathbf{M} \mathbf{v}_i}{\partial \mathbf{M}} = \mathbf{u}_i^* \mathbf{v}_i^T \tag{178}$$

$$\frac{\partial \sigma_i}{\partial (\Re e\ \mathbf{M})} = \Re e\ \mathbf{u}_i^* \mathbf{v}_i^T = \Re e\ \mathbf{u}_i \mathbf{v}_i^H \qquad \frac{\partial \sigma_i}{\partial (\Im m\ \mathbf{M})} = -\Im m\ \mathbf{u}_i^* \mathbf{v}_i^T = \Im m\ \mathbf{u}_i \mathbf{v}_i^H \tag{179}$$

$$\frac{\partial (\sigma_i^2)}{\partial (\Re e\ \mathbf{M})} = -(\Re e\ \mathbf{M})(\mathbf{\Lambda}_1^T + \mathbf{\Lambda}_1) - (\Im m\ \mathbf{M})(\mathbf{\Lambda}_2^T - \mathbf{\Lambda}_2) \tag{180}$$

$$\frac{\partial (\sigma_i^2)}{\partial (\Im m\ \mathbf{M})} = -(\Im m\ \mathbf{M})(\mathbf{\Lambda}_1^T + \mathbf{\Lambda}_1) + (\Re e\ \mathbf{M})(\mathbf{\Lambda}_2^T - \mathbf{\Lambda}_2)\ , \tag{181}$$

$$\text{where}\quad \mathbf{\Lambda}_1 = -\frac{\partial \lambda[\mathbf{M}^H \mathbf{M}]}{\partial (\Re e\ \mathbf{M}^H \mathbf{M})} = -\Re e\ \mathbf{v}_i^* \mathbf{v}_i^T, \qquad \mathbf{\Lambda}_2 = -\frac{\partial \lambda[\mathbf{M}^H \mathbf{M}]}{\partial (\Im m\ \mathbf{M}^H \mathbf{M})} = \Im m\ \mathbf{v}_i^* \mathbf{v}_i^T \tag{182}$$

(*Sevaston, G.E., and Longman, R.W., 1988; Freudenberg, J.S., et al. 1982*).

8.4 Derivative of the Spectral Condition Number

From

$$\kappa_s[\mathbf{M}] = \frac{\sigma_{\min}[\mathbf{M}]}{\sigma_{\min}[\mathbf{M}]} \quad \text{derive} \quad \Delta\kappa_s[\mathbf{M}] = \frac{\sigma_{\min}\Delta\sigma_{\max} - \sigma_{\max}\Delta\sigma_{\min}}{\sigma_{\min}^2}\ . \tag{183}$$

For $\partial\sigma_{\max}/\partial\mathbf{M} \triangleq \mathbf{u}_a \mathbf{v}_a^T$ and $\partial\sigma_{\min}/\partial\mathbf{M} \triangleq \mathbf{u}_i \mathbf{v}_i^T$, where $\mathbf{v}_i, \mathbf{v}_a, \mathbf{u}_i, \mathbf{u}_a$ are the real eigenvectors of $\mathbf{M}^H\mathbf{M}$ and $\mathbf{M}\mathbf{M}^H$ associated with $\sigma_{\min}$ (subscript i) and $\sigma_{\max}$ (subscript a), respectively, one has

$$\Delta\sigma_{\max} = \text{tr}\ [\Delta\sigma_{\max}] = \text{tr}\ [\mathbf{v}_a \mathbf{u}_a^T \Delta\mathbf{M}] \tag{184}$$

and

$$\Delta\sigma_{\min} = \text{tr}\ [\mathbf{v}_i \mathbf{u}_i^T \Delta\mathbf{M}], \tag{185}$$

$$\Delta\kappa_s[\mathbf{M}] = \frac{\text{tr}\ [(\sigma_{\min}[\mathbf{M}]\mathbf{v}_a\mathbf{u}_a^T - \sigma_{\max}[\mathbf{M}]\mathbf{v}_i\mathbf{u}_i^T)\Delta\mathbf{M}]}{\sigma_{\min}^2[\mathbf{M}]}, \tag{186}$$

$$\frac{\partial\kappa_s[\mathbf{M}]}{\partial\mathbf{M}} = \frac{\sigma_{\min}[\mathbf{M}]\mathbf{u}_a\mathbf{v}_a^T - \sigma_{\max}[\mathbf{M}]\mathbf{u}_i\mathbf{v}_i^T}{\sigma_{\min}^2[\mathbf{M}]}\ . \tag{187}$$

8.5 Differential Quotients of Determinant Expressions

Some determinant derivative properties are (*Athans, M., 1968; Vetter, W.J., 1971*)

$$\frac{\partial}{\partial \mathbf{M}} \det \mathbf{M} = (\mathbf{M}^{-1})^T \det \mathbf{M} = \mathbf{adj}^T(\mathbf{M}), \tag{188}$$

$$\frac{\partial}{\partial \mathbf{M}} \det (\mathbf{M}^n) = n(\mathbf{M}^{-1})^T (\det \mathbf{M})^n, \tag{189}$$

$$\frac{\partial}{\partial \mathbf{M}} \det e^{\mathbf{M}} = \det e^{\mathbf{M}}, \tag{190}$$

$$\frac{\partial}{\partial \mathbf{M}} \log \det \mathbf{M} = (\mathbf{M}^{-1})^T . \tag{191}$$

If $\mathbf{V}$ is independent of $\mathbf{M}$, $\quad \frac{\partial}{\partial \mathbf{M}} \det (\mathbf{M}\mathbf{V}\mathbf{M}^T) = 2(\mathbf{M}\mathbf{V}\mathbf{M}^T)^{-1}\mathbf{M}\mathbf{V}$. (192)

If $\mathbf{A}$ and $\mathbf{B}$ are square and independent of $\mathbf{M}$, and $\mathbf{M}$ is square, then

$$\frac{\partial}{\partial \mathbf{M}} \det (\mathbf{A}\mathbf{M}\mathbf{B}) = (\mathbf{M}^{-1})^T \det (\mathbf{A}\mathbf{M}\mathbf{B}) , \tag{193}$$

$$\frac{\partial \det[\mathbf{A}(\mathbf{H}+\mathbf{M})\mathbf{B}]}{\partial \mathbf{M}} = (\mathbf{H}+\mathbf{M})^{-T} \det[\mathbf{A}(\mathbf{H}+\mathbf{M})\mathbf{B}] . \tag{194}$$

Mostly used for controller coefficient matrix $\mathbf{M} = \mathbf{K}$, if $\mathbf{A}$ and $\mathbf{B}$ and $\mathbf{C}$ are independent of $\mathbf{M} = \mathbf{K}$, one finds, see Eq.(11), (12),

$$\frac{\partial}{\partial \mathbf{K}} \det(s\mathbf{I}_n - \mathbf{A} - \mathbf{B}\mathbf{K}\mathbf{C}) = -\mathbf{B}^T[\mathbf{adj}(s\mathbf{I}_n - \mathbf{A} - \mathbf{B}\mathbf{K}\mathbf{C})]^T \mathbf{C}^T \tag{195}$$

$$\frac{\partial \det(\mathbf{B}\mathbf{K}_y\mathbf{C})}{\partial \mathbf{K}_y} = [\mathbf{C} \cdot \mathbf{adj}(\mathbf{B}\mathbf{K}_y\mathbf{C}) \cdot \mathbf{B}]^T . \tag{196}$$

Example (8.5). Lyapunov subdeterminants:

From the Lyapunov equation $\quad \mathbf{A}_{cl}^T\mathbf{P}_n + \mathbf{P}_n\mathbf{A}_{cl} = -\mathbf{Q}$, $\quad \mathbf{A}_{cl} \triangleq \mathbf{A} + \mathbf{B}\mathbf{K}$, (197)

$$\operatorname{col} \Delta\mathbf{P}_n = \mathbf{H}_n \operatorname{col} \Delta\mathbf{K} \qquad \begin{cases} \mathbf{H}_n \triangleq -\mathbf{H}_P^{-1}\{[(\mathbf{P}_n\mathbf{B}) \otimes \mathbf{I}_n]\mathbf{U}_{m,n} + [\mathbf{I}_n \otimes (\mathbf{P}_n\mathbf{B})]\} \\ \mathbf{H}_P \triangleq [\mathbf{I}_n \otimes (\mathbf{A}+\mathbf{B}\mathbf{K})^T + (\mathbf{A}+\mathbf{B}\mathbf{K})^T \otimes \mathbf{I}_n] \end{cases} \tag{198}$$

where $\mathbf{K} \in \mathcal{R}^{m\times n}$, $\mathbf{B} \in \mathcal{R}^{n\times m}$, $\mathbf{H}_P \in \mathcal{R}^{n^2\times n^2}$, $\mathbf{U}_{n,m} \in \mathcal{R}^{mn\times mn}$

$$\frac{\partial \det \mathbf{P}_n}{\partial \mathbf{K}} = [\mathbf{I}_m \otimes (\mathbf{h}_n^T\mathbf{U}_{n,m})][(\operatorname{col} \mathbf{I}_m) \otimes \mathbf{I}_n] , \quad \text{where} \quad \mathbf{h}_n^T \triangleq (\det \mathbf{P}_n)[\operatorname{col} (\mathbf{P}_n^{-T})]^T \mathbf{H}_n . \tag{199}$$

For the $n-1$ northwest main minor of $\mathbf{P}_n$ and a reduction matrix $\mathbf{R}_n$ (*Weinmann, A., 2008d*)

$$\mathbf{P}_{n-i} = \mathbf{P}([1:n-i],[1:n-i]) \ \ \forall\, i = \{1, n-1\} \tag{200}$$

$$\frac{\partial \det \mathbf{P}_{n-i}}{\partial \mathbf{K}} = [\mathbf{I}_m \otimes (\mathbf{h}_{n-i}^T\mathbf{U}_{n,m})][(\operatorname{col} \mathbf{I}_m) \otimes \mathbf{I}_n] \tag{201}$$

$$\mathbf{h}_{n-i}^T \triangleq \{\operatorname{col}\ [\operatorname{adj}\mathbf{P}_{n-i}]^T\}^T \prod_{\nu=1}^{i} \mathbf{R}_{n-\nu+1}\mathbf{H}_n \in \mathcal{R}^{1\times n^2},\ \mathbf{R}_n \in \mathcal{R}^{(n-1)^2\times n^2} . \quad \square \tag{202}$$

8.6 Trace and Frobenius Norm Gradients

Use the cyclic property for a matrix product in a trace argument, Eq.(14), to modify the formulas in the following.

Some trace derivative properties are (*Athans, M., 1968*; *Vetter, W.J., 1971*; *Brewer, J.W., 1978; Weinmann, A., 1991*)

$$\frac{\partial}{\partial \mathbf{M}} \mathbf{a}^T \mathbf{M} \mathbf{b} = \mathbf{b}\mathbf{a}^T \qquad \mathbf{M} \in \mathcal{R}^{m\times m}, \quad \mathbf{a}, \mathbf{b} \in \mathcal{R}^m \ . \tag{203}$$

$$\text{If } \mathbf{B} \text{ and } \mathbf{C} \text{ independent of } \mathbf{M} \qquad \frac{\partial}{\partial \mathbf{M}} \mathrm{tr}[\mathbf{B}\mathbf{M}^T\mathbf{C}] = \mathbf{C}\mathbf{B}. \tag{204}$$

$$\text{If } \mathbf{A}, \mathbf{B} \text{ are independent of } \mathbf{M}\ , \qquad \frac{\partial}{\partial \mathbf{M}} \mathrm{tr}\ [\mathbf{A}\mathbf{M}\mathbf{B}] = \mathbf{A}^T\mathbf{B}^T (= \mathbf{A}^H\mathbf{B}^H). \tag{205}$$

$$\text{If } \mathbf{a}, \mathbf{B}, \mathbf{b} \text{ are independent of } \mathbf{M} \quad \frac{\partial}{\partial \mathbf{M}} (\mathbf{a}^T\mathbf{M}\mathbf{B}\mathbf{M}^T\mathbf{b}) = \mathbf{a}\mathbf{b}^T\mathbf{M}\mathbf{B}^T + \mathbf{b}\mathbf{a}^T\mathbf{M}\mathbf{B}\ . \tag{206}$$

$$\text{If } \mathbf{a} \text{ is independent of } \mathbf{M} \qquad \frac{\partial}{\partial \mathbf{M}} \mathbf{a}^T\mathbf{M}^{-1}\mathbf{a} = -(\mathbf{M}^{-1})^T\mathbf{a}\mathbf{a}^T(\mathbf{M}^{-1})^T. \tag{207}$$

$$n \geq 2: \quad \frac{\partial}{\partial \mathbf{M}} \mathrm{tr}\ [\mathbf{M}^n] = n(\mathbf{M}^{n-1})^T, \tag{208}$$

$$\frac{\partial \mathrm{tr}[\mathbf{M}]}{\partial \mathbf{M}} = \mathbf{I}\ , \tag{209}$$

$$\frac{\partial ||\mathbf{M}||_F^2}{\partial M_{ij}} = \mathrm{tr}[\mathbf{E}_{ij}^T\mathbf{M} + \mathbf{M}^T\mathbf{E}_{ij}] \stackrel{(17)}{=} 2\, M_{ij} \tag{210}$$

$$\frac{\partial ||\mathbf{M}||_F^2}{\partial \mathbf{M}} = \frac{\partial}{\partial \mathbf{M}} \mathrm{tr}[\mathbf{M}\mathbf{M}^T] = 2\,\mathbf{M}, \tag{211}$$

$$\frac{\partial}{\partial \mathbf{M}} ||\mathbf{M}||_F = \frac{\partial}{\partial \mathbf{M}} \sqrt{\mathrm{tr}[\mathbf{M}\mathbf{M}^T]} = \frac{\mathbf{M}}{||\mathbf{M}||_F}\ , \tag{212}$$

$$\frac{\partial}{\partial \mathbf{M}} \mathrm{tr}\ [e^{\mathbf{M}}] = e^{(\mathbf{M}^T)} = (e^{\mathbf{M}})^T\ . \tag{213}$$

Derivative of the $(2p)$-th power of the norm:

$$\frac{\partial}{\partial \mathbf{K}} ||\mathbf{K}||_F^{2p} \ = \ \frac{\partial}{\partial \mathbf{K}} (||\mathbf{K}||_F^2)^p \tag{214}$$

$$= \ \left[p(||\mathbf{K}||_F^2)^{p-1}\right] \ \cdot \ \frac{\partial}{\partial \mathbf{K}} ||\mathbf{K}||_F^2 = 2p||\mathbf{K}||_F^{2(p-1)}\ \mathbf{K}\ , \tag{215}$$

that is, the power $2p$ only entails a scalar factor $p||\mathbf{K}||_F^{2(p-1)}$.

$$\Delta||\mathbf{M}^{-1}||_F^2 \ = \ ||(\mathbf{M}+\Delta\mathbf{M})^{-1}||_F^2 - ||\mathbf{M}^{-1}||_F^2 \stackrel{(93)}{=} ||\mathbf{M}^{-1} - \mathbf{M}^{-1}\Delta\mathbf{M}\cdot\mathbf{M}^{-1}||_F^2 - ||\mathbf{M}^{-1}||_F^2$$

$$\Delta||\mathbf{M}^{-1}||_F^2 \ = \ -2\ \mathrm{tr}[\Delta\mathbf{M}^T\mathbf{M}^{-T}\mathbf{M}^{-1}\mathbf{M}^{-T}] \tag{216}$$

$$\frac{\partial ||\mathbf{M}^{-1}||_F}{\partial \mathbf{M}} \stackrel{(450)}{=} -2\ \mathbf{M}^{-T}\mathbf{M}^{-1}\mathbf{M}^{-T} \tag{217}$$

$$\frac{\partial \|\mathbf{A}\mathbf{M}^{-1}\mathbf{B}\|_F^2}{\partial \mathbf{M}} = -2\ \mathbf{M}^{-T}\mathbf{A}^T\mathbf{A}\ \mathbf{M}^{-1}\ \mathbf{B}\mathbf{B}^T\mathbf{M}^{-T} \tag{218}$$

$$\frac{\partial \|\mathbf{a}+\mathbf{R}\mathbf{k}\|_F^2}{\partial \mathbf{k}} = 2\ \mathbf{R}^T(\mathbf{a}+\mathbf{R}\mathbf{k}) \tag{219}$$

If $\mathbf{A}$ independent of $\mathbf{M}$
$$\frac{\partial}{\partial \mathbf{M}}\text{tr}\ [\mathbf{A}\mathbf{M}^m] = [\sum_{i=0}^{m-1} \mathbf{M}^i\mathbf{A}\mathbf{M}^{m-i-1}\ ]^T, \tag{220}$$

if $\mathbf{A}$ and $\mathbf{B}$ independent of $\mathbf{A}$
$$\frac{\partial}{\partial \mathbf{M}}\text{tr}\ [\mathbf{A}\mathbf{M}\mathbf{B}\mathbf{M}] = \mathbf{A}^T\mathbf{M}^T\mathbf{B}^T + \mathbf{B}^T\mathbf{M}^T\mathbf{A}^T, \tag{221}$$

$$\frac{\partial}{\partial \mathbf{M}}\text{tr}\ [\mathbf{A}\mathbf{M}\mathbf{B}\mathbf{M}^T] = \frac{\partial}{\partial \mathbf{M}}\text{tr}\ [\mathbf{B}\mathbf{M}^T\mathbf{A}\mathbf{M}] = \mathbf{A}^T\mathbf{M}\mathbf{B}^T + \mathbf{A}\mathbf{M}\mathbf{B}, \tag{222}$$

$$\frac{\partial}{\partial \mathbf{M}}\text{tr}\ [\mathbf{A}\mathbf{M}^{-1}\mathbf{B}] = -(\mathbf{M}^{-1}\mathbf{B}\mathbf{A}\mathbf{M}^{-1})^T\ . \tag{223}$$

For some $\Delta\omega_0 \triangleq \text{tr}[\mathbf{B}\ \Delta\mathbf{A}]$

$$\frac{\Delta\omega_0}{\Delta\mathbf{A}} \doteq \frac{\partial\omega_0}{\partial\Delta\mathbf{A}} = \mathbf{B}^T \quad \rightsquigarrow \quad \Delta\omega_0 = \text{tr}[(\frac{\partial\omega_0}{\partial\mathbf{A}})^T\Delta\mathbf{A}]\ . \tag{224}$$

$$\frac{\partial \|\mathbf{T}\|_F^2}{\partial \mathbf{K}} = 2\ \mathbf{A}^T\mathbf{A}\ \mathbf{K}\ \mathbf{D}\,\mathbf{D}^T\ ,\ \text{where}\ \ \mathbf{T} \triangleq \mathbf{A}\ \mathbf{K}\ \mathbf{D}\ . \tag{225}$$

For holistic controllers one finds for the trace of the inverse of the closed-loop coefficient matrix (*Weinmann, A., 2001a*)

$$\frac{\partial}{\partial \mathbf{K}}\text{tr}\ [(\mathbf{A}+\mathbf{B}\mathbf{K})^{-1}] = -\mathbf{B}^T(\mathbf{A}+\mathbf{B}\mathbf{K})^{-2T}\ . \tag{226}$$

Furthermore,

$$\frac{\partial}{\partial \mathbf{M}}\|\mathbf{A}+\mathbf{B}\mathbf{M}\mathbf{C}\|_F^2 = \frac{\partial}{\partial \mathbf{M}}\text{tr}[(\mathbf{A}+\mathbf{B}\mathbf{M}\mathbf{C})^T(\mathbf{A}+\mathbf{B}\mathbf{M}\mathbf{C})] = 2\ \mathbf{B}^T(\mathbf{A}+\mathbf{B}\mathbf{M}\mathbf{C})\mathbf{C}^T. \tag{227}$$

Example (8.6;1) :

$$\|\mathbf{K}\mathbf{B}-\mathbf{B}\mathbf{A}\|_F \to \min_{\mathbf{K}} \quad \rightsquigarrow \quad \mathbf{K} = \mathbf{B}\mathbf{A}\mathbf{B}^T(\mathbf{B}\mathbf{B}^T)^{-1}\ .\ \square \tag{228}$$

$$\frac{\partial}{\partial \mathbf{M}}\text{tr}[(\mathbf{A}+\mathbf{B}\mathbf{M}\mathbf{C})^T(\mathbf{A}+\mathbf{B}\mathbf{M}\mathbf{C})\mathbf{U}] = \mathbf{B}^T(\mathbf{A}+\mathbf{B}\mathbf{M}\mathbf{C})(\mathbf{U}+\mathbf{U}^T)\mathbf{C}^T\ . \tag{229}$$

$$\frac{\partial}{\partial \mathbf{A}}\ \text{tr}[\mathbf{A}^T\mathbf{B}(\mathbf{A})] = \mathbf{B}(\mathbf{A}) + \mathbf{matrix}_{jk}\{\text{tr}[\mathbf{A}^T\frac{\partial \mathbf{B}(\mathbf{A})}{\partial A_{jk}}]\}\ . \tag{230}$$

Two equivalent derivations of the trace with respect to a matrix. First

$$\frac{\partial I}{\partial K_{ij}} = \mathbf{x}_o^T\mathbf{M}^T\mathbf{E}_{ij}\mathbf{x}_o = \mathrm{tr}[\mathbf{x}_o\mathbf{x}_o^T\mathbf{M}^T\mathbf{E}_{ij}] = (\mathbf{M}\mathbf{x}_o\mathbf{x}_o^T)_{ij} \rightsquigarrow \frac{\partial I}{\partial \mathbf{K}} = \mathbf{M}\mathbf{x}_o\mathbf{x}_o^T. \quad (231)$$

Second, using $\mathbf{E}_{ij} = \mathbf{e}_i\mathbf{e}_j^T$,

$$\frac{\partial I}{\partial K_{ij}} = \mathbf{x}_o^T\mathbf{M}^T\mathbf{e}_i\mathbf{e}_j^T\mathbf{x}_o = \mathbf{e}_j^T\mathbf{x}_o\mathbf{x}_o^T\mathbf{M}^T\mathbf{e}_i = \mathbf{e}_i^T\mathbf{M}\mathbf{x}_o\mathbf{x}_o^T\mathbf{e}_j \quad (232)$$

yields the same result.

Based on Eq.(143),

$$\begin{aligned} \frac{\partial \mathrm{tr}[\mathbf{A}^{-1}]}{\partial A_{ij}} &= -\mathrm{tr}[\mathbf{A}^{-2}\mathbf{E}_{ij}] \overset{(17)}{=} -(\mathbf{A}^{-2})_{ji} && (233) \\ \frac{\partial \mathrm{tr}[\mathbf{A}^{-1}]}{\partial \mathbf{A}} &= -(\mathbf{A}^{-2})^T && (234) \\ \Re e \, \frac{\partial \mathrm{tr}[(\mathbf{C}+\Delta\mathbf{A})^{-1}]}{\partial \Delta\mathbf{A}} &= -\Re e \, (\mathbf{C}^{-2})^T, \quad \mathbf{C} \in \mathcal{C}^{n\times n} \; . && (235) \end{aligned}$$

$$\frac{\partial \sum_i 1/\lambda_i[\mathbf{A}_{cl}]}{\partial \mathbf{A}} = \frac{\partial \mathrm{tr}[\mathbf{A}_{cl}^{-1}]}{\partial \mathbf{A}} \overset{(226)}{=} -\mathbf{B}^T(\mathbf{A}+\mathbf{B}\mathbf{K})^{-2T} \quad \text{since} \quad \lambda[\mathbf{A}^{-1}] = (\lambda[\mathbf{A}])^{-1} \; . \quad (236)$$

$$\text{For} \quad \mathbf{A}_{cl} \overset{\triangle}{=} \mathbf{A} + \mathbf{B}\mathbf{K}_y\mathbf{C} \quad (237)$$

$$\frac{\partial}{\partial \mathbf{K}_y}\mathrm{tr}[\mathbf{A}_{cl}^{-1}] = -[\mathbf{C}(\mathbf{A}_{cl})^{-2}\mathbf{B}]^T \quad (238)$$

$$\frac{\partial}{\partial \mathbf{K}_y}\frac{1/\det \mathbf{A}_{cl}}{\mathrm{tr}[\mathbf{A}_{cl}^{-1}]} = -\frac{[\mathbf{C}\{\mathbf{A}_{cl}^{-2} + \mathbf{A}_{cl}^{-1}\mathrm{tr}[\mathbf{A}_{cl}^{-1}]\}\mathbf{B}]^T}{(\mathrm{tr}[\mathbf{A}_{cl}^{-1}])^2 \det \mathbf{A}_{cl}} \quad (239)$$

$$\frac{\partial}{\partial \mathbf{K}_y}\frac{1/\det \mathbf{A}_{cl}}{\mathrm{tr}[\mathbf{A}_{cl}]} = -\frac{[\mathbf{C}\{\mathbf{I}_n + \mathbf{A}_{cl}^{-1}\mathrm{tr}[\mathbf{A}_{cl}]\}\mathbf{B}]^T}{(\mathrm{tr}[\mathbf{A}_{cl}])^2 \det \mathbf{A}_{cl}} \quad (240)$$

$$\frac{\partial}{\partial \mathbf{K}_y}\frac{1}{\mathrm{tr}[\mathbf{A}_{cl}]} = -\frac{[\mathbf{C}\mathbf{B}]^T}{(\mathrm{tr}[\mathbf{A}_{cl}])^2} \quad (241)$$

(*Weinmann, A., 2005*). Derivative of the power of the trace:

$$\begin{aligned} \frac{\partial[-\{\mathrm{tr}[\mathbf{A}_{cl}^{-1}]\}^p]}{\partial \mathbf{K}_y} &\overset{(394)}{=} -p\Big(\mathrm{tr}[\mathbf{A}_{cl}^{-1}]\Big)^{p-1}\frac{\partial \mathrm{tr}[\mathbf{A}_{cl}^{-1}]}{\partial \mathbf{K}_y} \\ &\overset{(238)}{=} p\Big(\mathrm{tr}[\mathbf{A}_{cl}^{-1}]\Big)^{p-1}[\mathbf{C}\mathbf{A}_{cl}^{-2}\mathbf{B}]^T \; . \quad (242) \end{aligned}$$

$$\mathbf{M},\ \Delta\mathbf{M} \in \mathcal{R}^{n\times n} \qquad \frac{\partial \mathrm{tr}[\mathbf{I}_n \otimes \mathbf{M}]}{\partial M_{ij}} = n\delta_{ij} \; , \quad \frac{\partial \mathrm{tr}[\mathbf{I}_n \otimes \mathbf{M}]}{\partial \mathbf{M}} = n\mathbf{I}_n \quad (243)$$

$$\mathrm{tr}[\mathbf{M}^T \otimes \Delta\mathbf{M}] = \mathrm{tr}[\Delta\mathbf{M} \otimes \mathbf{M}^T] = (\mathrm{tr}\mathbf{M})(\mathrm{tr}\Delta\mathbf{M}) \tag{244}$$

$$\frac{\partial \mathrm{tr}[\mathbf{M}^T \otimes \Delta\mathbf{M}]}{\partial \Delta\mathbf{M}} = (\mathrm{tr}\mathbf{M})\ \mathbf{I}_n \tag{245}$$

$$\mathrm{tr}[\mathbf{I}_n \otimes (\mathbf{M} \cdot \Delta\mathbf{M}^T)] = \mathrm{tr}[\mathbf{I}_n \otimes (\Delta\mathbf{M} \cdot \mathbf{M}^T)] = n \cdot \mathrm{tr}[\mathbf{M} \cdot \Delta\mathbf{M}^T] \tag{246}$$

$$\frac{\partial \mathrm{tr}[\mathbf{I}_n \otimes (\mathbf{M} \cdot \Delta\mathbf{M}^T)]}{\partial \Delta\mathbf{M}} = n\mathbf{M}\ . \tag{247}$$

$$\text{For} \quad \check{\mathbf{M}} \triangleq \mathbf{I}_n \otimes \mathbf{M} + \mathbf{M} \otimes \mathbf{I}_n = \mathbf{M} \oplus \mathbf{M}, \tag{248}$$

$$\frac{\partial \|\check{\mathbf{M}}\|_F^2}{\partial \mathbf{M}} = \frac{\partial}{\partial \mathbf{M}} \|\mathbf{I}_n \otimes \mathbf{M} + \mathbf{M} \otimes \mathbf{I}_n\|_F^2 = 4(n\mathbf{M} + \mathbf{I}_n \mathrm{tr}\mathbf{M}) \tag{249}$$

$$\frac{\partial \|\check{\mathbf{M}}\|_F^2}{\partial M_{ij}} = \frac{\partial}{\partial M_{ij}} \|\mathbf{I}_n \otimes \mathbf{M} + \mathbf{M} \otimes \mathbf{I}_n\|_F^2 = 4(nM_{ij} + \delta_{ij} \mathrm{tr}\mathbf{M})\ . \tag{250}$$

For $\mathbf{G} = \mathbf{G}(\mathbf{x})$

$$\frac{\partial \mathbf{x}^T\mathbf{G}^T\mathbf{G}\mathbf{x}}{\partial \mathbf{x}} = \frac{\partial \mathbf{x}^T\mathbf{G}^T}{\partial \mathbf{x}}[\mathbf{I}_s \otimes (\mathbf{G}\mathbf{x})] + [\mathbf{I}_r \otimes (\mathbf{x}^T\mathbf{G}^T)]\frac{\partial \mathbf{G}\mathbf{x}}{\partial \mathbf{x}} \tag{251}$$

$$\frac{\partial \mathbf{x}^T\mathbf{G}^T\mathbf{G}\mathbf{x}}{\partial \mathbf{x}} \overset{(251)}{=} \mathbf{G}^T\mathbf{G}\mathbf{x} + (\mathbf{I}_n \otimes \mathbf{x}^T)\frac{\partial \mathbf{G}^T}{\partial \mathbf{x}}\mathbf{G}\mathbf{x} + [\mathbf{I}_n \otimes (\mathbf{x}^T\mathbf{G}^T)][\frac{\partial \mathbf{G}}{\partial \mathbf{x}}\mathbf{x} + (\mathbf{I}_n \otimes \mathbf{G})\mathrm{col}\,\mathbf{I}_n]\ . \tag{252}$$

Its zeros result from

$$\frac{\partial}{\partial \mathbf{x}}\frac{\mathbf{x}^T\mathbf{G}^T\mathbf{G}\mathbf{x}}{\mathbf{x}^T\mathbf{x}} = \mathbf{0} \overset{(164)(394)}{=} \frac{1}{\mathbf{x}^T\mathbf{x}}\frac{\partial \mathbf{x}^T\mathbf{G}^T\mathbf{G}\mathbf{x}}{\partial \mathbf{x}} - \frac{1}{(\mathbf{x}^T\mathbf{x})^2}\frac{\partial \mathbf{x}^T\mathbf{x}}{\partial \mathbf{x}}\mathbf{x}^T\mathbf{G}^T\mathbf{G}\mathbf{x} = \mathbf{0} \tag{253}$$

$$\Big\{2\mathbf{G}^T(\mathbf{x})\mathbf{G}(\mathbf{x}) + (\mathbf{I}_n \otimes \mathbf{x}^T)\frac{\partial \mathbf{G}^T(\mathbf{x})}{\partial \mathbf{x}}\mathbf{G}(\mathbf{x}) + [\mathbf{I}_n \otimes (\mathbf{x}^T\mathbf{G}^T(\mathbf{x}))]\frac{\partial \mathbf{G}(\mathbf{x})}{\partial \mathbf{x}}\Big\}\mathbf{x} - 2\mathbf{x}\frac{\mathbf{x}^T\mathbf{G}^T(\mathbf{x})\mathbf{G}(\mathbf{x})\mathbf{x}}{\mathbf{x}^T\mathbf{x}} \overset{(262)}{=} \mathbf{0}\ . \tag{254}$$

This result follows from the following derivation. Using an abbreviation for the partial differential quotient, for $\mathbf{x} \in \mathcal{R}^n$ (and $r = n,\ s = 1$)

$$\frac{\partial \mathbf{G}(\mathbf{x})}{\partial \mathbf{x}} \triangleq \mathbf{G}_x \tag{255}$$

we find, first, employing Eq.(292),

$$\frac{\partial \mathbf{G}(\mathbf{x}) \cdot \mathbf{x}}{\partial \mathbf{x}} \triangleq (\mathbf{G}\mathbf{x})_x \overset{(351)}{=} \mathbf{G}_x(\mathbf{I}_s \otimes \mathbf{x}) + (\mathbf{I}_r \otimes \mathbf{G})\mathbf{x}_x = \mathbf{G}_x\mathbf{x} + (\mathbf{I}_n \otimes \mathbf{G})\mathrm{col}\ \mathbf{I}_n \tag{256}$$

$$\frac{\partial \mathbf{x}^T\mathbf{G}^T}{\partial \mathbf{x}} \triangleq (\mathbf{x}^T\mathbf{G}^T)_x = (\mathbf{x}^T)_x(\mathbf{I}_s \otimes \mathbf{G}^T) + [\mathbf{I}_r \otimes \mathbf{x}^T](\mathbf{G}^T)_x = \mathbf{G}^T + [\mathbf{I}_r \otimes \mathbf{x}^T]\mathbf{G}_x^T \tag{257}$$

and, second,

$$
\begin{aligned}
(\mathbf{x}^T\mathbf{G}^T\mathbf{G}\mathbf{x})_x &\overset{(351)}{=} (\mathbf{x}^T\mathbf{G}^T)_x[\mathbf{I}_s \otimes (\mathbf{G}\mathbf{x})] + [\mathbf{I}_r \otimes (\mathbf{x}^T\mathbf{G}^T)](\mathbf{G}\mathbf{x})_x && (258)\\
&\overset{(256),(257)}{=} [\mathbf{G}^T + (\mathbf{I}_n \otimes \mathbf{x}^T)\mathbf{G}_x^T]\mathbf{G}\mathbf{x} + [\mathbf{I}_n \otimes (\mathbf{x}^T\mathbf{G}^T)][\mathbf{G}_x\mathbf{x} + (\mathbf{I}_n \otimes \mathbf{G})\mathbf{x}_x] && (259)\\
&= \mathbf{G}^T\mathbf{G}\mathbf{x} + (\mathbf{I}_n \otimes \mathbf{x}^T)\mathbf{G}_x^T\mathbf{G}\mathbf{x} + [\mathbf{I}_n \otimes (\mathbf{x}^T\mathbf{G}^T)]\mathbf{G}_x\mathbf{x} + [\mathbf{I}_n \otimes (\mathbf{x}^T\mathbf{G}^T)](\mathbf{I}_n \otimes \mathbf{G})\mathbf{x}_x\\
&\overset{(43)}{=} \mathbf{G}^T\mathbf{G}\mathbf{x} + (\mathbf{I}_n \otimes \mathbf{x}^T)\mathbf{G}_x^T\mathbf{G}\mathbf{x} + [\mathbf{I}_n \otimes (\mathbf{x}^T\mathbf{G}^T)]\mathbf{G}_x\mathbf{x} + [\mathbf{I}_n \otimes (\mathbf{x}^T\mathbf{G}^T\mathbf{G})]\mathbf{x}_x && (260)\\
&\overset{(63)}{=} \mathbf{G}^T\mathbf{G}\mathbf{x} + (\mathbf{I}_n \otimes \mathbf{x}^T)\mathbf{G}_x^T\mathbf{G}\mathbf{x} + [\mathbf{I}_n \otimes (\mathbf{x}^T\mathbf{G}^T)]\mathbf{G}_x\mathbf{x} + \mathbf{G}^T\mathbf{G}\ \mathbf{x} && (261)\\
&= 2\ \mathbf{G}^T\mathbf{G}\mathbf{x} + (\mathbf{I}_n \otimes \mathbf{x}^T)\mathbf{G}_x^T\mathbf{G}\mathbf{x} + [\mathbf{I}_n \otimes (\mathbf{x}^T\mathbf{G}^T)]\mathbf{G}_x\mathbf{x}\ , \quad \text{where} && (262)
\end{aligned}
$$

$$[\mathbf{I}_n \otimes \mathbf{x}^T\mathbf{G}^T)](\mathbf{I}_n \otimes \mathbf{G})\mathbf{x}_x = [\mathbf{I}_n \otimes (\mathbf{x}^T\mathbf{G}^T\mathbf{G})]\mathbf{x}_x = [\mathbf{I}_n \otimes (\mathbf{x}^T\mathbf{G}^T\mathbf{G})]\mathrm{col}\ \mathbf{I}_n = \overset{(55)}{=} \mathbf{G}\mathbf{G}^T\mathbf{x}\ . \quad (263)$$

For comparison,

$$
\begin{aligned}
R[\mathbf{A}] &\triangleq \frac{\mathbf{x}^H\mathbf{A}\mathbf{x}}{\mathbf{x}^H\mathbf{x}} \overset{(\mathbf{x}=\mathbf{a}_i)}{=} \lambda_i \quad \text{Rayleigh quotient} && (264)\\
\max_i \frac{\mathbf{x}^H\mathbf{A}\mathbf{x}}{\mathbf{x}^H\mathbf{x}} &= \max_i \lambda_i \quad \text{Rayleigh principle.} && (265)
\end{aligned}
$$

Example (8.6;2) Minimizing the regression error: Consider the measurement vector $\mathbf{y}$ and the mathematical model $\mathbf{M}$ of the process, then

$$I = \varepsilon^T\mathbf{W}\varepsilon \quad \to \quad \min_{\mathbf{k}} \quad \text{subject to} \quad \mathbf{k}^T\mathbf{V}\mathbf{k} = k_o\ , \quad \text{where} \quad \varepsilon = \mathbf{y} - \mathbf{M}\mathbf{k}\ , \quad (266)$$

$$\frac{\partial I}{\partial \mathbf{k}} \overset{(153)}{=} 2(\mathbf{M}^T\mathbf{W}\mathbf{M}\mathbf{k} - \mathbf{M}^T\mathbf{W}\mathbf{y} + \lambda\mathbf{V}\mathbf{k}) \quad \rightsquigarrow \quad \mathbf{k} = (\mathbf{M}^T\mathbf{W}\mathbf{M} + \lambda\mathbf{V})^{-1}\mathbf{M}^T\mathbf{W}\mathbf{y}\ . \quad (267)$$

The Lagrange multiplier results from the equality condition with k_o only numerically (or with MATLAB `fmincon`). □

Example (8.6;3) Linear equality condition: Consider I subject to $\mathbf{v}^T\mathbf{k} = k_o$, then an analytically concise result exists

$$\frac{\partial I}{\partial \mathbf{k}} \overset{(153)}{=} 2(\mathbf{M}^T\mathbf{W}\mathbf{M}\mathbf{k} - \mathbf{M}^T\mathbf{W}\mathbf{y} + \lambda\mathbf{v}) \quad \rightsquigarrow \quad \mathbf{k} = (\mathbf{M}^T\mathbf{W}\mathbf{M})^{-1}(\mathbf{M}^T\mathbf{W}\mathbf{y} - \lambda\mathbf{v}) \quad (268)$$

$$\lambda = [\mathbf{v}^T(\mathbf{M}^T\mathbf{W}\mathbf{M})^{-1}\mathbf{M}^T\mathbf{W}\mathbf{y} - k_o]/[\mathbf{v}^T(\mathbf{M}^T\mathbf{W}\mathbf{M})^{-1}\mathbf{v}]. \quad (269)$$

The inequality condition is not active. □

Example (8.6;4) In Fig. 3, a simple two-dimensional case is shown, with an additional inequality constraint. □

Example (8.6;5). Minimizing the determinant of the covariance matrix. Gauss Markov estimation: The observation vector $\mathbf{y}$ now is regarded as the sum of the model function $\mathbf{Mp}$ plus an additive zero mean noise ε where the matrix $\mathbf{M}$ is given, that is

$$\mathbf{y} = \mathbf{M}\mathbf{p} + \varepsilon\ . \quad (270)$$

The parameter $\mathbf{p}$ is to be estimated optimally. The expectation $E\{\varepsilon\}$ should be zero. The covariance (variance) matrix $\mathbf{V}$ is defined as the expectation $E\{\varepsilon\varepsilon^T\} = \mathbf{V} = \mathbf{cov}\ \varepsilon$ and is assumed nonzero.

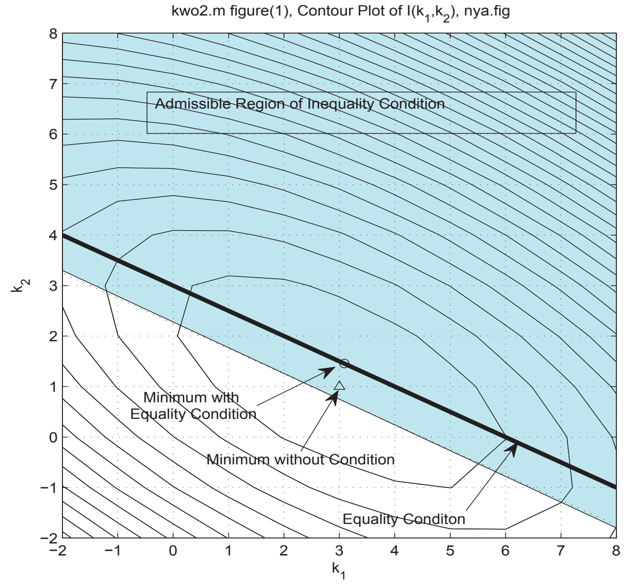

Figure 2: Example (8.6;3)
Linear equality condition

$\mathbf{M} = \begin{pmatrix} 2 & 0 \\ 1 & 2 \\ 0 & 1 \end{pmatrix}, \mathbf{y} = \begin{pmatrix} 6.1 \\ 5.2 \\ 1.3 \end{pmatrix}$

$\mathbf{W} = \mathtt{diag}(1,3,2), \quad \mathbf{v} = (1 \ \ 2)^T,$
$k_o = 6; \quad \mathbf{v}^T\mathbf{k} - k_o + 1.5 > 0;$
$\lambda = -2.55; \quad \mathbf{k}^\star = (3.09 \ \ 1.45)^T$

Suppose a linear relationship $\mathbf{p}^\star = \mathbf{\Gamma y}$ in order to derive the optimal estimation $\mathbf{p}^\star$ from the observation $\mathbf{y}$, that is $\mathbf{p}^\star = \mathbf{\Gamma y} = \mathbf{\Gamma}(\mathbf{Mp} + \varepsilon) = \mathbf{\Gamma Mp} + \mathbf{\Gamma}\varepsilon$. Applying the expectation operator yields $\hat{\mathbf{p}} = E\{\mathbf{p}^\star\} = E\{\mathbf{\Gamma Mp}\} + E\{\mathbf{\Gamma}\varepsilon\}$. The second term is the bias $\mathbf{b} = E\{\mathbf{\Gamma}\varepsilon\}$. The bias can be separated into $E\{\mathbf{\Gamma}\}$ and $E\{\varepsilon\}$ if both processes are uncorrelated. Regardless of $E\{\mathbf{\Gamma}\}$, the bias $\mathbf{b}$ vanishes if ε has zero mean as preassumed, and if ε is a white noise. If $\mathbf{b} = \mathbf{0}$ then only the first term remains: $\hat{\mathbf{p}} = E\{\mathbf{\Gamma Mp}\} = E\{\mathbf{\Gamma M}\}\mathbf{p}$. In order to obtain $\hat{\mathbf{p}} \stackrel{\triangle}{=} E\{\mathbf{p}^\star\} = \mathbf{p}$, the condition

$$\mathbf{\Gamma M} = \mathbf{I} \tag{271}$$

must be satisfied. Hence, $\mathbf{\Gamma}$ is a left pseudo-inverse of $\mathbf{M}$. To continue,

$$\begin{aligned} \mathbf{cov}\ \tilde{\mathbf{p}} &\stackrel{\triangle}{=} E\{\tilde{\mathbf{p}}\tilde{\mathbf{p}}^T\} = E\{(\mathbf{p}-\hat{\mathbf{p}})(\mathbf{p}-\hat{\mathbf{p}})^T\} = E\{[\mathbf{p} - \mathbf{\Gamma}(\mathbf{Mp}+\varepsilon)][\mathbf{p} - \mathbf{\Gamma}(\mathbf{Mp}+\varepsilon]^T\} && (272) \\ &= E\{(\mathbf{\Gamma}\varepsilon)(\mathbf{\Gamma}\varepsilon)^T\} = \mathbf{\Gamma} E\{\varepsilon\varepsilon^T\}\mathbf{\Gamma}^T = \mathbf{\Gamma}(\mathbf{cov}\ \varepsilon)\mathbf{\Gamma}^T = \mathbf{\Gamma V \Gamma}^T. && (273) \end{aligned}$$

As a scalar valuation of $\mathbf{cov\ p}$, the determinant is chosen. The objective is to minimize det cov $\tilde{\mathbf{p}}$ subject to the matrix condition $\hat{\mathbf{p}} = \mathbf{\Gamma y}$ or $\mathbf{\Gamma M} = \mathbf{I}$. Introducing a matrix-valued Lagrange multiplier $\mathbf{\Lambda}$ and selecting the trace as a scalar measure of the constraint relationship

$$\det\ (\mathbf{\Gamma V \Gamma}^T) + \mathrm{tr}\ \mathbf{\Lambda}(\mathbf{\Gamma M} - \mathbf{I}) \quad \to \quad \min_{\mathbf{\Gamma}} \tag{274}$$

is achieved. The minimum with respect to $\mathbf{\Gamma}$ is obtained by differentiation $\frac{\partial}{\partial \mathbf{\Gamma}}$ of Eq.(274), and the result by some manipulational effort

$$\begin{aligned} 2(\mathbf{\Gamma V \Gamma}^T)^{-1}\mathbf{\Gamma V} + \mathbf{\Lambda}^T\mathbf{M}^T &\stackrel{(192),(205)}{=} \mathbf{0} \quad | \times \mathbf{\Gamma}^T && (275) \\ 2(\mathbf{\Gamma V \Gamma}^T)^{-1}\mathbf{\Gamma V \Gamma}^T + \mathbf{\Lambda}^T\mathbf{M}^T\mathbf{\Gamma}^T &= 2 + \mathbf{\Lambda}^T(\mathbf{\Gamma M})^T = 2 + \mathbf{\Lambda}^T = \mathbf{0} && (276) \\ 2(\mathbf{\Gamma V \Gamma}^T)^{-1}\mathbf{\Gamma V} - 2\mathbf{M}^T &\stackrel{(275),(276)}{=} \mathbf{0} \quad | \times 0.5 \cdot \mathbf{V}^{-1} && (277) \\ (\mathbf{\Gamma V \Gamma}^T)^{-1}\mathbf{\Gamma} - \mathbf{M}^T\mathbf{V}^{-1} &= \mathbf{0} \quad | \times \mathbf{M} && (278) \\ (\mathbf{\Gamma V \Gamma}^T)^{-1}\mathbf{\Gamma M} &= \mathbf{M}^T\mathbf{V}^{-1}\mathbf{M} && (279) \\ (\mathbf{\Gamma V \Gamma}^T)^{-1}\mathbf{I} &= \mathbf{M}^T\mathbf{V}^{-1}\mathbf{M} && (280) \end{aligned}$$

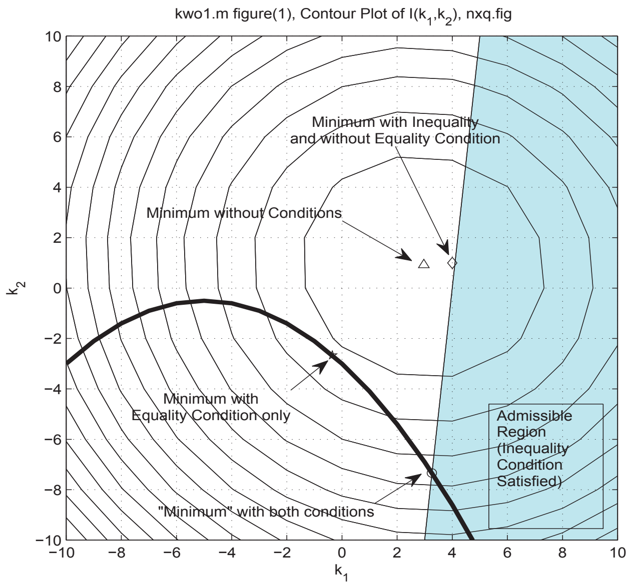

Figure 3: Example (8.6;4). Index contours and three minimum versions for the setup $I = (k_1-3)^2+(k_2-1)^2+0.1\ k_1k_2$, and equality condition $[3+k_1+k_2+0.1\ k_1^2]$, and inequality condition $[4-k_1+0.1\ k_2]$

$$(\mathbf{M}^T\mathbf{V}^{-1}\mathbf{M})\mathbf{\Gamma} \overset{(280),(278)}{=} \mathbf{M}^T\mathbf{V}^{-1} \tag{281}$$

$$\mathbf{\Gamma} = (\mathbf{M}^T\mathbf{V}^{-1}\mathbf{M})^{-1}\mathbf{M}^T\mathbf{V}^{-1}\ . \tag{282}$$

The appropriate covariance is $\mathbf{cov}\ \tilde{\mathbf{p}} = \mathbf{\Gamma V \Gamma}^T = (\mathbf{M}^T\mathbf{V}^{-1}\mathbf{M})^{-1}$. Under the condition for unbiased estimation $\mathbf{\Gamma M} = \mathbf{I}$, the minimization $\min_{\mathbf{\Gamma}}\ [\mathbf{cov}\ \tilde{\mathbf{p}}] = \min_{\mathbf{\Gamma}}\ [\mathbf{\Gamma}(\mathbf{cov}\ \varepsilon)\mathbf{\Gamma}^T]$ yields the same result $\hat{\mathbf{p}}$ as the optimum weighted least squares estimation, e.g., Eq.(267) with $\mathbf{p} = \mathbf{k}$

$$\min_{\mathbf{p}}\ I = \min_{\mathbf{p}}\ \boldsymbol{\varepsilon}^T\mathbf{W}\boldsymbol{\varepsilon} \quad \text{where} \quad \mathbf{W} = \mathbf{V}^{-1} = (\mathbf{cov}\ \boldsymbol{\varepsilon})^{-1} \quad \text{and} \quad \boldsymbol{\varepsilon} = \boldsymbol{\varepsilon}(\mathbf{p}) \tag{283}$$

(*Hoerl, A.E., 1962; Hoerl, A.E., and Kennard, R.W., 1970*). □

Example (8.6;6) Toeplitz type restriction:

The total-least-squares problem arises with the restriction of a Toeplitz scheme in the process equations of shifted data structure. Then, using a vector Lagrange multiplier $\boldsymbol{\lambda} \in \mathcal{R}^{n_m}$ (*Weinmann, A., 1995*)

$$(\mathbf{M}+\mathbf{E})\mathbf{p} = \mathbf{y} - \boldsymbol{\varepsilon} \qquad \|\mathbf{E} \vdots \boldsymbol{\varepsilon}\|_F \to \min_{\mathbf{E},\boldsymbol{\mathcal{E}},\mathbf{p}} \qquad \mathbf{M},\ \mathbf{E} \in \mathcal{R}^{n_m\times 2n} \tag{284}$$

$$\|\mathbf{E}\|_F^2 + \|\boldsymbol{\varepsilon}\|_F^2 + \boldsymbol{\lambda}^T(\mathbf{Mp}+\mathbf{Ep}+\boldsymbol{\varepsilon}-\mathbf{y}) \to \min_{\mathbf{E},\boldsymbol{\mathcal{E}},\mathbf{p}} \quad \text{subject to} \quad \boldsymbol{\varepsilon} = -\mathbf{P}_o\boldsymbol{\eta} \quad \text{and} \quad \mathrm{col}\,\mathbf{E} = -\mathrm{vec}\{\mathbf{P}_{.\,i}\boldsymbol{\eta}\} \tag{285}$$

The matrices $\mathbf{P}$ follow the specific structure, e.g.,

$$\mathbf{e}_{.\,2} = -\mathbf{P}_{.\,2}\boldsymbol{\eta} \quad \in \mathcal{R}^{n_m} \quad \text{where} \quad \mathbf{P}_{.\,2} \overset{\triangle}{=} \begin{pmatrix} 0&1&0&0&0&0&0&0&0&0&0&0&0\\ 0&0&1&0&0&0&0&0&0&0&0&0&0\\ 0&0&0&1&0&0&0&0&0&0&0&0&0\\ 0&0&0&0&1&0&0&0&0&0&0&0&0 \end{pmatrix}, \tag{286}$$

$$\mathrm{col}\,\mathbf{E} = -\begin{pmatrix} \mathbf{P}_{.\,1}\boldsymbol{\eta} \\ \mathbf{P}_{.\,2}\boldsymbol{\eta} \\ \vdots \\ \mathbf{P}_{.\,,2n}\boldsymbol{\eta} \end{pmatrix} = -\begin{pmatrix} \mathbf{P}_{.\,1} \\ \mathbf{P}_{.\,2} \\ \vdots \\ \mathbf{P}_{.\,,2n} \end{pmatrix}\boldsymbol{\eta} \overset{\triangle}{=} -\mathbf{Q}\boldsymbol{\eta}\ , \qquad \mathbf{Q} \in \mathcal{R}^{2n_m n\times r}\ . \tag{287}$$

Hence, from

$$(\mathbf{Q}\boldsymbol{\eta})^T\mathbf{Q}\boldsymbol{\eta} + (\mathbf{P}_o\boldsymbol{\eta})^T\mathbf{P}_o\boldsymbol{\eta} + \boldsymbol{\lambda}^T\mathbf{M}\mathbf{p} - (\mathbf{p}^T \otimes \boldsymbol{\lambda}^T)\mathbf{Q}\boldsymbol{\eta} - \boldsymbol{\lambda}^T\mathbf{P}_o\boldsymbol{\eta} - \boldsymbol{\lambda}^T\mathbf{y} \to \min_{\mathbf{p},\boldsymbol{\eta},\boldsymbol{\lambda}} \ , \tag{288}$$

the results are

$$\boldsymbol{\eta} = \mathbf{F}\boldsymbol{\lambda} \quad \text{where} \quad \mathbf{F}(\mathbf{p}) \ \triangleq\ 0.5(\mathbf{Q}^T\mathbf{Q} + \mathbf{P}_o^T\mathbf{P}_o)^{-1}(\mathbf{R} + \mathbf{P}_o^T) \ \in \mathcal{R}^{r\times n_m}, \tag{289}$$

$$\boldsymbol{\lambda} = -\mathbf{G}^{-1}[\mathbf{y} - \mathbf{M}\mathbf{p}] \quad \text{where} \quad \mathbf{G} \ \triangleq\ [(\mathbf{p}^T \otimes \mathbf{I}_{n_m})\mathbf{Q} + \mathbf{P}_o]\mathbf{F} = \mathbf{G}(\mathbf{p}) \ \in \mathcal{R}^{n_m\times n_m} \ , \tag{290}$$

$$\Big[\mathbf{M}^T + \{\mathbf{I}_{2n} \otimes [(\mathbf{y} - \mathbf{M}\mathbf{p})^T\mathbf{G}^{-1}]\}\mathbf{Q}\mathbf{F}\Big]\mathbf{G}^{-1}(\mathbf{y} - \mathbf{M}\mathbf{p}) = \mathbf{0} \quad \text{where } \mathbf{G},\ \mathbf{F},\ \mathbf{R} \text{ depend on } \mathbf{p}\ .\square \tag{291}$$

9 Derivatives of a Column with Respect to a Row or Column

Basic relations are

$$\frac{\partial \mathbf{x}}{\partial \mathbf{x}^T} = \mathbf{I} \quad \text{and} \quad \frac{\partial \mathbf{x}^T}{\partial \mathbf{x}} = \mathbf{I}\ . \tag{292}$$

$$\frac{\partial \mathbf{r}^T(\mathbf{x})}{\partial \mathbf{x}^T} = \Big(\frac{\partial r_1}{\partial x_1} \vdots \frac{\partial r_2}{\partial x_1} \vdots \frac{\partial r_3}{\partial x_1} \dots \frac{\partial r_1}{\partial x_2} \vdots \frac{\partial r_2}{\partial x_2} \vdots \frac{\partial r_3}{\partial x_2} \dots \frac{\partial r_n}{\partial x_r}\Big) = \Big(\frac{\partial \mathbf{r}}{\partial \mathbf{x}}\Big)^T \tag{293}$$

Furthermore,

$$\frac{\partial}{\partial \mathbf{x}^T}\mathbf{A}\mathbf{x} = \mathbf{A}\frac{\partial \mathbf{x}}{\partial \mathbf{x}^T} = \mathbf{A}\ \mathbf{I} = \mathbf{A}, \tag{294}$$

(resulting from Eq.(353) with $\mathbf{B} = \mathbf{x},\ \mathbf{M} = \mathbf{x}^T, r = 1,\ q = 1,\ s = m,\ n = m$), and

$$\frac{\partial \mathbf{A}\mathbf{x}}{\partial \mathbf{x}} = \mathrm{col}\mathbf{A} \tag{295}$$

$$\frac{\partial \mathbf{A}^{[k\times l]}\mathbf{x}}{\partial \mathbf{A}} \overset{(351)}{=} \bar{\mathbf{U}}_{k,l}(\mathbf{I}_l \otimes \mathbf{x}) \in \mathcal{R}^{k^2\times l}. \tag{296}$$

Disassembling $\mathbf{B}$ in what follows

$$\frac{\partial}{\partial \mathbf{x}}\mathbf{x}^T\mathbf{B} \triangleq \frac{\partial}{\partial \mathbf{x}}\mathbf{x}^T(\mathbf{h}_1 \vdots \mathbf{h}_2 \vdots \dots) = (\mathbf{h}_1 \vdots \mathbf{h}_2 \vdots \dots) = \mathbf{B}\ , \quad \text{or see Eq.(354)}\ . \tag{297}$$

For $\mathbf{A}$ and $\mathbf{C}$ independent of p,

$$\frac{\partial \mathrm{col}\,[\mathbf{A}\mathbf{B}\mathbf{C}]}{\partial p} = \mathrm{col}(\mathbf{A}\frac{\partial \mathbf{B}}{\partial p}\mathbf{C}). \tag{298}$$

Derivative of a set of functions in the vector $\mathbf{r}(\mathbf{x}) \in \mathcal{R}^m$ (where $\mathbf{x} \in \mathcal{R}^n$) with respect to the transposed vector $\mathbf{x}^T \in \mathcal{R}^{1\times n}$, i.e., the Jacobi matrix $\frac{\partial \mathbf{r}(\mathbf{x})}{\partial \mathbf{x}^T}$

$$\mathbf{J}_{rx} \triangleq \frac{\partial \mathbf{r}(\mathbf{x})}{\partial \mathbf{x}^T} = \begin{pmatrix} \frac{\partial r_1}{\partial \mathbf{x}^T} \\ \frac{\partial r_2}{\partial \mathbf{x}^T} \\ \vdots \\ \frac{\partial r_m}{\partial \mathbf{x}^T} \end{pmatrix} = \Big(\frac{\partial \mathbf{r}(\mathbf{x})}{\partial x_1} \vdots \frac{\partial \mathbf{r}(\mathbf{x})}{\partial x_2} \ldots \frac{\partial \mathbf{r}(\mathbf{x})}{\partial x_r}\Big) = \Big(\frac{\partial \mathbf{r}^T}{\partial \mathbf{x}}\Big)^T = \tag{299}$$

$$= \begin{pmatrix} \frac{\partial r_1}{\partial x_1} & \frac{\partial r_1}{\partial x_2} & \cdots & \frac{\partial r_1}{\partial x_n} \\ \frac{\partial r_2}{\partial x_1} & \frac{\partial r_2}{\partial x_2} & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial r_n}{\partial x_1} & \cdots & \cdots & \frac{\partial r_n}{\partial x_n} \end{pmatrix} \in \mathcal{R}^{m\times n} . \tag{300}$$

(Inconsistently, in the literatur often $\frac{\partial \mathbf{r}(\mathbf{x})}{\partial \mathbf{x}}$ is written.)

We decided $\frac{\partial r}{\partial \mathbf{x}}$ to be a vector (column matrix), according to the vector display in the denominator. The fraction $\frac{\partial \mathbf{r}}{\partial x}$ is also a vector, referring to the vector in the numerator. Hence, to consistently define the Jacobi matrix, one has to select the notation $\frac{\partial \mathbf{r}}{\partial \mathbf{x}^T}$. This is simpler than to define alternatively (*Ludyk, G., 1995*)

$$\left[\frac{\partial \mathbf{r}}{\partial \mathbf{x}} \triangleq \begin{pmatrix} (\frac{\partial r_1}{\partial \mathbf{x}})^T \\ \vdots \\ (\frac{\partial r_m}{\partial \mathbf{x}})^T \end{pmatrix}\right] . \tag{301}$$

Derivative of a scalar product of $\mathbf{f}(\mathbf{x}),\ \mathbf{r}(\mathbf{x}) \in \mathcal{R}^m$ with respect to the vector $\mathbf{x} \in \mathcal{R}^n$

$$\frac{\partial}{\partial \mathbf{x}} \mathbf{f}(\mathbf{x})^T \mathbf{r}(\mathbf{x}) = \left(\frac{\partial \mathbf{f}^T(\mathbf{x})}{\partial \mathbf{x}}\right)^{[n\times m]} \mathbf{r}(\mathbf{x}) + \left(\frac{\partial \mathbf{r}^T(\mathbf{x})}{\partial \mathbf{x}}\right)^{[n\times m]} \mathbf{f}(\mathbf{x}) \quad \in \mathcal{R}^n. \tag{302}$$

Scalar/vector product with respect to a transposed vector

$$\frac{\partial [f(\mathbf{x})\mathbf{r}(\mathbf{x})]}{\partial \mathbf{x}^T} = f(\mathbf{x}) \left(\frac{\partial \mathbf{r}(\mathbf{x})}{\partial \mathbf{x}^T}\right)^{[m\times n]} + \mathbf{r}(\mathbf{x}) \left(\frac{\partial f}{\partial \mathbf{x}^T}\right)^{[1\times n]} \quad \in \mathcal{R}^{m\times n} . \tag{303}$$

Frequently, the Jacobi matrix is used for linear approximation (small-scale linearization)

$$\dot{\mathbf{x}}(t) = \mathbf{g}(\mathbf{x}, \mathbf{u}) \ , \quad \mathbf{x}(0^+) = \mathbf{x}_o \qquad \rightsquigarrow \qquad \dot{\mathbf{x}} + \Delta\dot{\mathbf{x}} = \mathbf{g}(\mathbf{x} + \Delta\mathbf{x},\ \mathbf{u} + \Delta\mathbf{u}) \tag{304}$$

$$\frac{d}{dt}\Delta x_i(t) = \sum_{j=1}^{n} \frac{\partial g_i}{\partial x_j}\Big|_{x_{jo}} \Delta x_j(t) + \frac{\partial g_i}{\partial u_j}\Big|_{u_{jo}} \Delta u_j(t) \qquad \forall i = 1, 2...n \tag{305}$$

$$\Delta\dot{\mathbf{x}}(t) = \mathbf{J}_{gx}^T \Delta\mathbf{x} + \mathbf{J}_{gu}^T \Delta\mathbf{u} . \tag{306}$$

10 Derivatives of a Matrix

10.1 Derivations of Matrix-Valued Functions with Respect to a Scalar

$$\frac{\partial}{\partial p}\mathbf{A}(p)\mathbf{B}(p)\mathbf{C}(p) = \frac{\partial \mathbf{A}}{\partial p}\mathbf{B}\mathbf{C} + \mathbf{A}\frac{\partial \mathbf{B}}{\partial p}\mathbf{C} + \mathbf{A}\mathbf{B}\frac{\partial \mathbf{C}}{\partial p} \tag{307}$$

$$\text{or} \qquad d(\mathbf{A}\mathbf{B}\mathbf{C}) = (d\mathbf{A})\mathbf{B}\mathbf{C} + \mathbf{A}(d\mathbf{B})\mathbf{C} + \mathbf{A}\mathbf{B}(d\mathbf{C}) \ , \tag{308}$$

$$\text{e.g.,} \quad \mathbf{A}^T\mathbf{X}+\mathbf{X}\mathbf{A} = \text{constant} \quad \rightsquigarrow \quad (d\mathbf{A}^T)\mathbf{X}+\mathbf{A}^T d\mathbf{X}+(d\mathbf{X})\mathbf{A}+\mathbf{X}d\mathbf{A} = \mathbf{0} \ . \tag{309}$$

$$\frac{\partial \mathbf{A}^m}{\partial p} = \frac{\partial \mathbf{A}}{\partial p}\mathbf{A}^{m-1} + \mathbf{A}\frac{\partial \mathbf{A}}{\partial p}\mathbf{A}^{m-2} + \ldots + \mathbf{A}^{m-1}\frac{\partial \mathbf{A}}{\partial p} \ . \tag{310}$$

For a matrix $\mathbf{A} = \text{matrix}[a_{ij}]$ and using Eq.(312)

$$\frac{\partial \mathbf{A}^m}{\partial a_{ij}} = \sum_{k=0}^{k=m-1} \mathbf{A}^k\mathbf{E}_{ij}\mathbf{A}^{m-k-1}, \qquad \text{tr}[\frac{\partial \mathbf{A}^m}{\partial a_{ij}}] = (m-1)\text{tr}[\mathbf{A}^{m-1}\mathbf{E}_{ij}] \ , \tag{311}$$

$$\text{where Kronecker matrix } \frac{\partial \mathbf{K}_y}{\partial K_{y,ij}} = \mathbf{E}_{ij} \stackrel{\triangle}{=} \mathbf{e}_i\mathbf{e}_j^T \text{ with unit vector } \mathbf{e}_i \ . \tag{312}$$

From *Doyle, J.C. et al., 1996,* the derivative of the inverse is

$$\frac{\partial \mathbf{A}^{-1}}{\partial p} = -\mathbf{A}^{-1}\frac{\partial \mathbf{A}}{\partial p}\mathbf{A}^{-1} \ . \tag{313}$$

$$\frac{\partial}{\partial \Delta P_{ij}}\mathbf{Y}(\mathbf{N}+\Delta\mathbf{P}\cdot\mathbf{W})^{-1}\mathbf{Z} = -\mathbf{Y}\mathbf{N}^{-1}\mathbf{E}_{ij}\mathbf{W}\mathbf{N}^{-1}\mathbf{Z} \tag{314}$$

for $\Delta\mathbf{P}\cdot\mathbf{W} \ll \mathbf{N}$ small in norm sense, or

$$\frac{\partial}{\partial e}\mathbf{Y}(\mathbf{N}+\Delta\mathbf{P}\cdot\mathbf{W})^{-1}\mathbf{Z} = -\mathbf{Y}\mathbf{N}^{-1}\frac{\partial \Delta\mathbf{P}}{\partial e}\mathbf{W}\mathbf{N}^{-1}\mathbf{Z} \tag{315}$$

where only $\Delta\mathbf{P}$ depends on e.

$$\frac{\partial \mathbf{A}^m}{\partial a_{ij}} = \sum_{k=0}^{k=m-1} \mathbf{A}^k\mathbf{E}_{ij}\mathbf{A}^{m-k-1}, \qquad \text{tr}[\frac{\partial \mathbf{A}^m}{\partial a_{ij}}] = (m-1)\text{tr}[\mathbf{A}^{m-1}\mathbf{E}_{ij}] \ . \tag{316}$$

10.2 Gradient for the Phase Variable Form

The companion form $\mathbf{A}$ to describe an unforced linear time-invariant system is

$$\dot{\mathbf{x}}(t) = \mathbf{A}(\mathbf{p})\mathbf{x}(t) = \begin{pmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & & 1 \\ -p_1 & -p_2 & -p_3 & \ldots & -p_n \end{pmatrix} \mathbf{x}(t) \ , \tag{317}$$

with the last row of $\mathbf{A}$ and the coefficients c_i of the characteristic polynomial $c(\lambda)$ associated with $\mathbf{A}$. The coefficient matrix $\mathbf{A}$ in companion form is

$$\mathbf{A} = \sum_{i=1}^{n-1} \mathbf{E}_{i,i+1} - \mathbf{e}_n \otimes \mathbf{p}^T \qquad \in \mathcal{R}^{n\times n} \tag{318}$$

$$\mathbf{p} = (p_1 \;\; p_2 \ldots p_n)^T = (c_0 \;\; c_1 \ldots c_{n-1})^T \in \mathcal{R}^n \;\;, \quad r = n \;. \tag{319}$$

If there are no interdependencies within $\mathbf{p}$ (*Brewer, J.W., 1978*), using Eq.(379)

$$\begin{aligned}\frac{\partial \mathbf{A}}{\partial \mathbf{p}^T} &= -(\mathbf{I}_1 \otimes \mathbf{U}_{n1})\Big(\frac{\partial \mathbf{p}^T}{\partial \mathbf{p}^T} \otimes \mathbf{e}_n\Big)(\mathbf{I}_n \otimes \mathbf{U}_{n1}), \qquad \mathbf{U}_{n1} = \mathbf{U}_{1n} = \mathbf{I}_n & (320)\\ &= -\mathbf{I}_n[(\mathrm{col}^T \mathbf{I}_n) \otimes \mathbf{e}_n](\mathbf{I}_n \otimes \mathbf{I}_n) = -(\mathrm{col}\;^T \mathbf{I}_n) \otimes \mathbf{e}_n \;. & (321)\end{aligned}$$

10.3 Basic Columnized Matrices with Respect to a Matrix

Find the following differential quotients for $\mathbf{K} \in \mathcal{R}^{m\times r}$, $\Psi \in \mathcal{R}^{r\times m}$

$$\frac{\partial}{\partial \mathbf{K}} \;\mathrm{col}\; [\mathbf{\Psi K}] \in \mathcal{R}^{mpr\times r} \tag{322}$$

$$\frac{\partial}{\partial \mathbf{K}} \;\mathrm{col}\; [\mathbf{K}^T \mathbf{\Psi}^T] \in \mathcal{R}^{mpr\times r} \;. \tag{323}$$

For omitted $\mathbf{\Psi}$, the results are given by using the Kronecker product $\otimes$

$$\frac{\partial}{\partial \mathbf{K}} \;\mathrm{col}\; \mathbf{K} = (\mathbf{I}_m \otimes \mathbf{U}_{r,m})[(\mathrm{col}\; \mathbf{I}_m) \otimes \mathbf{I}_r] \overset{(\text{only for } m=1)}{=} \mathbf{I}_r \tag{324}$$

$$\frac{\partial}{\partial \mathbf{K}} \;\mathrm{col}\; \mathbf{K}^T = (\mathrm{col}\; \mathbf{I}_m) \otimes \mathbf{I}_r \;. \tag{325}$$

For $\mathbf{K}_y \in \mathcal{R}^{m\times r}$ and using the permuation matrix $\mathbf{U}_{r,m}$, see Eq.(25), one has

$$\begin{aligned}\frac{\partial}{\partial \mathbf{K}_y} \;\mathrm{col}\; \mathbf{K}_y &= (\mathbf{I}_m \otimes \mathbf{U}_{r,m})[(\mathrm{col}\; \mathbf{I}_m) \otimes \mathbf{I}_r] & (326)\\ &= \sum_{j=1,\; k=1}^{m,\; r} \mathbf{E}_{jk} \otimes \mathrm{col}\; \mathbf{E}_{jk} \in \mathcal{R}^{m^2 r\times r} & (327)\end{aligned}$$

and

$$\begin{aligned}\frac{\partial}{\partial \mathbf{K}_y} \;\mathrm{col}\; \mathbf{K}_y^T &= (\mathrm{col}\; \mathbf{I}_m) \otimes \mathbf{I}_r \in \mathcal{R}^{m^2 r\times r} & (328)\\ &= \sum_{j=1,\; k=1}^{m,\; r} \mathbf{E}_{jk} \otimes \mathrm{col}\mathbf{E}_{jk}^T & (329)\\ &= \frac{\partial}{\partial \mathbf{K}_y}\mathbf{U}_{m,r}\mathrm{col}\mathbf{K}_y = (\mathbf{I}_m \otimes \mathbf{U}_{m,r})\frac{\partial \mathrm{col}\mathbf{K}_y}{\partial \mathbf{K}_y} \;. & (330)\end{aligned}$$

Consider the controller matrix $\mathbf{K} \in \mathcal{R}^{m \times r}$ and an arbitrary matrix $\mathbf{\Psi} \in \mathcal{R}^{p \times m}$

$$\begin{aligned}
\frac{\partial}{\partial \mathbf{K}} \operatorname{col}[\mathbf{K}^T \mathbf{\Psi}^T] &\overset{(51)}{=} \frac{\partial}{\partial \mathbf{K}} (\mathbf{\Psi} \otimes \mathbf{I}_r) \operatorname{col} \mathbf{K}^T && (331)\\
&\overset{(353)}{=} [\mathbf{I}_m \otimes (\mathbf{\Psi} \otimes \mathbf{I}_r)] \frac{\partial}{\partial \mathbf{K}} \operatorname{col} \mathbf{K}^T && (332)\\
&\overset{(328)}{=} [(\mathbf{I}_m \otimes \mathbf{\Psi}) \otimes \mathbf{I}_r][(\operatorname{col} \mathbf{I}_m) \otimes \mathbf{I}_r] && (333)\\
&\overset{(43)}{=} [(\mathbf{I}_m \otimes \mathbf{\Psi}) \operatorname{col} \mathbf{I}_m] \otimes (\mathbf{I}_r \mathbf{I}_r) && (334)\\
\frac{\partial}{\partial \mathbf{K}} \operatorname{col}[\mathbf{K}^T \mathbf{\Psi}^T] &= \operatorname{col}(\mathbf{\Psi} \mathbf{I}_m \mathbf{I}_m) \otimes \mathbf{I}_r = (\operatorname{col} \mathbf{\Psi}) \otimes \mathbf{I}_r \ \in \mathcal{R}^{mpr \times r} \ . && (335)
\end{aligned}$$

Using the same correspondences once more and Eq.(324)

$$\begin{aligned}
\frac{\partial}{\partial \mathbf{K}} \operatorname{col} [\mathbf{\Psi} \mathbf{K}] &= \frac{\partial}{\partial \mathbf{K}} (\mathbf{I}_r \otimes \mathbf{\Psi}) \operatorname{col}\mathbf{K} && (336)\\
&\overset{(351)}{=} [\mathbf{I}_m \otimes (\mathbf{I}_r \otimes \mathbf{\Psi})] \frac{\partial}{\partial \mathbf{K}} \operatorname{col} \mathbf{K} && (337)\\
\frac{\partial}{\partial \mathbf{K}} \operatorname{col}[\mathbf{\Psi} \mathbf{K}] &= [\mathbf{I}_m \otimes \mathbf{I}_r \otimes \mathbf{\Psi}](\mathbf{I}_m \otimes \mathbf{U}_{r,m})[(\operatorname{col} \mathbf{I}_m) \otimes \mathbf{I}_r] && (338)
\end{aligned}$$

which can be simplified to

$$\frac{\partial}{\partial \mathbf{K}} \operatorname{col}[\mathbf{\Psi} \mathbf{K}] = \mathbf{U}_{rm,p}[(\operatorname{col}\mathbf{\Psi}^T) \otimes \mathbf{I}_r] \in \mathcal{R}^{mpr \times r} \tag{339}$$

or to

$$\frac{\partial \operatorname{col}(\mathbf{\Psi} \mathbf{K})}{\partial \mathbf{K}} = \begin{pmatrix} \mathbf{I}_r \otimes (\operatorname{col}\mathbf{\Psi})_1 \\ \vdots \\ \mathbf{I}_r \otimes (\operatorname{col}\mathbf{\Psi})_m \end{pmatrix} , \tag{340}$$

where $(\operatorname{col}\mathbf{\Psi})_m$ is the m-th column of $\mathbf{\Psi}$.

For the sake of systematic presentation, note that from Eq.(337)

$$\frac{\partial}{\partial \mathbf{K}} \operatorname{col}[\mathbf{\Psi} \mathbf{K}] = (\mathbf{I}_m \otimes \mathbf{I}_r \otimes \mathbf{\Psi}) \frac{\partial}{\partial \mathbf{K}} \operatorname{col}\mathbf{K} = (\mathbf{I}_{mr} \otimes \mathbf{\Psi}) \frac{\partial}{\partial \mathbf{K}} \operatorname{col}\mathbf{K} \ , \tag{341}$$

and from Eqs.(335) and (353)

$$\frac{\partial}{\partial \mathbf{K}} \operatorname{col} [\mathbf{K}^T \mathbf{\Psi}^T] = (\mathbf{I}_m \otimes \mathbf{\Psi} \otimes \mathbf{I}_r) \frac{\partial}{\partial \mathbf{K}} \operatorname{col} \mathbf{K}^T \ . \tag{342}$$

Generalizing with a matrix $\mathbf{C} \in \mathcal{R}^{r \times v}$ yields on the one hand

$$\begin{aligned}
\frac{\partial}{\partial \mathbf{K}} \operatorname{col} [\mathbf{\Psi} \mathbf{K} \mathbf{C})] &= [\mathbf{I}_m \otimes \mathbf{C}^T \otimes \mathbf{\Psi}](\mathbf{I}_m \otimes \mathbf{U}_{r,m})[(\operatorname{col} \mathbf{I}_m) \otimes \mathbf{I}_r] && (343)\\
&\overset{(339)}{=} \mathbf{U}_{mv,p}[(\operatorname{col} \mathbf{\Psi}^T) \otimes \mathbf{C}^T] \in \mathcal{R}^{mpr \times v}, && (344)
\end{aligned}$$

and on the other hand

$$\frac{\partial}{\partial \mathbf{K}} \operatorname{col}[\mathbf{C}^T\mathbf{K}^T\mathbf{\Psi}^T] = [\mathbf{I}_m \otimes \mathbf{\Psi} \otimes \mathbf{C}^T][(\operatorname{col}\mathbf{I}_m) \otimes \mathbf{I}_r] \tag{345}$$

$$= [(\mathbf{I}_m \otimes \mathbf{\Psi})\operatorname{col}\mathbf{I}_m] \otimes [\mathbf{C}^T\mathbf{I}_r] \tag{346}$$

$$\frac{\partial}{\partial \mathbf{K}} \operatorname{col}[\mathbf{C}^T\mathbf{K}^T\mathbf{\Psi}^T] = (\operatorname{col}\mathbf{\Psi}) \otimes \mathbf{C}^T \in \mathcal{R}^{mpr\times v}\ . \tag{347}$$

Derivative of a vector-valued function (a j-th column $\mathbf{A}_{:,j}$ of a matrix $\mathbf{A}$) with respect to a matrix $\mathbf{M}$, i.e., $\frac{\partial \mathbf{f}}{\partial \mathbf{M}}$, if the entire matrix-valued function $\mathbf{A}$ has been used in a derivation before

$$\frac{\partial \mathbf{A}_{:,j}}{\partial \mathbf{M}} = \frac{\partial \mathbf{A}}{\partial \mathbf{M}}(\mathbf{I}_s \otimes \mathbf{e}_j^{(m\times 1)}) \qquad \mathbf{A}\in\mathcal{R}^{n\times m},\quad \mathbf{M}\in\mathcal{R}^{r\times s} \tag{348}$$

The matrix $(\mathbf{I}_s \otimes \mathbf{e}_j^{(m\times 1)})$ provides the selection of the j-th column of the result $\partial\mathbf{A}/\partial\mathrm{M}$.

10.4 Matrix Gradients with Respect to a Matrix

General definition

$$\frac{\partial \mathbf{G}}{\partial \mathrm{M}} \triangleq \sum_{ij} \mathbf{E}_{ij}^{(r\times s)} \otimes \frac{\partial \mathbf{G}}{\partial M_{ij}} = \begin{pmatrix} \frac{\partial \mathbf{G}}{\partial M_{11}} & \frac{\partial \mathbf{G}}{\partial M_{12}} & \cdots & \frac{\partial \mathbf{G}}{\partial M_{1s}} \\ \frac{\partial \mathbf{G}}{\partial M_{21}} & \frac{\partial \mathbf{G}}{\partial M_{22}} & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial \mathbf{G}}{\partial M_{r1}} & \frac{\partial \mathbf{G}}{\partial M_{r2}} & \cdots & \frac{\partial \mathbf{G}}{\partial M_{rs}} \end{pmatrix} \tag{349}$$

$$\mathbf{G}\in\mathcal{R}^{n\times m},\ \mathbf{M}\in\mathcal{R}^{r\times s},\ \frac{\partial \mathbf{G}}{\partial \mathbf{M}} \in \mathcal{R}^{nr\times ms}\ . \tag{350}$$

The matrix product rule with $\mathbf{A}\in\mathcal{R}^{n\times m}, \mathbf{B}\in\mathcal{R}^{m\times q}, \mathbf{M}\in\mathcal{R}^{r\times s}$, is

$$\frac{\partial}{\partial \mathbf{M}}\mathbf{A}(\mathbf{M})\mathbf{B}(\mathbf{M}) = \frac{\partial \mathbf{A}}{\partial \mathbf{M}}(\mathbf{I}_s\otimes\mathbf{B}) + (\mathbf{I}_r\otimes\mathbf{A})\frac{\partial \mathbf{B}}{\partial \mathbf{M}} \qquad \frac{\partial(\mathbf{AB})}{\partial \mathbf{M}} \in \mathcal{R}^{nr\times qs}\ . \tag{351}$$

$$\frac{\partial}{\partial \mathbf{M}}\mathbf{A}(\mathbf{M})\mathbf{B}(\mathbf{M})\mathbf{C}(\mathbf{M}) \overset{(351)}{=} \frac{\partial \mathbf{A}}{\partial \mathbf{M}}[\mathbf{I}_s\otimes(\mathbf{BC})] + (\mathbf{I}_r\otimes\mathbf{A})\frac{\partial \mathbf{B}}{\partial \mathbf{M}}(\mathbf{I}_s\otimes\mathbf{C}) + [\mathbf{I}_r\otimes(\mathbf{AB})]\frac{\partial \mathbf{C}}{\partial \mathbf{M}} \tag{352}$$

For $\mathbf{A}$ constant,

$$\frac{\partial}{\partial \mathbf{M}}\mathbf{A}\mathbf{B}^{(n\times m)}(\mathbf{M}) = (\mathbf{I}_r\otimes\mathbf{A})\frac{\partial \mathbf{B}}{\partial \mathbf{M}}. \tag{353}$$

For $\mathbf{B}$ constant, $\mathbf{A}=\mathbf{x}^T,\ \mathbf{M}=\mathbf{x},\ s=n=1,\ q=r=m$,

$$\frac{\partial}{\partial \mathbf{x}}\mathbf{x}^T\mathbf{B} = \frac{\partial \mathbf{A}}{\partial \mathbf{M}}(\mathbf{I}_s\otimes\mathbf{B}) = \frac{\partial \mathbf{x}^T}{\partial \mathbf{x}}\mathbf{B} = \mathbf{I}\ \mathbf{B} = \mathbf{B}\ . \tag{354}$$

$$\frac{\partial}{\partial \mathbf{M}}a(\mathbf{M})\mathbf{B}(\mathbf{M}) = \frac{\partial a}{\partial \mathbf{M}}\otimes\mathbf{B} + \frac{\partial \mathbf{B}}{\partial \mathbf{M}}\otimes a = \frac{\partial a(\mathbf{M})}{\partial \mathbf{M}}\otimes\mathbf{B} + a(\mathbf{M})\frac{\partial \mathbf{B}}{\partial \mathbf{M}}\ . \tag{355}$$

Important examples are

$$\frac{\partial(\mathbf{A}+\mathbf{BK})}{\partial\mathbf{K}} = \frac{\partial\mathbf{BK}}{\partial\mathbf{K}} = (\mathbf{I}_m \otimes \mathbf{B})\frac{\partial\mathbf{K}}{\partial\mathbf{K}} = (\mathbf{I}_m \otimes \mathbf{B})\bar{\mathbf{U}}_{mn} \tag{356}$$

$$\frac{\partial(\mathbf{A}+\mathbf{BK})^T}{\partial\mathbf{K}} = \frac{\partial\mathbf{K}^T\mathbf{B}^T}{\partial\mathbf{K}} = \mathbf{U}_{mn}(\mathbf{I}_n \otimes \mathbf{B}^T) \tag{357}$$

$$\frac{\partial\mathbf{P}(\mathbf{A}+\mathbf{BK})}{\partial\mathbf{K}} = \frac{\partial\mathbf{P}}{\partial\mathbf{K}}[\mathbf{I}_n \otimes (\mathbf{A}+\mathbf{BK})] + (\mathbf{I}_m \otimes \mathbf{P})\frac{\partial(\mathbf{A}+\mathbf{BK})}{\partial\mathbf{K}} \tag{358}$$

$$\frac{\partial(\mathbf{A}+\mathbf{BK})^T\mathbf{P}}{\partial\mathbf{K}} = \frac{\partial(\mathbf{A}+\mathbf{BK})^T}{\partial\mathbf{K}}(\mathbf{I}_n \otimes \mathbf{P}) + [\mathbf{I}_m \otimes (\mathbf{A}+\mathbf{BK})^T]\frac{\partial\mathbf{P}}{\partial\mathbf{K}} \tag{359}$$

$$\begin{aligned}\frac{\partial\mathbf{K}^T\mathbf{RK}}{\partial\mathbf{K}} &= \mathbf{U}_{mn}(\mathbf{I}_n \otimes \mathbf{RK}) + (\mathbf{I}_m \otimes \mathbf{K}^T)\frac{\partial\mathbf{RK}}{\partial\mathbf{K}} && (360)\\ &= \mathbf{U}_{mn}[\mathbf{I}_n \otimes (\mathbf{RK})] + [\mathbf{I}_m \otimes (\mathbf{K}^T\mathbf{R})]\bar{\mathbf{U}}_{mn}\ . && (361)\end{aligned}$$

Derivative of the inverse matrix (*Vetter, W.J., 1973*)

$$\frac{\partial\mathbf{A}^{-1}}{\partial\mathbf{M}} = -(\mathbf{I}_r \otimes \mathbf{A}^{-1})\frac{\partial\mathbf{A}}{\partial\mathbf{M}}(\mathbf{I}_s \otimes \mathbf{A}^{-1}) \qquad \mathbf{A} \in \mathcal{R}^{n\times n}\ . \tag{362}$$

$$\text{For}\ \ \mathbf{X} \in \mathcal{R}^{n\times n} \qquad \frac{\partial\mathbf{X}^{-1}}{\partial p} = -\mathbf{X}^{-1}\frac{\partial\mathbf{X}}{\partial p}\mathbf{X}^{-1}\ . \tag{363}$$

$$\begin{aligned}\frac{\partial(\mathbf{X}\mathbf{X}^{-1})}{\partial\mathbf{X}} = \frac{\partial\mathbf{I}_n}{\partial\mathbf{X}} &= \mathbf{0} && (364)\\ \frac{\partial\mathbf{X}}{\partial\mathbf{X}}(\mathbf{I}_n \otimes \mathbf{X}^{-1}) + (\mathbf{I}_n \otimes \mathbf{X})\frac{\partial\mathbf{X}^{-1}}{\partial\mathbf{X}} &\overset{(351)}{=} \mathbf{0} && (365)\\ \frac{\partial\mathbf{X}^{-1}}{\partial\mathbf{X}} &\overset{(41),(44)}{=} -(\mathbf{I}_n \otimes \mathbf{X}^{-1})\bar{\mathbf{U}}_{n,n}(\mathbf{I}_n \otimes \mathbf{X}^{-1})\ . && (366)\end{aligned}$$

$$\begin{aligned}\frac{\partial}{\partial\mathbf{M}}\mathbf{A}^v &\overset{(351)}{=} \frac{\partial\mathbf{A}}{\partial\mathbf{M}}(\mathbf{I}_s \otimes \mathbf{A}^{v-1}) + (\mathbf{I}_r \otimes \mathbf{A})\frac{\partial\mathbf{A}^{v-1}}{\partial\mathbf{M}} \qquad \in \mathcal{R}^{nr\times qs} && (367)\\ &= \frac{\partial\mathbf{A}}{\partial\mathbf{M}}(\mathbf{I}_s \otimes \mathbf{A}^{v-1}) + \sum_{i=1}^{v-2}(\mathbf{I}_r \otimes \mathbf{A}^i)\frac{\partial\mathbf{A}}{\partial\mathbf{M}}(\mathbf{I}_s \otimes \mathbf{A}^{v-i-1}) + (\mathbf{I}_r \otimes \mathbf{A}^{v-1})\frac{\partial\mathbf{A}}{\partial\mathbf{M}}\\ &= \sum_{i=1}^{v}(\mathbf{I}_r \otimes \mathbf{A}^{i-1})\frac{\partial\mathbf{A}}{\partial\mathbf{M}}(\mathbf{I}_s \otimes \mathbf{A}^{v-i})\ . \qquad (v \geq 3) && (368)\end{aligned}$$

and for direct transfer purposes, only, $\mathbf{k}_r \in \mathcal{R}^r$

$$\frac{\partial\mathbf{A}^{v+1}}{\partial\mathbf{k}_r} = \sum_{i=1}^{v+1}(\mathbf{I}_r \otimes \mathbf{A}^{i-1})\frac{\partial\mathbf{A}}{\partial\mathbf{k}_r}\mathbf{A}^{v-i+1})\ , \qquad v \geq 0\ , \tag{369}$$

$$\frac{\partial \mathbf{A}^{k-1}}{\partial \mathbf{k}_r} = \sum_{j=1}^{k-1} (\mathbf{I}_r \otimes \mathbf{A}^{j-1}) \frac{\partial \mathbf{A}}{\partial \mathbf{k}_r} \mathbf{A}^{k-j-1} \ , \quad k \geq 2 \ . \tag{370}$$

For differentiation with respect to $A_{i,j}$

$$\frac{\partial \mathbf{A}^{k-1}}{\partial A_{i,j}} = \sum_{v=1}^{v=k-1} \mathbf{A}^{v-1} \mathbf{E}_{i,j} \mathbf{A}^{k-v-1} \ , \quad k \geq 2 \ , \tag{371}$$

$$\frac{\partial \mathbf{A}^{N}}{\partial A_{i,j}} = \sum_{v=1}^{v=N} \mathbf{A}^{v-1} \mathbf{E}_{i,j} \mathbf{A}^{N-v} \ , \quad n \geq 1 \ . \tag{372}$$

Derivatives of square terms are

$$\frac{\partial \mathbf{A}^2}{\partial \mathbf{A}} \overset{(351)}{=} \frac{\partial \mathbf{A}}{\partial \mathbf{A}} (\mathbf{I} \otimes \mathbf{A}) + (\mathbf{I} \otimes \mathbf{A}) \frac{\partial \mathbf{A}}{\partial \mathbf{A}} \tag{373}$$

$$\overset{(41)}{=} \bar{\mathbf{U}}_{n,n} (\mathbf{I} \otimes \mathbf{A}) + (\mathbf{I} \otimes \mathbf{A}) \bar{\mathbf{U}}_{n,n} \tag{374}$$

$$\frac{\partial \mathbf{A}^T \mathbf{A}}{\partial \mathbf{A}} \overset{(351)}{=} \frac{\partial \mathbf{A}^T}{\partial \mathbf{A}} (\mathbf{I} \otimes \mathbf{A}) + (\mathbf{I} \otimes \mathbf{A}^T) \frac{\partial \mathbf{A}}{\partial \mathbf{A}} \tag{375}$$

$$\overset{(30),(41)}{=} \mathbf{U}_{n,n} (\mathbf{I} \otimes \mathbf{A}) + (\mathbf{I} \otimes \mathbf{A}^T) \bar{\mathbf{U}}_{n,n} \tag{376}$$

$$\frac{\partial \mathbf{b} \mathbf{k}_r^T}{\partial \mathbf{k}_r} = (\mathbf{I}_n \otimes \mathbf{b}) \mathbf{U}_{n,1} = (\mathbf{I}_n \otimes \mathbf{b}) \ , \quad \mathbf{b} = \text{constant}, \mathbf{k}_r \in \mathcal{R}^n \tag{377}$$

$$\frac{\partial \mathbf{k}_r \mathbf{b}^T}{\partial \mathbf{k}_r} = \bar{\mathbf{U}}_{n,1} \mathbf{b}^T \ , \quad \mathbf{b} = \text{constant}, \mathbf{k}_r \in \mathcal{R}^n \ . \tag{378}$$

The derivative of the Kronecker product of $\mathbf{A} \in \mathcal{R}^{n \times m}$ and $\mathbf{B} \in \mathcal{R}^{k \times l}$ with respect to a matrix $\mathbf{M} \in \mathcal{R}^{r \times s}$ is

$$\frac{\partial (\mathbf{A} \otimes \mathbf{B})}{\partial \mathbf{M}} = \frac{\partial \mathbf{A}}{\partial \mathbf{M}} \otimes \mathbf{B} + (\mathbf{I}_r \otimes \mathbf{U}_{nk}) \Big(\frac{\partial \mathbf{B}}{\partial \mathbf{M}} \otimes \mathbf{A} \Big) (\mathbf{I}_s \otimes \mathbf{U}_{lm}) \ . \tag{379}$$

Note that $\mathbf{I}_r, \mathbf{U}_{nk}$ etc. are not conformable for matrix multiplication.

The derivative with respect to a scalar p

$$\frac{\partial (\mathbf{A}_M \otimes \mathbf{B}_M)}{\partial p} \overset{(379)}{=} \frac{\partial \mathbf{A}_M}{\partial p} \otimes \mathbf{B}_M + \mathbf{U}_{nk} \Big(\frac{\partial \mathbf{B}_M}{\partial p} \otimes \mathbf{A}_M \Big) \mathbf{U}_{lm} \ . \tag{380}$$

$$\frac{\partial \mathbf{e} \otimes \mathbf{B}}{\partial p} \overset{(380)}{=} \frac{\partial \mathbf{B}}{\partial p} \otimes \mathbf{e} \ , \quad \mathbf{e} = \text{constant. Note the change of order.} \ (\mathbf{U}_{n,1} = \mathbf{I}_n, \ \mathbf{U}_{1,n} = \mathbf{I}_n) \tag{381}$$

$$\frac{\partial \mathbf{A} \otimes \mathbf{e}}{\partial p} \overset{(380)}{=} \frac{\partial \mathbf{A}}{\partial p} \otimes \mathbf{e} \ , \quad \mathbf{e} = \text{constant} \ . \tag{382}$$

Hadamard product derivatives: Following Eq.(85), and using $\mathbf{1}^{[r\times s]} = \mathtt{ones(n,m)}$,

$$\begin{aligned}
\frac{\partial \mathbf{A}^{[n\times m]} \cdot * \mathbf{B}(\mathbf{M})}{\partial \mathbf{M}^{[r\times s]}} &= \left(\mathbf{1}^{[r\times s]} \otimes \mathbf{A}^{[n\times m]}\right) \cdot * \left(\frac{\partial \mathbf{B}^{[n\times m]}}{\partial \mathbf{M}}\right) && (383)\\
&\overset{(\mathbf{M}=\mathbf{B})}{=} \left(\mathbf{1}^{[n\times m]} \otimes \mathbf{A}^{[n\times m]}\right) \cdot * \sum_{i,j} \mathbf{E}_{ij}^{[n\times m]} \otimes \frac{\partial \mathbf{B}}{\partial B_{ij}} && (384)\\
&= \left(\mathbf{1}^{[n\times m]} \otimes \mathbf{A}^{[n\times m]}\right) \cdot * \sum_{i,j} \mathbf{E}_{ij}^{[n\times m]} \mathbf{E}_{ij}^{[n\times m]}. && (385)
\end{aligned}$$

Alternatively,

$$\begin{aligned}
\frac{\partial \mathbf{A}^{[n\times m]} \cdot * \mathbf{B}(\mathbf{M})}{\partial \mathbf{M}^{[r\times s]}} &= \sum_{i,j}^{r,s} \mathbf{E}_{ij}^{[r\times s]} \otimes \frac{\partial \mathbf{A} \cdot * \mathbf{B}(\mathbf{M})}{\partial M_{ij}} && (386)\\
&\overset{(\mathbf{M}=\mathbf{B})}{=} \sum_{i,j}^{n,m} \mathbf{E}_{ij} \otimes (A_{ij}\mathbf{E}_{ij}) = \sum_{i,j}^{n,m} A_{ij}\mathbf{E}_{ij} \otimes \mathbf{E}_{ij}\ . && (387)
\end{aligned}$$

$$\begin{aligned}
\frac{\partial \mathbf{adj}[\mathbf{X}]}{\partial \mathbf{X}} &= \frac{\partial \mathbf{X}^{-1} \det \mathbf{X}}{\partial \mathbf{X}} \overset{(355)}{=} \frac{\partial \det \mathbf{X}}{\partial \mathbf{X}} \otimes \mathbf{X}^{-1} + (\det \mathbf{X}) \frac{\partial \mathbf{X}^{-1}}{\partial \mathbf{X}} && (388)\\
&\overset{(188),(366),(41)}{=} [\mathbf{adj}^T(\mathbf{X})] \otimes \mathbf{X}^{-1} \\
&\quad -(\det \mathbf{X}) \cdot (\mathbf{I}_n \otimes \mathbf{X}^{-1}) \cdot \bar{\mathbf{U}}_{n,n} \cdot (\mathbf{I}_n \otimes \mathbf{X}^{-1})\ . && (389)
\end{aligned}$$

Based on

$$\mathbf{\Phi}(t) = e^{\mathbf{A}t}, \qquad \frac{d}{dt}\mathbf{\Phi}(t) = \mathbf{A}\mathbf{\Phi}(t), \qquad \mathbf{\Phi}(0) = \mathbf{I}\ , \tag{390}$$

differentiating with respect to $\mathbf{K}$ and using the product rule Eq.(351) yields

$$\frac{\partial}{\partial \mathbf{K}} \frac{d}{dt}\mathbf{\Phi}(t) = (\mathbf{I}_m \otimes \mathbf{A}) \frac{\partial \mathbf{\Phi}}{\partial \mathbf{K}} + \frac{\partial \mathbf{A}}{\partial \mathbf{K}} (\mathbf{I}_r \otimes \mathbf{\Phi}) \tag{391}$$

$$\frac{d}{dt}\frac{\partial \mathbf{\Phi}(t)}{\partial \mathbf{K}} - (\mathbf{I}_m \otimes \mathbf{A}) \frac{\partial \mathbf{\Phi}(t)}{\partial \mathbf{K}} = \frac{\partial \mathbf{A}}{\partial \mathbf{K}} \left(\mathbf{I}_r \otimes \mathbf{\Phi}(t)\right)\ . \tag{392}$$

For further derivatives $\partial\mathbf{\Phi}/\partial\mathbf{K}$ see *Brewer, J.W., 1977; 1978; 1978a.*

10.5 Chain Rules

The chain rule for scalar functions $\mathbf{M} \in \mathcal{R}^{r\times s}$ is (*Brewer, J.W., 1977; 1977a; 1978a*)

$$\begin{aligned}
\frac{\partial a[b(\mathbf{M})]}{\partial \mathbf{M}} &= \left(\mathbf{I}_r \otimes \frac{\partial a}{\partial b}\right) \left(\frac{\partial b}{\partial \mathbf{M}} \otimes \mathbf{I}_m\right) && (393)\\
&= \left(\frac{\partial a}{\partial b}\mathbf{I}_r\right) \left(\frac{\partial b}{\partial \mathbf{M}} \otimes 1\right) = \frac{\partial a}{\partial b}\mathbf{I}_r \frac{\partial b}{\partial \mathbf{M}} = \frac{\partial a}{\partial b}\frac{\partial b}{\partial \mathbf{M}}\ . && (394)
\end{aligned}$$

$$\frac{\partial}{\partial \mathbf{q}}\ \frac{1}{\mathbf{a}^T\mathbf{q}} = \frac{-1}{(\mathbf{a}^T\mathbf{q})^2} \cdot \mathbf{a} = \frac{-\mathbf{a}}{(\mathbf{a}^T\mathbf{q})^2}\ . \tag{395}$$

10.6 Chain Rule for Vector Functions

$$\frac{\partial f[\mathbf{r}(\mathbf{x})]}{\partial \mathbf{x}} = \left(\frac{\partial \mathbf{r}^T(\mathbf{x})}{\partial \mathbf{x}}\right)^{[n\times m]} \left(\frac{\partial f}{\partial \mathbf{r}}\right)^{[m\times 1]} = \mathbf{J}_{rx}^T \frac{\partial f}{\partial \mathbf{r}} \quad \in \mathcal{R}^n \ . \tag{396}$$

10.7 Jacobi Matrix

From Eq.(300),

$$\text{Jacobi matrix} \quad \mathbf{J}_{rx} \triangleq \frac{\partial \mathbf{r}}{\partial \mathbf{x}^T} \quad \text{from source } \mathbf{x} \text{ to target } \mathbf{r}, \tag{397}$$

in detail, equivalent to Eq.(300),

$$\frac{\partial \mathbf{r}(\mathbf{x})}{\partial \mathbf{x}^T} = \left(\frac{\partial \mathbf{r}(\mathbf{x})}{\partial x_1} \vdots \frac{\partial \mathbf{r}(\mathbf{x})}{\partial x_2} \vdots \ldots \vdots \frac{\partial \mathbf{r}(\mathbf{x})}{\partial x_n}\right) = \left(\frac{\partial \mathbf{r}^T}{\partial \mathbf{x}}\right)^T = \begin{pmatrix} \frac{\partial r_1}{\partial x_1} & \frac{\partial r_1}{\partial x_2} & \frac{\partial r_1}{\partial x_3} & \cdots \\ \frac{\partial r_2}{\partial x_1} & \frac{\partial r_2}{\partial x_2} & \cdots & \\ \vdots & \vdots & \ddots & \\ \frac{\partial r_m}{\partial x_1} & \cdots & & \frac{\partial r_m}{\partial x_n} \end{pmatrix}, \tag{398}$$

where in the Jacobi matrix $\mathbf{J}_{rx} \in \mathcal{R}^{\dim(\mathbf{r})\times\dim(\mathbf{x})} = \mathcal{R}^{m\times n}$ the rows are oriented at $\mathbf{r} \in \mathcal{R}^m$ and the columns at $\mathbf{x} \in \mathcal{R}^n$. For the special case $\mathbf{y}(\mathbf{x}) = \mathbf{Q}\mathbf{x}$, $\mathbf{Q} \in \mathcal{R}^{m\times n}$, one has the Jacobi matrix $\frac{\partial \mathbf{y}}{\partial \mathbf{x}^T} = \mathbf{Q}$. A partition $\frac{\partial \mathbf{y}}{\partial x_i}$ is the ith column of $\mathbf{Q}$ or the ith column of $\frac{\partial \mathbf{y}}{\partial \mathbf{x}^T}$. In transposed form $\frac{\partial \mathbf{y}^T}{\partial \mathbf{x}} = \mathbf{Q}^T$.

$$\text{Scalar } f: \quad \frac{\partial f[\mathbf{y}(\mathbf{x})]}{\partial \mathbf{x}} = \left(\frac{\partial \mathbf{y}^T}{\partial \mathbf{x}}\right)^{[n\times p]} \left(\frac{\partial f}{\partial \mathbf{y}}\right)^{[p\times 1]} = \mathbf{J}_{yx}^T \frac{\partial f}{\partial \mathbf{y}} \quad \in \mathcal{R}^n \tag{399}$$

$$\frac{\partial f[\mathbf{z}(\mathbf{y}(\mathbf{x}))]}{\partial \mathbf{x}} = \left(\frac{\partial \mathbf{y}^T}{\partial \mathbf{x}}\right)^{[n\times p]} \left(\frac{\partial \mathbf{z}^T}{\partial \mathbf{y}}\right)^{[p\times s]} \left(\frac{\partial f}{\partial \mathbf{z}}\right)^{[s\times 1]} = \mathbf{J}_{yx}^T \mathbf{J}_{zy}^T \frac{\partial f}{\partial \mathbf{z}} \quad \in \mathcal{R}^n \tag{400}$$

$$\text{Vector } \mathbf{r}: \quad \frac{\partial \mathbf{r}[\mathbf{z}(\mathbf{y}(\mathbf{x}))]}{\partial \mathbf{x}^T} = \left(\frac{\partial \mathbf{r}}{\partial \mathbf{z}^T}\right)^{[m\times s]} \left(\frac{\partial \mathbf{z}}{\partial \mathbf{y}^T}\right)^{[s\times p]} \left(\frac{\partial \mathbf{y}}{\partial \mathbf{x}^T}\right)^{[p\times n]} = \mathbf{J}_{rz}\mathbf{J}_{zy}\mathbf{J}_{yx} \quad \in \mathcal{R}^{m\times n} \ . \tag{401}$$

Comparing Eq.(401) with Eq.(396) looks as some irregularity at a first glance, as far as the order and the transpositions are concerned. However, all is true because

$$\frac{\partial f[\mathbf{r}(\mathbf{x})]}{\partial \mathbf{x}} = \frac{\partial \mathbf{r}^T}{\partial \mathbf{x}} \frac{\partial f}{\partial \mathbf{r}} = \mathbf{J}_{rx}^T \frac{\partial f}{\partial \mathbf{r}} \ . \tag{402}$$

From Eq.(401)

$$\frac{\partial \mathbf{f}[\mathbf{r}(\mathbf{x})]}{\partial \mathbf{x}^T} = \mathbf{J}_{fr}\mathbf{J}_{rx} \tag{403}$$

$$\text{transposing} \quad \rightsquigarrow \quad \frac{\partial \mathbf{f}^T[\mathbf{r}(\mathbf{x})]}{\partial \mathbf{x}} = \mathbf{J}_{rx}^T \mathbf{J}_{fr}^T = \mathbf{J}_{rx}^T \frac{\partial \mathbf{f}^T}{\partial \mathbf{r}} \ . \tag{404}$$

Comparing Eq.(402) and Eq.(404): Eq.(404) comprises a complete row $\mathbf{f}^T$, Eq.(402) is only one element of this row.

10.8 Nonlinear System Linearization

Consider the nonlinear system $\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}, \mathbf{u}), \quad \mathbf{y}(t) = \mathbf{g}(\mathbf{x}, \mathbf{u})$ and small increments $\Delta\mathbf{x}$ etc. in a specific point of operation. Evaluating the first-order Taylor expansion,

$$\Delta\dot{\mathbf{x}}(t) = \frac{\partial \mathbf{f}}{\partial \mathbf{x}^T}\Delta\mathbf{x} + \frac{\partial \mathbf{f}}{\partial \mathbf{u}^T}\Delta\mathbf{u} \tag{405}$$

$$\Delta\mathbf{y}(t) = \frac{\partial \mathbf{g}}{\partial \mathbf{x}^T}\Delta\mathbf{x} + \frac{\partial \mathbf{g}}{\partial \mathbf{u}^T}\Delta\mathbf{u} \; . \tag{406}$$

Reusing $\mathbf{x}$ instead of $\Delta\mathbf{x}$ etc. for the sake of simplification, and using the Jacobi matrix, one has

$$\dot{\mathbf{x}}(t) = \mathbf{J}_{fx}\mathbf{x} + \mathbf{J}_{xu}\mathbf{u} \tag{407}$$

$$\mathbf{y}(t) = \mathbf{J}_{yx}\mathbf{x} + \mathbf{J}_{yu}\mathbf{u} \; , \tag{408}$$

where

$$\mathbf{J}_{fx} \triangleq \frac{\partial \mathbf{f}}{\partial \mathbf{x}^T} = \mathbf{matrix}_{i,j}[\frac{\partial f_i}{\partial x_j}] \; . \tag{409}$$

10.9 Second Derivative (Hesse Matrix)

$$\frac{\partial}{\partial \mathbf{x}^T}\left(\frac{\partial f}{\partial \mathbf{x}}\right) = \begin{pmatrix} \frac{\partial}{\partial x_1}(\frac{\partial f}{\partial x_1}) & \frac{\partial}{\partial x_2}(\frac{\partial f}{\partial x_1}) & \frac{\partial}{\partial x_3}(\frac{\partial f}{\partial x_1}) & \cdots & \frac{\partial}{\partial x_n}(\frac{\partial f}{\partial x_1}) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial}{\partial x_n}(\frac{\partial f}{\partial x_1}) & \frac{\partial}{\partial x_n}(\frac{\partial f}{\partial x_2}) & \frac{\partial}{\partial x_n}(\frac{\partial f}{\partial x_3}) & \cdots & \frac{\partial^2 f}{\partial x_n^2}) \end{pmatrix} \in \mathcal{R}^{n\times n} \; . \tag{410}$$

(The Hesse matrix is often written as $\frac{\partial^2 f}{\partial \mathbf{x}^2}$.)

10.10 Amplitude Scaling

Scaling of variables provides a numerically different (but physically identic) system of variables the range of which is smaller and yields less computational errors. Consider an arbitrary index of performance $f(\mathbf{x})$. Apply a similarity transform $\mathbf{y} = \mathbf{Q}\mathbf{x}$ with an identic index $f(\mathbf{x}) \equiv h(\mathbf{y})$.

$$\frac{\partial f}{\partial \mathbf{x}} = \frac{\partial h(\mathbf{y})}{\partial \mathbf{x}} \overset{(396)}{=} \left(\frac{\partial \mathbf{y}}{\partial \mathbf{x}^T}\right)^T \frac{\partial h}{\partial \mathbf{y}} = \mathbf{Q}^T \frac{\partial h}{\partial \mathbf{y}} \tag{411}$$

$$\frac{\partial f}{\partial \mathbf{x}^T} = \frac{\partial h(\mathbf{y})}{\partial \mathbf{y}^T}\mathbf{Q} \tag{412}$$

$$\frac{\partial}{\partial \mathbf{x}}\left(\frac{\partial f}{\partial \mathbf{x}^T}\right) \overset{(411)}{=} \mathbf{Q}^T \frac{\partial}{\partial \mathbf{y}}\left(\frac{\partial f}{\partial \mathbf{x}^T}\right) \overset{(412)}{=} \mathbf{Q}^T \frac{\partial}{\partial \mathbf{y}}\left(\frac{\partial h}{\partial \mathbf{y}^T}\mathbf{Q}\right) = \mathbf{Q}^T \frac{\partial^2 h}{\partial \mathbf{y}\partial \mathbf{y}^T}\mathbf{Q} \tag{413}$$

$$\frac{\partial^2 h}{\partial \mathbf{y}\partial \mathbf{y}^T} = \mathbf{Q}^{-T} \frac{\partial^2 f}{\partial \mathbf{x}\partial \mathbf{x}^T}\mathbf{Q}^{-1} . \tag{414}$$

The right-hand side of Eq.(414) should be close to the identity matrix to optimize the condition.

Example (10.10): $f(\mathbf{x}) = a^2 x_1^2 + b^2 x_2^2 \equiv h(\mathbf{y})$, where $a \gg b$. For f =constant there result ellipses in the (x_1, x_2)-plane.

$$\frac{\partial^2 f}{\partial \mathbf{x}\partial \mathbf{x}^T} = \frac{\partial}{\partial \mathbf{x}^T}\frac{\partial}{\partial \mathbf{x}} f = \frac{\partial}{\partial \mathbf{x}^T}\begin{pmatrix} 2a^2 x_1 & 0 \\ 0 & 2b^2 x_2 \end{pmatrix} = \begin{pmatrix} 2a^2 & 0 \\ 0 & 2b^2 \end{pmatrix} . \tag{415}$$

Selecting $\mathbf{Q} = \left(\begin{smallmatrix} a & 0 \\ 0 & b \end{smallmatrix}\right)$, from Eq.(414)

$$\frac{\partial^2 h}{\partial \mathbf{y}\partial \mathbf{y}^T} = \begin{pmatrix} 1/a & 0 \\ 0 & 1/b \end{pmatrix} \frac{\partial^2 f}{\partial \mathbf{x}\partial \mathbf{x}^T} \begin{pmatrix} 1/a & 0 \\ 0 & 1/b \end{pmatrix} = 2\,\mathbf{I}_2 \tag{416}$$

$$f(\mathbf{x}) = \mathbf{x}^T\mathbf{Q}^T\mathbf{Q}\mathbf{x} = \mathbf{x}^T\mathbf{Q}^2\mathbf{x} = \mathbf{y}^T\mathbf{Q}^{-T}\mathbf{Q}^2\mathbf{Q}^{-1}\mathbf{y} = \mathbf{y}^T\mathbf{y} \equiv h(\mathbf{y}) . \quad \square \tag{417}$$

In $\mathbf{y}$-space, $h(\mathbf{y})$ = constant is circle shaped (*Franklin, G.F., et al., 2002; Papageorgiou, M., 1991*).

10.11 Derivative of a Partitioned Matrix

For $\mathbf{H} \triangleq \mathbf{I}_2 + \mathbf{I}_2^{\swarrow}$, and using the rotation matrix of Eq.(33) with appropriate dimension,

$$\mathbf{A}_D \triangleq \begin{pmatrix} \mathbf{M}^{(r\times r)} & \mathbf{M} \\ \mathbf{M} & \mathbf{M} \end{pmatrix} = \mathbf{I}_2^{\swarrow} \otimes \mathbf{M} + \mathbf{I}_2 \otimes \mathbf{M} = \mathbf{H} \otimes \mathbf{M} \in \mathcal{R}^{2r\times 2r} \tag{418}$$

$$\frac{\partial \mathbf{A}_D}{\partial \mathbf{M}} \overset{(379)}{=} (\mathbf{I}_r \otimes \mathbf{U}_{2,r})[\frac{\partial \mathbf{M}}{\partial \mathbf{M}} \otimes \mathbf{H}](\mathbf{I}_r \otimes \mathbf{U}_{2,r}) \tag{419}$$

$$\overset{(41),(30)}{=} (\mathbf{I}_r \otimes \mathbf{U}_{2,r})[\bar{\mathbf{U}}_{r,r} \otimes (\mathbf{I}_2 + \mathbf{I}_2^{\swarrow})](\mathbf{I}_r \otimes \mathbf{U}_{2,r}) \in \mathcal{R}^{[(2r^2)\times(2r^2)]} . \tag{420}$$

11 Complex Closed-Loop Transfer Matrices and Their Derivatives

Argument and logarithmic expression for the following elementary correspondences $v = a + jb \equiv |v|e^{\arg(v)}$, $|v| = \sqrt{a^2 + b^2}$, $\arg(v) = \arctan(b/a)$

$$\arg F = \Im m\ [\ln F] \tag{421}$$

$$\frac{\partial \arg F}{\partial \omega} = \Im m\ [\frac{1}{F}\frac{\partial F}{\partial \omega}] \tag{422}$$

$$\Delta \arg F = \Im m\ [\frac{1}{F}\Delta F] \quad \text{for} \quad \Delta F \quad \text{small in norm sense.} \tag{423}$$

Evolving the symmetric expression

$$K^*\frac{\partial K}{\partial \mathbf{p}} + K\frac{\partial K^*}{\partial \mathbf{p}} = (\Re e\ K - j\Im m\ K)\frac{\partial \Re e\ K + j\Im m\ K}{\partial \mathbf{p}} + \ \ldots \tag{424}$$

$$= 2\ \Re e\ K\frac{\partial \Re e\ K}{\partial \mathbf{p}} + 2\ \Im m\ K\frac{\partial \Im m\ K}{\partial \mathbf{p}}\ . \tag{425}$$

Since

$$K\frac{\partial K^*}{\partial \mathbf{p}} = \left(K^*\frac{\partial K}{\partial \mathbf{p}}\right)^*\ , \tag{426}$$

from Eq.(424), left-hand side, one finds

$$\Re e\ [K^*\frac{\partial K}{\partial \mathbf{p}}] = \Re e\ K\frac{\partial \Re e\ K}{\partial \mathbf{p}} + \Im m\ K\frac{\partial \Im m\ K}{\partial \mathbf{p}}\ . \tag{427}$$

For complex $\mathbf{G}$, the derivatives of $\mathrm{tr}[\mathbf{G}^H\mathbf{G}]$ and $||\mathbf{G}||_F$ with respect to the real and imaginary part obey

$$\frac{\partial\ \mathrm{tr}\ \mathbf{G}^H\mathbf{G}}{\partial(\Re e\ \mathbf{G})} = 2\ \Re e\ \mathbf{G} \qquad \frac{\partial\ \mathrm{tr}\ \mathbf{G}^H\mathbf{G}}{\partial(\Im m\ \mathbf{G})} = 2\ \Im m\ \mathbf{G} \tag{428}$$

$$\frac{\partial||\mathbf{G}||_F}{\partial(\Re e\ \mathbf{G})} = \frac{\Re e\ \mathbf{G}}{||\mathbf{G}||_F} \qquad \frac{\partial||\mathbf{G}||_F}{\partial(\Im m\ \mathbf{G})} = \frac{\Im m\ \mathbf{G}}{||\mathbf{G}||_F}\ . \tag{429}$$

Generalizing Eq.(227) for $\mathbf{B}, \mathbf{C}$ complex and $\mathbf{A}, \mathbf{M}$ real

$$\frac{\partial}{\partial \mathbf{M}}||\mathbf{A} + \mathbf{BMC}||_F^2 = 2\Re e\ \mathbf{B}^H(\mathbf{A} + \mathbf{BMC})\mathbf{C}^H \tag{430}$$

since

$$\frac{\partial}{\partial \mathbf{M}}||\mathbf{A} + \mathbf{BMC}||_F^2 = \frac{\partial}{\partial \mathbf{M}}\mathrm{tr}[(\mathbf{A} + \mathbf{BMC})^H(\mathbf{A} + \mathbf{BMC})] \tag{431}$$

$$= \frac{\partial}{\partial \mathbf{M}}\mathrm{tr}\ [\mathbf{A}^H\mathbf{A} + \mathbf{C}^H\mathbf{M}^H\mathbf{B}^H\mathbf{A} + \mathbf{A}^H\mathbf{BMC} + \mathbf{C}^H\mathbf{M}^H\mathbf{B}^H\mathbf{BMC}] \tag{432}$$

$$= \frac{\partial}{\partial \mathbf{M}}\mathrm{tr}\ [\mathbf{A}^T\mathbf{B}^*\mathbf{MC}^* + \mathbf{A}^T\mathbf{BMC} + \mathbf{B}^H\mathbf{BMCC}^H\mathbf{M}^T] \tag{433}$$

$$= \mathbf{B}^H\mathbf{AC}^H + \mathbf{B}^T\mathbf{AC} + \mathbf{B}^T\mathbf{B}^*\mathbf{MCC}^T + \mathbf{B}^H\mathbf{BMCC}^H \tag{434}$$

$$= 2\Re e\ (\mathbf{B}^H\mathbf{AC}^H + \mathbf{B}^H\mathbf{BMCC}^H) = 2\Re e\ \mathbf{B}^H(\mathbf{A} + \mathbf{BMC})\mathbf{C}^H\ . \tag{435}$$

The combined expression leads to

$$
\begin{aligned}
\frac{\partial}{\partial \mathbf{M}} \|\mathbf{N}_1\mathbf{M}\mathbf{N}_2 + \mathbf{N}_3\mathbf{M}\mathbf{N}_4 + \mathbf{N}_5\|_F = \\
= 2\Re e\ \{\mathbf{N}_3^H(\mathbf{N}_1\mathbf{M}\mathbf{N}_2 + \mathbf{N}_3\mathbf{M}\mathbf{N}_4 + \mathbf{N}_5)\mathbf{N}_4^H \\
+ \mathbf{N}_1^H(\mathbf{N}_1\mathbf{M}\mathbf{N}_2 + \mathbf{N}_3\mathbf{M}\mathbf{N}_4 + \mathbf{N}_5)\mathbf{N}_2^H\}\ ,
\end{aligned}
\tag{436}
$$

where $\mathbf{N}_1....\mathbf{N}_5$ are complex and $\mathbf{M}$ is real. To continue this result,

$$
\frac{\partial}{\partial \mathbf{K}} \|\mathbf{F} + \mathbf{LKR} + \mathbf{R}^H\mathbf{K}^T\mathbf{L}^H\|_F^2 = 4\Re e\ \{\mathbf{L}^H(\mathbf{F} + \mathbf{LKR} + \mathbf{R}^H\mathbf{K}^T\mathbf{L}^H)\mathbf{R}^H\}\ . \tag{437}
$$

For the closed-loop matrix-valued transfer function

$$
T_{cl} \ \stackrel{\triangle}{=}\ \|\mathbf{T}_{ref}(s) - \mathbf{C}(s\mathbf{I} - \mathbf{A} - \mathbf{BK})^{-1}\mathbf{B}\|_F^2\ , \tag{438}
$$

using $\mathbf{H} \stackrel{\triangle}{=} s\mathbf{I} - \mathbf{A} - \mathbf{BK} = \mathbf{\Phi}_{cl}^{-1}(s)$ leads to (*Weinmann, A., 2001*)

$$
\frac{\partial T_{cl}}{\partial \mathbf{K}} \stackrel{(205)}{=} \frac{\partial \|\mathbf{T}_{ref}(s) - \mathbf{C}\mathbf{H}^{-1}\mathbf{B}\|_F^2}{\partial \mathbf{K}} \tag{439}
$$

$$
= -2\Re e\ \{\mathbf{B}^T\mathbf{H}^{-H}\mathbf{C}^T(\mathbf{T}_{ref} - \mathbf{C}\mathbf{H}^{-1}\mathbf{B})^H\mathbf{B}^T\mathbf{H}^{-H}\}\ . \tag{440}
$$

Consider the transfer matrix $\mathbf{F}_{St}(s)$ (or a related type) of the closed-loop system

$$
\mathbf{F}_{St}(s) = \mathbf{F}_1(s\mathbf{I} - \mathbf{A} - \mathbf{BK})^{-1}\mathbf{F}_2 \ \in \mathcal{C}^{r\times p} \qquad \mathbf{y} = \mathbf{F}_{St}\mathbf{w}_d \tag{441}
$$

$$
\mathbf{B} \in \mathcal{R}^{n\times m}\ ,\quad \mathbf{K} \in \mathcal{R}^{m\times n}\ ,\quad \mathbf{A} \in \mathcal{R}^{n\times n}\ ,\ \ \mathbf{w}_d \in \mathcal{C}^p\ ,\ \ \mathbf{y} \in \mathcal{C}^r\ ,\ \ \mathbf{F}_2 \in \mathcal{C}^{n\times p}\ \ \mathbf{F}_1 \in \mathcal{C}^{r\times n}\ . \tag{442}
$$

Abbreviating $\mathbf{F}_3 = \mathbf{F}_1^H\mathbf{F}_1,\ \mathbf{F}_4 = \mathbf{F}_2\mathbf{F}_2^H$, $\mathbf{H} = s\mathbf{I} - \mathbf{A} - \mathbf{BK} = \mathbf{\Phi}_{cl}^{-1}(s)$,

$$
\frac{\partial \|\mathbf{F}_{St}\|_F^2}{\partial \mathbf{K}} = 2\mathbf{B}^T\Re e\ \{\mathbf{H}^{-H}\mathbf{F}_3\mathbf{H}^{-1}\mathbf{F}_4\mathbf{R}^{-H}\}\ . \tag{443}
$$

The derivative of the prefilter $\mathbf{V} = -[\mathbf{C}(\mathbf{A}+\mathbf{BK})^{-1}\mathbf{B}]^{-1}$ with respect to $\mathbf{K}$ leads to the differential quotient as

$$
\frac{\partial \|[\mathbf{C}(\mathbf{A} + \mathbf{BK})^{-1}\mathbf{B}]^{-1}\|_F^2}{\partial \mathbf{K}} = 2\mathbf{T}_1^{-1}\mathbf{T}_1^{-T}\mathbf{B}^T(\mathbf{A} + \mathbf{BK})^{-T}\ , \tag{444}
$$

where $\mathbf{T}_1 \stackrel{\triangle}{=} \mathbf{C}(\mathbf{A} + \mathbf{BK})^{-1}\mathbf{B}$.

11.1 Condition for Additional Placement on a Hypersphere

As an additional condition, the controller parameters $\mathbf{p}$ should be located on a hypersphere with radius r_s, that is, $\sum_i(\mathbf{v}_i^T\mathbf{p})^2 = r_s^2$. Using $\mathbf{V} \triangleq \sum_i \mathbf{v}_i\mathbf{v}_i^T$, from

$$0.5\|\mathbf{W}\mathbf{p}\|_F^2 + 0.5\mu_s \sum_i (\mathbf{v}_i^T\mathbf{p})^2 + \boldsymbol{\mu}^T(\mathbf{M}\mathbf{p} - \mathbf{f}) \to \min_{\mathbf{p}} \tag{445}$$

$$\frac{\partial}{\partial \mathbf{p}} \rightsquigarrow \quad (\mathbf{W}^T\mathbf{W} + \mu_s\mathbf{V})\mathbf{p} + \mathbf{M}^T\boldsymbol{\mu} = \mathbf{0} \quad \rightsquigarrow \quad \mathbf{p} = -(\mathbf{W}^T\mathbf{W} + \mu_s\mathbf{V})^{-1}\mathbf{M}^T\boldsymbol{\mu} \ . \tag{446}$$

From $\mathbf{M}\mathbf{p} = \mathbf{f}$ and Eq.(446) one has $\boldsymbol{\mu} = -[\mathbf{M}(\mathbf{W}^T\mathbf{W} + \mu_s\mathbf{V})^{-1}\mathbf{M}^T]^{-1}\mathbf{f}$. Inserting into $\sum_i(\mathbf{v}_i^T\mathbf{p})^2 = r_s^2$ yields an equation in the unknown scalar μ_s

$$\sum_i \{\mathbf{v}_i^T(\mathbf{W}^T\mathbf{W} + \mu_s\mathbf{V})^{-1}\mathbf{M}^T[\mathbf{M}(\mathbf{W}^T\mathbf{W} + \mu_s\mathbf{V})^{-1}\mathbf{M}^T]^{-1}\mathbf{f}\} = r_s^2 \ . \tag{447}$$

12 Incremental Notation, Fréchet Derivatives

The incremental notation is addressed as advantageously applicable in control theory. The increment of a matrix $\mathbf{A}$ is termed $\Delta\mathbf{A}$.

For $\mathbf{A}, \mathbf{M} \in \mathcal{R}^{n\times m}$ and $f(\mathbf{M})$ given as

$$f(\mathbf{M}) \triangleq \mathrm{tr}[\mathbf{A}^T\mathbf{M}] \tag{448}$$

$$\Delta f(\mathbf{M}, \Delta\mathbf{M}) \triangleq f(\mathbf{M} + \Delta\mathbf{M}) - f(\mathbf{M}) = \mathrm{tr}[\mathbf{A}^T\Delta\mathbf{M}] = \mathrm{tr}[\Delta\mathbf{M}^T \cdot \mathbf{A}] \tag{449}$$

and

$$\Delta f(\mathbf{M}) = \Delta\mathrm{tr}[\mathbf{A}^T\mathbf{M}] = \mathrm{tr}[\mathbf{A} \cdot \Delta\mathbf{M}^T] \quad \rightsquigarrow \quad \frac{\partial f(\mathbf{M})}{\partial \mathbf{M}} = \mathbf{A} \ . \tag{450}$$

The factor of the increment $\Delta\mathbf{M}$, having been transposed, equals the derivative.

$$\Delta f(\mathbf{M}) = \sum_{i,j} \frac{\partial f}{\partial M_{ij}} \Delta M_{ij} = \mathrm{tr}[\frac{\partial f}{\partial \mathbf{M}} \Delta\mathbf{M}^T] = \mathrm{tr}[\frac{\partial f}{\partial \mathbf{M}^T} \Delta\mathbf{M}] \ . \tag{451}$$

There are two methodes for some more complicated (e.g., nonlinear) conditions $f(\mathbf{L})$. For the statement $f(\mathbf{L})$ derive $f(\mathbf{L} + \Delta\mathbf{L})$, especially using Taylor expansion, and $\Delta f = f(\mathbf{L} + \Delta\mathbf{L}) - f(\mathbf{L})$:

- If $\Delta f(\mathbf{L})$ can be rewritten to $\Delta f(\mathbf{L}) \triangleq \mathrm{tr}[\mathbf{A}\Delta\mathbf{L}]$, then $\frac{\partial f}{\partial \mathbf{L}} \overset{(450)}{=} \mathbf{A}^T$.
- If $\Delta f(\mathbf{L})$ can be rewritten to $\Delta f(\mathbf{L}) = \mathrm{tr}[\mathbf{R}\Delta\mathbf{T}]$, and if instead of the matrix $\Delta\mathbf{T}$ only the vector $\mathrm{col}(\Delta\mathbf{T}) = \mathbf{M}\ \mathrm{col}\Delta\mathbf{L}$ is available, then (see, e.g., Section 31)

$$\begin{aligned} \Delta f(\mathbf{L}) &= \mathrm{tr}[\mathbf{R}\Delta\mathbf{T}] \overset{(49)}{=} (\mathrm{col}\ \mathbf{I}_n)^T\mathrm{col}\ (\mathbf{R}\Delta\mathbf{T}) && (452)\\ &= (\mathrm{col}\ \mathbf{I}_n)^T(\mathbf{I}_n \otimes \mathbf{R})\mathrm{col}\Delta\mathbf{T} = (\mathrm{col}\ \mathbf{I}_n)^T(\mathbf{I}_n \otimes \mathbf{R})\ \mathbf{M}\ \mathrm{col}\Delta\mathbf{L} && (453)\\ \frac{\partial f(\mathbf{L})}{\partial \mathrm{col}\ \mathbf{L}} &= \mathbf{M}^T(\mathbf{I}_n \otimes \mathbf{R}^T)\ \mathrm{col}\ \mathbf{I}_n \ . && (454) \end{aligned}$$

- If $\Delta\mathbf{T} = \mathbf{B}\ \Delta\mathbf{K}$ from $\mathbf{T}^{[n\times n]} = \mathbf{B}^{[n\times m]}\mathbf{K}^{[m\times n]}$ and $\mathbf{B}$ constant, then

$$\frac{\partial\mathbf{T}}{\partial\mathbf{K}} = \frac{\partial\mathbf{BK}}{\partial\mathbf{K}} \overset{(351)}{=} (\mathbf{I}_m \otimes \mathbf{B})\bar{\mathbf{U}}_{m,n}^{[m^2\times n^2]}) \ . \tag{455}$$

The result is a specific arrangement of the coefficients of the matrix $\mathbf{B}$; of course independent of $\mathbf{K}$. E.g., for $\mathbf{B} = (1\ \ 2\ \ 5)^T$, $\mathbf{K} = (0.4\ \ 0.2\ \ 0.9)$

```
B        dT/dK                at the operation point B*K
1    1 0 0 0 1 0 0 0 1        0.4000    0.2000    0.9000
2    2 0 0 0 2 0 0 0 2        0.8000    0.4000    1.8000
5    5 0 0 0 5 0 0 0 5        2.0000    1.0000    4.5000
```

12.1 Increments and Differential Quotients

$$\text{For}\quad v(a) \qquad \Delta v(a) \doteq \frac{\partial v}{\partial a}\Delta a \ , \tag{456}$$

$$\text{for}\quad v(\mathbf{A}) \qquad \Delta v(\mathbf{A}) \doteq \sum_{ij}(\frac{\partial v}{\partial\mathbf{A}})_{ij}\Delta A_{ij} = (\text{col}\frac{\partial v}{\partial\mathbf{A}})^T\text{col}\Delta\mathbf{A} \ , \tag{457}$$

$$\text{for}\quad \mathbf{v}(\mathbf{A}) \qquad \Delta\mathbf{v}(\mathbf{A}) \doteq (\frac{\partial\mathbf{v}^T}{\partial\text{col}\mathbf{A}})^T\text{col}\Delta\mathbf{A} = (\frac{\partial\mathbf{v}}{\partial(\text{col}\mathbf{A})^T})\text{col}\Delta\mathbf{A} \ . \tag{458}$$

Example (12.1;1): $\mathbf{v} = (v_1 \vdots v_2)^T$; $v_1 \overset{\triangle}{=} \text{tr}[\mathbf{A}]$; $v_2 \overset{\triangle}{=} (\text{tr}[\mathbf{A}])^2$. The two versions are

$$\frac{\partial v_1}{\partial\mathbf{A}} = \mathbf{I} \ , \qquad \frac{\partial v_1}{\partial\text{col}\mathbf{A}} = \text{col}\frac{\partial v_1}{\partial\mathbf{A}} = \text{col}\ \mathbf{I} \tag{459}$$

$$\Delta v_1 \overset{(457)}{=} \sum_{ij}\frac{\partial v_1}{\partial A_{ij}}\Delta A_{ij} = \sum_{ij}(\text{col}\ \mathbf{I})_{ij}\Delta A_{ij} = \sum_{ij}\Delta A_{ij} = \text{tr}[\Delta\mathbf{A}] \tag{460}$$

$$\Delta v_2 \overset{(457)}{=} \sum_{ij}\frac{\partial v_2}{\partial A_{ij}}\Delta A_{ij} = \sum_{ij}2(\text{tr}\mathbf{A})(\text{col}\ \mathbf{I})_{ij}\Delta A_{ij} = 2(\text{tr}\mathbf{A})\text{tr}[\Delta\mathbf{A}] \tag{461}$$

$$\Delta\mathbf{v} \overset{(458)}{=} \begin{pmatrix} (\text{col}\ \mathbf{I})^T \\ 2\text{tr}[\mathbf{A}](\text{col}\mathbf{I})^T \end{pmatrix}\text{col}\Delta\mathbf{A} = \begin{pmatrix} \sum_{ij}\Delta A_{ij} \\ 2\text{tr}[\mathbf{A}]\sum_{ij}\Delta A_{ij} \end{pmatrix} = \begin{pmatrix} \text{tr}[\Delta\mathbf{A}] \\ 2\text{tr}[\mathbf{A}]\ \text{tr}[\Delta\mathbf{A}] \end{pmatrix} \ . \quad \square \tag{462}$$

For changes in ΔM_{ij}, only,

$$\Delta c \doteq (\frac{\partial c}{\partial M_{ij}})\Delta M_{ij} \overset{(6)}{=} \mathbf{e}_i^T\frac{\partial c}{\partial\mathbf{M}}\mathbf{e}_j\Delta M_{ij} \tag{463}$$

Referring to Eq.(349), if only ΔM_{kl} varies

$$(\Delta G)_{ij} \doteq \frac{\partial G_{ij}}{\partial M_{kl}}\Delta M_{kl} \overset{(6)}{=} (\mathbf{e}_{(k-1)r+k_1}^{[nr\times 1]})^T\ \frac{\partial\mathbf{G}}{\partial\mathbf{M}}\ \mathbf{e}_{(l-1)r+l_1}^{[ms\times 1]}\Delta M_{kl} \ . \tag{464}$$

Example (12.1;2): $\mathbf{G} = \mathbf{M}\mathbf{M}^T$, $r = 2$; $s = 3$; $k = 2$; $l = 3$; $k_1 = 1$; $l_1 = 2$; $(k_1 \in \{1, n\}, l_1 \in \{1, m\})$; $\Delta M_{kl} \stackrel{\triangle}{=} 10^{-3}\mathbf{E}_{kl}^{[r\times s]}$; $\mathbf{G} \in \mathcal{R}^{n\times m}$, $\mathbf{M} \in \mathcal{R}^{r\times s}$, $\frac{\partial \mathbf{G}}{\partial \mathbf{M}} \in \mathcal{R}^{nr\times ms}$; $n = m = r$;

$$\mathbf{M} \stackrel{\triangle}{=} \begin{pmatrix} 27.8498 & 95.7507 & 15.7613 \\ 54.6882 & 96.4889 & 97.0593 \end{pmatrix}, \quad \Delta\mathbf{G} = \begin{pmatrix} 0 & 0.0158 \\ 0.0158 & 0.1941 \end{pmatrix} \tag{465}$$

$$\frac{\partial \mathbf{G}}{\partial \mathbf{M}} \stackrel{(351)}{=} \begin{pmatrix} 55.6996 & 54.6882 & 191.5014 & 96.4889 & 31.5226 & 97.0593 \\ 54.6882 & 0 & 96.4889 & 0 & 97.0593 & 0 \\ 0 & 27.8498 & 0 & 95.7507 & 0 & \mathbf{15.7613} \\ 27.8498 & 109.3764 & 95.7507 & 192.9778 & 15.7613 & 194.1186 \end{pmatrix} \tag{466}$$

$$\frac{\partial \mathbf{G}_{k_1,l_1}}{\partial \mathbf{M}_{k,l}} = 15.7613 \qquad \Delta\mathbf{G}(k_1, l_1) = 0.0158\ . \qquad \square \tag{467}$$

12.2 Increment of the Frobenius Norm

$$\Delta||\mathbf{M}||_F \stackrel{(212)}{=} \frac{1}{||\mathbf{M}||_F}\text{tr}[\mathbf{M}^T \cdot \Delta\mathbf{M}] \tag{468}$$

Suppose real matrices $\mathbf{K}$, $\mathbf{L}$, $\mathbf{R}$ with appropriate dimension, the increment $\Delta\mathbf{K}$ of $\mathbf{K}$ causes an increment (*Vetter, W.J., 1971*)

$$\begin{aligned} \Delta||\mathbf{LKR}||_F^2(\mathbf{K}, \Delta\mathbf{K}) &\stackrel{\triangle}{=} ||\mathbf{L}(\mathbf{K} + \Delta\mathbf{K})\mathbf{R}||_F^2 - ||\mathbf{LKR}||_F^2 \\ &= \text{tr}\ \{2\mathbf{L}^T\mathbf{LKRR}^T\Delta\mathbf{K}^T\}\ . \end{aligned} \tag{469}$$

Hence,

$$\frac{\partial}{\partial \mathbf{K}}||\mathbf{LKR}||_F^2 = 2\mathbf{L}^T\mathbf{LKRR}^T\ . \tag{470}$$

The prefactor matrix in Eq.(469) corresponds to the differential quotient with respect to $\mathbf{K}$.

The differential or increment of $||\mathbf{A} - \mathbf{KC}||_F^2$ with respect to $\mathbf{A}$ is

$$\Delta||\mathbf{A} - \mathbf{KC}||_F^2(\mathbf{A}, \Delta\mathbf{A}) \stackrel{\triangle}{=} ||\mathbf{A} + \Delta\mathbf{A} - \mathbf{KC}||_F^2 - ||\mathbf{A} - \mathbf{KC}||_F^2 = \tag{471}$$

$$= 2\text{tr}\ \{[\mathbf{A}^T - (\mathbf{KC})^T]\Delta\mathbf{A}\}. \tag{472}$$

12.3 Increment of the Trace and Determinant

Important relations are (*Vetter, J.W., 1971*)

$$\Delta M_{ij} = \Delta(\mathbf{e}_i^T\mathbf{M}\mathbf{e}_j) = \text{tr}[\mathbf{E}_{ij} \cdot \Delta\mathbf{M}] \equiv \Delta M_{ij} \tag{473}$$

$$\Delta\text{tr}[\mathbf{M}] = \text{tr}[\Delta\mathbf{M}]\ , \qquad \Delta\text{tr}[\mathbf{M}^n] \stackrel{(208)}{=} n\ \text{tr}[\mathbf{M}^{n-1} \cdot \Delta\mathbf{M}] \tag{474}$$

$$\Delta\text{tr}[e^{\mathbf{M}}] \stackrel{(213)}{=} \text{tr}[e^{\mathbf{M}} \cdot \Delta\mathbf{M}] \tag{475}$$

$$\Delta \det \mathbf{M} = (\det \mathbf{M})\ \mathrm{tr}[\mathbf{M}^{-1} \cdot \Delta\mathbf{M}] = \mathrm{tr}[(\mathbf{adjM}) \cdot \Delta\mathbf{M}] \tag{476}$$

$$\Delta \det(\mathbf{M}^n) \overset{(189)}{=} n(\det \mathbf{M})^n \mathrm{tr}[\mathbf{M}^{-1}\Delta\mathbf{M}] = n(\det \mathbf{M})^{n-1}\mathrm{tr}[(\mathbf{adjM}) \cdot \Delta\mathbf{M}] \tag{477}$$

$$\Delta \det e^{\mathbf{M}} \overset{(190)}{=} (\det e^{\mathbf{M}})\ \mathrm{tr}[\Delta\mathbf{M}] \tag{478}$$

$$\Delta\lambda_i[\mathbf{M}] \overset{(169)}{=} \frac{\{\mathrm{tr}[\mathbf{adj}(\lambda_i \mathbf{I} - \mathbf{M})\}\Delta\mathbf{M}]}{\mathrm{tr}[\mathbf{adj}(\lambda_i \mathbf{I} - \mathbf{M})]}\ . \tag{479}$$

12.4 Taylor Expansion of the Determinant

$$\begin{aligned} \det(\mathbf{A} + \mathbf{V}\Delta p)\Big|_{\Delta p \ll 1} &\doteq \det \mathbf{A} + \frac{\partial \det(\mathbf{A} + \mathbf{V}\Delta p)}{\partial \Delta p}\Delta p & (480)\\ &\overset{(130)}{=} \det \mathbf{A} + \mathrm{tr}[\mathbf{VadjA}]\Delta p & (481)\\ &\doteq \det \mathbf{A} + \mathrm{tr}[\mathbf{VA}^{-1}] \cdot \det \mathbf{A} \cdot \Delta p\ . & (482) \end{aligned}$$

$$\begin{aligned} \det(\mathbf{A} + \mathbf{E}_{ij}\Delta p)\Big|_{\Delta p \ll 1} &\doteq \det \mathbf{A} + \frac{\partial \det(\mathbf{A} + \mathbf{E}_{ij}\Delta p)}{\partial \Delta p}\Delta p & (483)\\ &\doteq \det \mathbf{A} + \mathrm{tr}[\mathbf{E}_{ij}\mathbf{adjA}]\Delta p & (484)\\ &\doteq \det \mathbf{A} + (\mathbf{adjA})_{ji} \cdot \Delta p\ . & (485) \end{aligned}$$

$$\begin{aligned} \det(\mathbf{A} + \Delta\mathbf{A})\Big|_{\|\Delta\mathbf{A}\|_F \ll 1} &\doteq \det \mathbf{A} + \sum_{ij} (\mathbf{adjA})_{ji}\Delta A_{ij} & (486)\\ &\doteq \det \mathbf{A} + \sum_{ij} [(\mathbf{adjA})^T \cdot *\ \Delta\mathbf{A}]_{ij} & (487)\\ &\doteq \det \mathbf{A} + [\mathrm{col}(\mathbf{adjA}^T)]^T \cdot \mathrm{col}\Delta\mathbf{A}\ . & (488) \end{aligned}$$

12.5 Increment of the Inverse

$$(\mathbf{X} + \Delta\mathbf{X})^{-1} \doteq \mathbf{X}^{-1} - \mathbf{X}^{-1}\Delta\mathbf{X} \cdot \mathbf{X}^{-1} + \mathbf{X}^{-1}\Delta\mathbf{X} \cdot \mathbf{X}^{-1}\Delta\mathbf{X} \cdot \mathbf{X}^{-1} - + \ldots \tag{489}$$

$$\Delta(\mathbf{X}^{-1}) \doteq -\mathbf{X}^{-1}(\Delta\mathbf{X})\mathbf{X}^{-1}\ . \tag{490}$$

If $\mathbf{A} = \mathbf{GX}^{-1}$ and both $\mathbf{G}$ and $\mathbf{X}$ are variable

$$\begin{aligned} \Delta\mathbf{A} &= (\Delta\mathbf{G})\mathbf{X}^{-1} + \mathbf{G}\ \Delta(\mathbf{X}^{-1}) = (\Delta\mathbf{G})\mathbf{X}^{-1} - \mathbf{GX}^{-1}(\Delta\mathbf{X})\mathbf{X}^{-1} & (491)\\ &= (\Delta\mathbf{G})\mathbf{X}^{-1} - \mathbf{A}(\Delta\mathbf{X})\mathbf{X}^{-1} = (\Delta\mathbf{G} - \mathbf{A}\ \Delta\mathbf{X})\mathbf{X}^{-1}\ , & (492) \end{aligned}$$

leading to

$$\Delta\|\mathbf{A} - \mathbf{KC}\|_F^2(\mathbf{A}, \Delta\mathbf{X}) = 2\mathrm{tr}\ \{[\mathbf{A}^T - (\mathbf{KC})^T](\Delta\mathbf{G} - \mathbf{A}\ \Delta\mathbf{X})\mathbf{X}^{-1}\}\ . \tag{493}$$

12.6 Increment of the Adjoint Matrix and the Matrix Exponential

$$\mathbf{adj}(\mathbf{X}^{[n\times n]}+\Delta\mathbf{X}) = (\mathbf{X}+\Delta\mathbf{X})^{-1}\det(\mathbf{X}+\Delta\mathbf{X}) \tag{494}$$

$$\overset{(489),(482)}{=} (\mathbf{X}^{-1}-\mathbf{X}^{-1}\Delta\mathbf{X}\cdot\mathbf{X}^{-1})[\det\mathbf{X}+\mathrm{tr}(\Delta\mathbf{X}\cdot\mathbf{X}^{-1})\det\mathbf{X}]$$

$$= \mathbf{adj}(\mathbf{X})\cdot[\mathbf{I}_n-\Delta\mathbf{X}\cdot\mathbf{X}^{-1}+\mathbf{I}_n\mathrm{tr}(\mathbf{X}^{-1}\Delta\mathbf{X})] \tag{495}$$

$$\Delta\mathbf{adj}(\mathbf{X})(\mathbf{X},\Delta\mathbf{X}) = \mathbf{I}_n\mathrm{tr}(\mathbf{X}^{-1}\Delta\mathbf{X})-\Delta\mathbf{X}\cdot\mathbf{X}^{-1}\ . \tag{496}$$

Furthermore for the increments of the matrix exponential

$$g(\mathbf{M})\overset{\triangle}{=}\mathrm{tr}[e^{\mathbf{M}}] \tag{497}$$

$$\Delta g(\mathbf{M},\Delta\mathbf{M})=\mathrm{tr}[e^{\mathbf{M}}\Delta\mathbf{M}] \tag{498}$$

$$\frac{\partial g(\mathbf{M})}{\partial\mathbf{M}}=e^{\mathbf{M}^T}\ . \tag{499}$$

Finally (*Levine, W.S., and Athans,M., 1970; Bellman, R., 1960*)

$$\mathbf{h}(\mathbf{M})\overset{\triangle}{=}e^{\mathbf{BM}t} \tag{500}$$

$$\Delta\mathbf{h}(\mathbf{M},\Delta\mathbf{M})=\int_0^t e^{\mathbf{BM}(t-\sigma)}\mathbf{B}\Delta\mathbf{M}e^{\mathbf{BM}\sigma}d\sigma\ . \tag{501}$$

13 Miscellaneous Correspondences

The Eqs.(454),(455), and Eq.(503) through (506), are susceptible to mistakes and confusion. Hence, the variables are termed with their dimensions in detail. The symbol $\times$ is used for denoting the dimensions of a matrix and a vector. We develop the differential operations referring to the simple matrix product $\mathbf{T}=\mathbf{BK}$ as used for input matrix and state controller matrix. The function to be analyzed is $\mathbf{B}$, operating as a simple matrix-valued factor of $\mathbf{K}$.

$$\mathbf{T}^{[n\times n]}=\mathbf{B}^{[n\times m]}\mathbf{K}^{[m\times n]} \tag{502}$$

$$\frac{\partial\mathbf{BK}}{\partial\mathbf{K}}\overset{(356)}{=}(\mathbf{I}_m\otimes\mathbf{B})\frac{\partial\mathbf{K}}{\partial\mathbf{K}}\overset{(41)}{=}(\mathbf{I}_m\otimes\mathbf{B})\bar{\mathbf{U}}_{m,n}^{[m^2\times n^2]} \tag{503}$$

$$(\mathrm{col}\ \Delta\mathbf{T})^{[n^2\times 1]}\overset{(51)}{=}\underbrace{(\mathbf{I}_n\otimes\mathbf{B})}_{\overset{\triangle}{=}\mathbf{M}^{[n^2\times mn]}}\left(\mathrm{col}\ \Delta\mathbf{K}^{[m\times n]}\right)^{[mn\times 1]} \tag{504}$$

$$\frac{\partial\,(\mathrm{col}\ \mathbf{T})^{[n^2\times 1]}}{\partial\,(\mathrm{col}\ \mathbf{K})^{[mn\times 1]}}\overset{(356)}{=}(\mathbf{I}_{mn}\otimes\mathbf{M})\bar{\mathbf{U}}_{mn,1}^{[(mn)^2\times 1]}=(\mathbf{I}_{mn}\otimes\mathbf{I}_n\otimes\mathbf{B})\bar{\mathbf{U}}_{mn,1}^{[(mn)^2\times 1]} \tag{505}$$

$$\left(\frac{\partial\mathrm{col}\ \mathbf{T}}{\partial\mathrm{col}\ \mathbf{K}}\right)^{[mn^3\times 1]} = (\mathbf{I}_{mn^2}\otimes\mathbf{B})^{[mn^3\times(mn)^2]}\,\bar{\mathbf{U}}_{mn,1}^{[(mn)^2\times 1]}\ , \tag{506}$$

having used $\mathbf{I}_{mn} \otimes \mathbf{I}_n = \mathbf{I}_{mn^2}$.

From Lyapunov equation we get an expression col$\mathbf{T} = \mathbf{M}$col$\mathbf{K}$ as Eq.(504), what we need is an expression Eq.(503), i.e., we have to apply a structural change. Via Eq.(505)

$$\frac{\mathrm{col}\mathbf{T}}{\mathrm{col}\mathbf{K}} = (\mathbf{I}_{mn} \otimes \mathbf{I}_n \otimes \mathbf{B})\bar{\mathbf{U}}_{mn,1}^{[(mn)^2 \times 1]} \tag{507}$$

we change the order by applying, first, loc$_m$, and, second, bloc loc$_m$ of a scheme of matrices $n \times n$.

In the special case, where $\mathbf{M}$ can be developed in $\mathbf{I}_n \otimes \mathbf{B}$, we can directly find the map

$$(\mathbf{I}_n \otimes \mathbf{B}) \;\mapsto\; (\mathbf{I}_m \otimes \mathbf{B})\bar{\mathbf{U}}_{m,n}^{[m^2 \times n^2]} \; . \tag{508}$$

Example (13;1): For illustration, a special case is shown in an example with Fig. 4. □

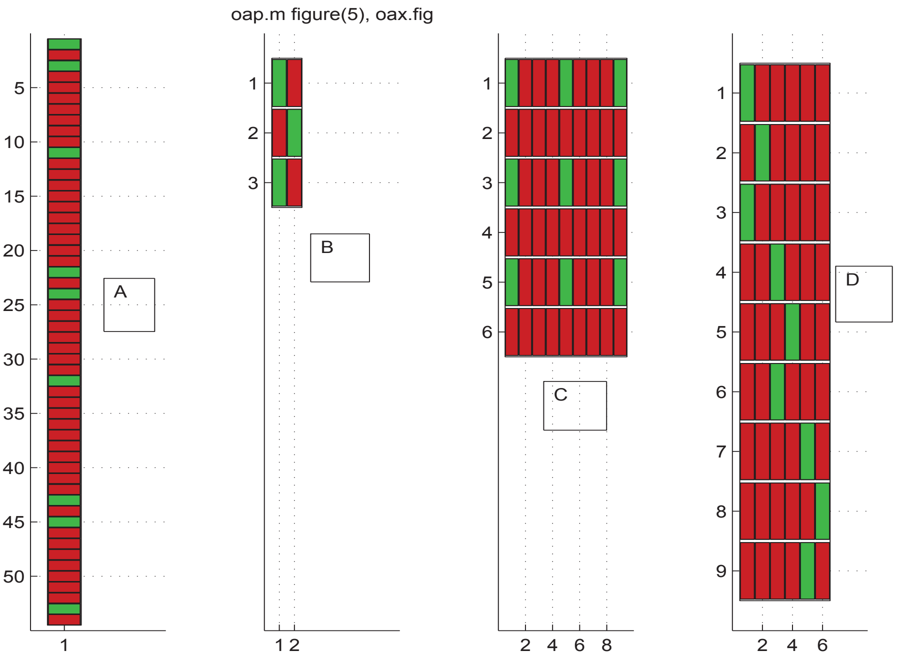

Figure 4: Graphic interpretation of the coefficient matrix details for $n = 3,\ m = 2$,

Subfigure A: $\frac{\partial col\mathbf{T}}{\partial col\mathbf{K}} = \boxed{(\mathbf{I}_{18} \otimes \mathbf{B})\bar{\mathbf{U}}_{6,1}}$; Subfigure B: $\mathbf{B} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \end{pmatrix}$;

Subfigure C: $\frac{\partial \mathbf{T}}{\partial \mathbf{K}} = \boxed{(\mathbf{I}_2 \otimes \mathbf{B})\bar{\mathbf{U}}_{2,3}}$; Subfigure D: Δcol $\mathbf{T} = \boxed{(\mathbf{I}_3 \otimes \mathbf{B})}\ \Delta$col $\mathbf{K}$.

Further Examples (13;2): For a single-input single-output controller $\mathbf{K} \in \mathcal{R}^{1\times n}$ one has

$$\begin{aligned} \frac{\partial \mathbf{K}^T}{\partial \mathbf{K}} &\overset{(357)}{=} \mathbf{U}_{1,n}(\mathbf{I}_n \otimes 1) = \mathbf{U}_{1,n} \; ; & \frac{\partial \text{col}\mathbf{K}^T}{\partial \text{col}\mathbf{K}} &\overset{(335)}{=} \text{col}\mathbf{I}_n \; ; \quad \text{col}\mathbf{K}^T \overset{(57)}{=} \mathbf{U}_{1,n}\text{col}\mathbf{K} \; . \\ \frac{\partial \mathbf{K}}{\partial \mathbf{K}} &= \bar{\mathbf{U}}_{1,n} \; ; & \frac{\partial \text{col}\mathbf{K}}{\partial \mathbf{K}} &= (\mathbf{I}_{mn^2} \otimes 1)\bar{\mathbf{U}}_{n,1} = \text{col}\mathbf{I}_n \; . \end{aligned} \tag{509}$$

For scalar numerators and $\mathbf{K}^{[m\times 1]}$

$$\text{col}\frac{\partial r}{\partial \mathbf{K}^{[m\times n]}} = \frac{\partial r}{\partial \text{col }\mathbf{K}} \; ; \;\; \frac{\partial r}{\partial \mathbf{K}^{[m\times n]}} = \text{loc}_m\left(\frac{\partial r}{\partial \text{col }\mathbf{K}}\right) \; ; \;\; \frac{\partial \mathbf{b}^T\mathbf{k}}{\partial \mathbf{k}} = \mathbf{b}; \;\; \Delta(\mathbf{b}^T\mathbf{k}) = \mathbf{b}^T\Delta\mathbf{k} \, . \tag{510}$$

For $n = 1$, $\mathbf{B}^{[1\times m]} = \mathbf{b}^T$, $\mathbf{K} = \mathbf{k}^{[m\times 1]}$, $T = \mathbf{b}^T\mathbf{k}$, Eq.(503) and Eq.(506) yield the same result

$$\frac{\partial T}{\partial \mathbf{k}} \overset{(503),(506)}{=} \frac{\partial \text{col}T}{\partial \text{col}\mathbf{k}} = (\mathbf{I}_m \otimes \mathbf{b}^T)\bar{\mathbf{U}}_{m,1} = \mathbf{b} \; . \tag{511}$$

In Eq.(511) a formula surfaces for simply transposing a row to a column, just as Eq.(62) or Eq.(63).

For $\mathbf{B}^{[n\times m]}$, $\mathbf{K} = \mathbf{k}^{[m\times 1]}$,

$$\frac{\partial (\mathbf{B}\mathbf{k})^{[n\times 1]}}{\partial \mathbf{k}} = (\mathbf{I}_m \otimes \mathbf{B})\bar{\mathbf{U}}_{m,1} \; , \qquad \Delta \text{ col } (\mathbf{B}\mathbf{k}) = \mathbf{B} \text{ col}\Delta\mathbf{k} \; . \qquad \square \tag{512}$$

Numerically checking the presented results or new ones can be carried out as follows. For $\frac{\partial f}{\partial \mathbf{M}}$, the (i, j) position in the matrix differential is $\frac{\partial f}{\partial M_{ij}}$. It is approximated by

$$\frac{\partial f}{\partial M_{ij}} = \lim_{\Delta M_{ij}\to 0} \frac{\Delta f(\mathbf{M}, \Delta M_{ij})}{\Delta M_{ij}} = \lim_{\Delta M_{ij}\to 0} \frac{f(\mathbf{M}) - f(\mathbf{M} + \mathbf{E}_{ij} \cdot \Delta M_{ij})}{\Delta M_{ij}} \; . \tag{513}$$

14 Derivatives with Respect to Time

Total derivative if c depends on $\mathbf{x}(t)$ and on t, additionally,

$$\frac{d}{dt}c[\mathbf{x}(t), t] = \frac{d\mathbf{x}^T(t)}{dt}\frac{\partial c}{\partial \mathbf{x}} + \frac{\partial c}{\partial t} \; . \tag{514}$$

Vector (column) function $\mathbf{z}$ dependent on $\mathbf{x}(t)$ and t: $\mathbf{z} \in \mathcal{R}^m$, $\mathbf{x}(t) \in \mathcal{R}^q$, $\mathbf{M} := t$ (*Vetter, W.J., 1970*)

$$\frac{d\mathbf{z}[\mathbf{x}(t)]}{dt} = \Big(\frac{d\mathbf{x}^T}{dt} \otimes \mathbf{I}_m\Big)\Big(1 \otimes \frac{\partial \mathbf{z}}{\partial \mathbf{x}}\Big) = \Big(\frac{d\mathbf{x}^T}{dt} \otimes \mathbf{I}_m\Big)\frac{\partial \mathbf{z}}{\partial \mathbf{x}} \tag{515}$$

$$\frac{d\mathbf{z}[\mathbf{x}(t), t]}{dt} = \Big(\frac{d\mathbf{x}^T}{dt} \otimes \mathbf{I}_m\Big)\frac{\partial \mathbf{z}}{\partial \mathbf{x}} + \frac{\partial \mathbf{z}}{\partial t} \; . \tag{516}$$

Row function $\mathbf{y}^T$ dependent on $\mathbf{x}(t)$ and t with $\mathbf{y} \in \mathcal{R}^n$

$$\frac{d\mathbf{y}^T[\mathbf{x}(t),t]}{dt} = \Big(\frac{d\mathbf{x}^T}{dt} \otimes \mathbf{I}_1\Big)\Big(\mathbf{I}_1 \otimes \frac{\partial \mathbf{y}^T}{\partial \mathbf{x}}\Big) = \frac{d\mathbf{x}^T}{dt}\frac{\partial \mathbf{y}^T}{\partial \mathbf{x}} \ . \tag{517}$$

Matrix-valued function $\mathbf{A}$ dependent on the vector $\mathbf{x}(t) \in \mathcal{R}^q$, $\mathbf{A} \in \mathcal{R}^{n\times m}$,

$$\frac{d\mathbf{A}[\mathbf{x}(t),t]}{dt} = \Big(\frac{d\mathbf{x}^T}{dt} \otimes \mathbf{I}_n\Big)\Big(\mathbf{I}_1 \otimes \frac{\partial \mathbf{A}}{\partial \mathbf{x}}\Big)+\frac{\partial \mathbf{A}}{\partial t} = \Big(\frac{d\mathbf{x}^T}{dt} \otimes \mathbf{I}_n\Big)\frac{\partial \mathbf{A}}{\partial \mathbf{x}} + \frac{\partial \mathbf{A}}{\partial t} \ . \tag{518}$$

Example (14;1) : For $\mathbf{y} \in \mathcal{R}^n$, $\mathbf{z} \in \mathcal{R}^m$, $\mathbf{x} \in \mathcal{R}^q$, $\mathbf{A} \in \mathcal{R}^{n\times m}$, derivatives of the product $c = \mathbf{y}^T[\mathbf{x}(t),t]\ \mathbf{A}[\mathbf{x}(t),t]\ \mathbf{z}[\mathbf{x}(t),t]$ (*Vetter, W.J., 1970*) are

$$\frac{d}{dt}\ \mathbf{y}^T[\mathbf{x}(t),t]\ \mathbf{A}[\mathbf{x}(t),t]\ \mathbf{z}[\mathbf{x}(t),t] = \frac{d\mathbf{y}^T}{dt}\mathbf{A}\mathbf{z} + \mathbf{y}^T\frac{d\mathbf{A}}{dt}\mathbf{z} + \mathbf{y}^T\mathbf{A}\frac{d\mathbf{z}}{dt} \ . \tag{519}$$

$$\frac{\partial c}{\partial \mathbf{y}} = \mathbf{A}\mathbf{z} \qquad \frac{\partial c}{\partial \mathbf{z}} = (\mathbf{y}^T\mathbf{A})^T = \mathbf{A}^T\mathbf{y} \tag{520}$$

$$\frac{\partial c}{\partial \mathbf{A}} = \frac{\partial}{\partial \mathbf{A}}\mathrm{tr}\ (\mathbf{y}^T\mathbf{A}\mathbf{z}) = \frac{\partial}{\partial \mathbf{A}}\mathrm{tr}\ (\mathbf{z}\mathbf{y}^T\mathbf{A}) = (\mathbf{z}\mathbf{y}^T)^T = \mathbf{y}\mathbf{z}^T \ . \ \square \tag{521}$$

Example (14;2):

$$\frac{\partial}{\partial \mathbf{x}^T}\ \mathbf{x}^T\mathbf{Q}(\mathbf{x})\ \mathbf{x} = \mathbf{x}^T[\mathbf{Q}^T(\mathbf{x}) + \mathbf{Q}(\mathbf{x}) + \frac{\partial \mathbf{Q}(\mathbf{x})}{\partial \mathbf{x}^T}(\mathbf{I}_q \otimes \mathbf{x})] \qquad \mathbf{x} \in \mathcal{R}^q \ . \ \square \tag{522}$$

Example (14;3):

$$\frac{d}{dt}e^{\mathbf{A}t} = \mathbf{A}e^{\mathbf{A}t} = e^{\mathbf{A}t}\mathbf{A} \qquad \text{and} \qquad \int_0^t e^{\mathbf{A}\tau}d\tau = \mathbf{A}^{-1}(e^{\mathbf{A}t} - \mathbf{I}) \ . \ \square \tag{523}$$

Remember that in a derivative of a matrix with respect to another matrix, the denominator matrix of the derivative determines the order scheme of the derivative, i.e., referring to Eq.(349), $\mathbf{M}$ determines the order. For a two-dimensional vector $\mathbf{f}$, the elements of which depend on $x_1(t),\ x_2(t),\ \ldots$ and for $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$,

$$\frac{d\mathbf{f}}{dt} = \frac{d\binom{f_1}{f_2}}{dt} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots \end{pmatrix}\begin{pmatrix} \frac{dx_1}{dt} \\ \frac{dx_2}{dt} \\ \vdots \end{pmatrix} = \frac{\partial \mathbf{f}}{\partial \mathbf{x}^T}\dot{\mathbf{x}} = \frac{\partial \mathbf{f}}{\partial \mathbf{x}^T}\mathbf{f} \ . \tag{524}$$

Alternatively,

$$\frac{d\mathbf{f}^T}{dt} = \Big(\frac{df_1}{dt} \vdots \frac{df_2}{dt}\Big) \tag{525}$$

$$= \Big(\frac{\partial f_1}{\partial x_1}\dot{x}_1 + \frac{\partial f_1}{\partial x_2}\dot{x}_2 + \ldots \vdots \frac{\partial f_2}{\partial x_1}\dot{x}_1 + \frac{\partial f_2}{\partial x_2}\dot{x}_2 + \ldots\Big) \tag{526}$$

$$= (\dot{x}_1 \quad \dot{x}_2 \quad \ldots)\begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_2}{\partial x_1} \\ \frac{\partial f_1}{\partial x_2} & \frac{\partial f_2}{\partial x_2} \\ \vdots & \vdots \end{pmatrix} = \mathbf{f}^T\frac{\partial \mathbf{f}^T}{\partial \mathbf{x}} \ . \tag{527}$$

From the above correspondences,

$$\frac{d\mathbf{f}^T\mathbf{f}}{dt} = \mathbf{f}^T\frac{\partial\mathbf{f}}{\partial\mathbf{x}^T}\mathbf{f} + \mathbf{f}^T\frac{\partial\mathbf{f}^T}{\partial\mathbf{x}}\mathbf{f} \tag{528}$$

$$= \mathbf{f}^T(\frac{\partial\mathbf{f}}{\partial\mathbf{x}^T} + \frac{\partial\mathbf{f}^T}{\partial\mathbf{x}})\mathbf{f}\ . \tag{529}$$

Hence,

$$\frac{\partial\mathbf{f}^T\mathbf{Q}\mathbf{f}}{\partial t} = \mathbf{f}^T(\mathbf{Q}\frac{\partial\mathbf{f}}{\partial\mathbf{x}^T} + \frac{\partial\mathbf{f}^T}{\partial\mathbf{x}}\mathbf{Q})\mathbf{f}\ . \tag{530}$$

For an exponentially weighted Lyapunov function $V = e^{at}\mathbf{f}^T\mathbf{Q}\mathbf{f}$ and a system $\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x})$, the stability condition with stability degree a is

$$V = e^{at}\mathbf{f}^T(\mathbf{x})\mathbf{Q}\mathbf{f}(\mathbf{x}) \tag{531}$$

$$\dot{V} \stackrel{(530)}{=} e^{at}\mathbf{f}^T(\mathbf{Q}\frac{\partial\mathbf{f}}{\partial\mathbf{x}^T} + \frac{\partial\mathbf{f}^T}{\partial\mathbf{x}}\mathbf{Q})\mathbf{f} + ae^{at}\mathbf{f}^T\mathbf{Q}\mathbf{f} \tag{532}$$

$$= e^{at}\mathbf{f}^T[\mathbf{Q}\frac{\partial\mathbf{f}}{\partial\mathbf{x}^T} + \frac{\partial\mathbf{f}^T}{\partial\mathbf{x}}\mathbf{Q} + a\mathbf{Q}]\mathbf{f} \stackrel{\triangle}{=} e^{at}\mathbf{f}^T\mathbf{P}\mathbf{f}\ . \tag{533}$$

Hence, to guarantee stability the matrix

$$\mathbf{P} = \mathbf{Q}\frac{\partial\mathbf{f}}{\partial\mathbf{x}^T} + \frac{\partial\mathbf{f}^T}{\partial\mathbf{x}}\mathbf{Q} + a\mathbf{Q} \tag{534}$$

must be negative definite.

15 Per-Unit Square Index of Performance

Find the maximum of a per-unit energy expression versus $\mathbf{x}$, i.e.,

$$\sup_{\mathbf{x}} \frac{\mathbf{x}^T\mathbf{R}\mathbf{x}}{\mathbf{x}^T\mathbf{x}}\ . \tag{535}$$

(The expression $\langle\mathbf{A}\mathbf{x},\mathbf{x}\rangle \stackrel{\triangle}{=} \mathbf{x}^H\mathbf{A}\mathbf{x}$ is also termed inner product.) For $\mathbf{R} = \mathbf{R}^H$, the result is given by the Rayleigh Principle. Comparison of Eq. 535 with the definition Eq.(536) of the singular value

$$\sigma_{\max}[\mathbf{M}] = \sup_{\mathbf{x}} \frac{||\mathbf{M}\mathbf{x}||_F}{||\mathbf{x}||_F} = \sup_{\mathbf{x}} \sqrt{\frac{\mathbf{x}^T\mathbf{M}^T\mathbf{M}\mathbf{x}}{\mathbf{x}^T\mathbf{x}}} \tag{536}$$

yields the correspondences $\mathbf{R} = \mathbf{M}^T\mathbf{M} \in \mathcal{R}^{n\times n}$ and

$$\sup_{\mathbf{x}} \frac{\mathbf{x}^T\mathbf{R}\mathbf{x}}{\mathbf{x}^T\mathbf{x}} = \sigma^2_{\max}[\mathbf{M}] = \sigma^2_{\max}[\sqrt{\mathbf{R}}]\ . \tag{537}$$

The worst initial condition $\mathbf{x}_0$ which yields the maximum per-unit index of performance I_{pu} results from

$$\begin{aligned}\frac{\partial}{\partial \mathbf{x}} \frac{\mathbf{x}^T \mathbf{M}^T \mathbf{M} \mathbf{x}}{\mathbf{x}^T \mathbf{x}} &= \mathbf{0} & (538)\\ \mathbf{M}^T \mathbf{M} \mathbf{x} &= \frac{\mathbf{x}^T \mathbf{M}^T \mathbf{M} \mathbf{x}}{\mathbf{x}^T \mathbf{x}} \mathbf{x} & (539)\\ \mathbf{M}^T \mathbf{M} \mathbf{x} &= \sigma_{\max}^2[\mathbf{M}] \cdot \mathbf{x}\,. & (540)\end{aligned}$$

The last but one line equals the correspondence of eigenvalues and eigenvectors $\mathbf{Aa} = \lambda(\mathbf{A}) \cdot \mathbf{a}$, in detail

$$(\mathbf{M}^T\mathbf{M}) \cdot \mathrm{eigv}[\mathbf{M}^T\mathbf{M}] = \mathrm{eig}_{\max}[\mathbf{M}^T\mathbf{M}] \cdot \mathrm{eigv}[\mathbf{M}^T\mathbf{M}] = \sigma_{\max}^2[\mathbf{M}] \cdot \mathrm{eigv}[\mathbf{M}^T\mathbf{M}]\,. \quad (541)$$

The largest eigenvalue is selected in order to find the largest I_{pu}. The largest eigenvalue of $\mathbf{M}^T\mathbf{M}$ is the squared singular value $\sigma_{\max}[\mathbf{M}]$. The worst vector $\mathbf{x}$ is $\mathbf{x}_0 = \mathrm{eigv}[\mathbf{R}]$ and has the manifold of the eigenvector. Only the direction is determined, not the length.

Referring to Eq.(537) and anticipating Eq.(611) (*Weinmann, A., 2003; 2004a*), an upper bound for the maximum per-unit energy of the actuating variable is

$$I_{um} = \sigma_{\max}^2[\sqrt{\mathbf{P}_{Ry}}] \le \|\sqrt{\mathbf{P}_{Ry}}\|_F^2 = \mathrm{tr}[(\sqrt{\mathbf{P}_{Ry}})^T(\sqrt{\mathbf{P}_{Ry}})] = \mathrm{tr}[\mathbf{P}_{Ry}]\,, \quad (542)$$

and the worst initial condition is $\mathbf{x}_0 = \mathrm{eigv}[\mathbf{P}_{Ry}]$.

16 Derivatives of the Matrix Pencil

With the definition of the inverse resolvent matrix (equivalent to the so-called matrix pencil, *Laub, A.J., 2005*)

$$\mathbf{H} \stackrel{\triangle}{=} s\mathbf{I} - \mathbf{A} - \mathbf{BK}\,, \quad (543)$$

the first and the higher derivatives with respect to s are presented (*Weinmann, A., 2004*). We presuppose $\mathbf{H}$ nonsingular.

16.1 First Derivative

$$\frac{\partial \det \mathbf{H}}{\partial s} \stackrel{(130)}{=} \mathrm{tr}[\frac{\partial \mathbf{H}}{\partial s} \mathbf{adj}\ \mathbf{H}] = \mathrm{tr}[\mathbf{I}\ \mathbf{adj}\ \mathbf{H}] = \mathrm{tr}[\mathbf{adj}\ \mathbf{H}] = \mathrm{tr}[\mathbf{H}^{-1} \det \mathbf{H}]\,. \quad (544)$$

With the definition [not to be confused with the unit vector $\mathbf{e}_i$ of Eq.(4)]

$$e_i \stackrel{\triangle}{=} \mathrm{tr}[\mathbf{H}^{-i}(s)] \quad (545)$$

and the definition of Q_i in Eq.(559), one finds

$$\frac{\partial \det \mathbf{H}}{\partial s} \stackrel{(543)}{=} \det \mathbf{H} \cdot \mathrm{tr}[\mathbf{H}^{-1}] \stackrel{(545)}{=} \det \mathbf{H} \cdot e_1 \ \stackrel{\triangle}{=} \det \mathbf{H} \cdot Q_1 \ . \tag{546}$$

For $\mathbf{H}$ singular, see *Weinmann, A., 2004, Section 6*. Furthermore, utilize the relations and abbreviations

$$\mathbf{L}(s) \ \stackrel{\triangle}{=} \ \mathbf{adj}\ \mathbf{H}(s) \tag{547}$$

$$\frac{\partial \mathrm{tr}[\mathbf{H}^{-1}]}{\partial s} = -\mathrm{tr}[\mathbf{H}^{-2}] \tag{548}$$

$$\frac{\partial \mathrm{tr}\mathbf{H}^{-i}}{\partial s} = -i \ \mathrm{tr}\mathbf{H}^{-i-1} \quad \text{or} \quad \frac{\partial e_i}{\partial s} = -i \ e_{i+1} \tag{549}$$

$$\frac{\partial^n}{\partial s^n}\mathrm{tr}\mathbf{H}^{-1} = (-1)^n \ n! \ \mathrm{tr}\mathbf{H}^{-n-1} \ . \tag{550}$$

In general,

$$e_i^k \stackrel{(545)}{=} (\mathrm{tr}\mathbf{H}^{-i})^k = \frac{(\mathrm{tr}\mathbf{L}^i)^k}{(\det \mathbf{H})^{ik}} \ . \tag{551}$$

16.2 Second Derivative

Continuing the derivations using differentiation of the product

$$\begin{aligned}
\frac{\partial^2 \det \mathbf{H}}{\partial s^2} &= \frac{\partial}{\partial s}\frac{\partial \det \mathbf{H}}{\partial s} = \frac{\partial}{\partial s} \ \det \mathbf{H} \cdot \mathrm{tr}\mathbf{H}^{-1} && (552)\\
&= \frac{\partial \det \mathbf{H}}{\partial s}\mathrm{tr}\mathbf{H}^{-1} + \det \mathbf{H}\frac{\partial \mathrm{tr}\mathbf{H}^{-1}}{\partial s} && (553)\\
&= \det \mathbf{H} \cdot (\mathrm{tr}\mathbf{H}^{-1})^2 + \det \mathbf{H} \cdot \mathrm{tr}[-\mathbf{H}^{-2}] && (554)\\
&\stackrel{(545)}{=} \det \mathbf{H} \cdot [e_1^2 - e_2] \ \stackrel{\triangle}{=} \ \det \mathbf{H} \cdot Q_2 && (555)\\
&= (\mathrm{tr}\ \mathbf{adj}\ \mathbf{H})(\mathrm{tr}\mathbf{H}^{-1} - \frac{\mathrm{tr}\mathbf{H}^{-2}}{\mathrm{tr}\mathbf{H}^{-1}}) && (556)\\
&= \mathrm{tr}\ \mathbf{adj}\ \mathbf{H} \cdot \mathrm{tr}\mathbf{H}^{-1} - \frac{1}{\det \mathbf{H}}\mathrm{tr}(\mathbf{adj}\ \mathbf{H})^2 && (557)\\
&= \frac{(\mathrm{tr}\ \mathbf{adj}\ \mathbf{H})^2 - \mathrm{tr}\ (\mathbf{adj}\ \mathbf{H})^2}{\det \mathbf{H}} \ . && (558)
\end{aligned}$$

16.3 i-th Derivative

For the i-th derivative one has

$$\frac{\partial^i \det \mathbf{H}}{\partial s^i} = Q_i \det \mathbf{H} \ . \tag{559}$$

The list of factors Q_i of $\det \mathbf{H}$ for its i-th derivative is

$$
\begin{aligned}
Q_0 &= 1 && (560)\\
Q_1 &\overset{(545)}{=} e_1 && (561)\\
Q_2 &= e_1^2 - e_2 && (562)\\
Q_3 &= e_1^3 - 3e_1e_2 + 2e_3 && (563)\\
Q_4 &= e_1^4 - 6e_1^2e_2 + 8e_1e_3 + 3e_2^2 - 6e_4 && (564)\\
Q_5 &= e_1^5 - 10e_1^3e_2 + 20e_1^2e_3 + 15e_1e_2^2 - 30e_1e_4 - 20e_2e_3 + 24e_5 && (565)\\
Q_6 &\overset{(545)}{=} e_1^6 - 15e_1^4e_2 + 40e_1^3e_3 + 45e_1^2e_2^2 - 90e_1^2e_4\\
&\quad -80e_1e_2e_3 + 144e_1e_5 + 120e_1e_2e_4 - 15e_2^3\\
&\quad +90e_2e_4 + 40e_3^2 - 120e_6\ . && (566)
\end{aligned}
$$

For numerical consideration concerning the inverse of $\mathbf{H}(s)$ and its powers see Section 6 of *Weinmann, A., 2004*.

16.4 Multiple Differential Quotients of the Matrix Pencil

Select the matrix pencil referring to Eq.(567) or Eq.(543). Then, $(\mathbf{H}_{n\,\mapsto\,(n-1),\,md})_i$ is defined as the operation, which selects the submatrix i along the main diagonal "md" with dimension reduction from $n \times n$ to $(n-1) \times (n-1)$, having striked out the i-th row and i-the column. According to the trace operator in Eq.(544), the determinants of all these matrices have to be summed up. $(\mathbf{H}_{n\,\mapsto\,z,\,md})_i$ is defined analogously irrespective how many intermediate steps are needed.

$$
\begin{aligned}
\mathbf{H}_n(s) &\triangleq s\mathbf{I}_n - \mathbf{A}^{[n\times n]} && (567)\\
\frac{\partial}{\partial s}\det \mathbf{H}_n(s) &\triangleq \frac{\partial}{\partial s}\det\left(s\mathbf{I}_n - \mathbf{A}^{[n\times n]}\right) \overset{(544)}{=} \mathrm{tr}[\left(\mathbf{adjH}^{[n\times n]}_{n\mapsto(n-1),md}\right)^{[(n-1)\times(n-1)]} && (568)\\
\frac{\partial}{\partial s}\det \mathbf{H}_n(s) &= \sum_{i=1}^{n} \det(\mathbf{H}_{n\mapsto(n-1),\,md})_i && (569)\\
\frac{\partial^2}{\partial s^2}\mathbf{H}_n(s) &= \sum_{i=1}^{n} \mathrm{tr}[(\mathbf{adjH}_{n\mapsto(n-1),\,md})_i^{[(n-2)\times(n-2)]}] && (570)\\
&= \sum_{j=1}^{n-1}\sum_{i=1}^{n} \det(\mathbf{H}_{n\,\mapsto\,(n-2),\,md})_{i,j}\ . && (571)
\end{aligned}
$$

Select, for example, $n = 4$ and the arbitrary matrix $\mathbf{H}_4$ the determinants of which should be differentiated with respect to s three times:

$$
\frac{\partial \det \mathbf{H}_4(s)}{\partial s} = \mathrm{tr}\ \mathbf{adjH}_4 = \sum_{i=1}^{4} \det(\mathbf{H}_{4\mapsto 3})_i \qquad (572)
$$

$$\frac{\partial^2 \det \mathbf{H}_4}{\partial s^2} = \sum_{i=1}^{4} \frac{\partial}{\partial s} \det(\mathbf{H}_{4\mapsto 3})_i \tag{573}$$

$$= \sum_{i=1}^{4} \operatorname{tr} \mathbf{adj}(\mathbf{H}_{4\mapsto 3})_i \tag{574}$$

$$= \sum_{j=1}^{3} \sum_{i=1}^{4} \det[(\mathbf{H}_{4\mapsto 2})_i]_j \tag{575}$$

$$\frac{\partial^3 \det \mathbf{H}_4}{\partial s^3} = \sum_{k=1}^{2} \sum_{j=1}^{3} \sum_{i=1}^{4} \det\{[(\mathbf{H}_{4\mapsto 1})_i]_j\}_k \ . \tag{576}$$

Detecting a multiple root (at -1) of det $\mathbf{H}_4$ without applying the derivatives of powers of s

$$\mathbf{H}_4 = \begin{pmatrix} s & -1 & 0 & 0 \\ 0 & s & -1 & 0 \\ 0 & 0 & s & -1 \\ 1 & 4 & 6 & s+4 \end{pmatrix}, \quad \det \mathbf{H}_4 = (s+1)^4 \tag{577}$$

$$(\mathbf{H}_{4\mapsto 3})_i : \begin{pmatrix} s & -1 & 0 \\ 0 & s & -1 \\ 4 & 6 & s+4 \end{pmatrix}, \begin{pmatrix} s & 0 & 0 \\ 0 & s & -1 \\ 1 & 6 & s+4 \end{pmatrix}, \begin{pmatrix} s & -1 & 0 \\ 0 & s & 0 \\ 1 & 4 & s+4 \end{pmatrix}, \begin{pmatrix} s & -1 & 0 \\ 0 & s & -1 \\ 0 & 0 & s \end{pmatrix} \tag{578}$$

Now there are four matrices, the sum of their determinants (minors) is the result.

$$\det \begin{pmatrix} s & -1 & 0 \\ 0 & s & -1 \\ 4 & 6 & s+4 \end{pmatrix} + \det \begin{pmatrix} s & 0 & 0 \\ 0 & s & -1 \\ 1 & 6 & s+4 \end{pmatrix} + \det \begin{pmatrix} s & -1 & 0 \\ 0 & s & 0 \\ 1 & 4 & s+4 \end{pmatrix} + \det \begin{pmatrix} s & -1 & 0 \\ 0 & s & -1 \\ 0 & 0 & s \end{pmatrix} . \tag{579}$$

$$\sum_i \det(\mathbf{H}_{4\mapsto 3})_i = (s^3+4s^2+6s+4)+(s^3+4s^2+6s)+(s^3+4s^2)+(s^3) = 4s^3+12s^2+12s+4 = 4(s+1)^3 . \tag{580}$$

For the next differentiation step, *each of the four* (3×3)-determinants is treated as the (4×4)-determinant of the setup. For each of the four (3×3)-matrices, calculating the determinants, we have 3 minors, yielding 12 determinants (cofactors, if $(-1)^{i+j}$ would be included)

$$((\mathbf{H}_{4\mapsto 2})_i)_j : \quad \begin{pmatrix} s & -1 \\ 6 & s+4 \end{pmatrix}, \begin{pmatrix} s & 0 \\ 4 & s+4 \end{pmatrix}, \begin{pmatrix} s & -1 \\ 0 & s \end{pmatrix} ; \tag{581}$$

$$\begin{pmatrix} s & -1 \\ 6 & s+4 \end{pmatrix}, \begin{pmatrix} s & 0 \\ 1 & s+4 \end{pmatrix}, \begin{pmatrix} s & 0 \\ 0 & s \end{pmatrix} ; \tag{582}$$

$$\begin{pmatrix} s & 0 \\ 4 & s+4 \end{pmatrix}, \begin{pmatrix} s & 0 \\ 1 & s+4 \end{pmatrix}, \begin{pmatrix} s & -1 \\ 0 & s \end{pmatrix} ; \tag{583}$$

$$\begin{pmatrix} s & -1 \\ 0 & s \end{pmatrix}, \quad \begin{pmatrix} s & 0 \\ 0 & s \end{pmatrix}, \quad \begin{pmatrix} s & -1 \\ 0 & s \end{pmatrix} \tag{584}$$

$$\sum_{j=1}^{3} \sum_{i=1}^{4} \det[(\mathbf{H}_{4\mapsto 2})_i]_j = 12s^2 + 24s + 12 = 12(s+1)^2 \tag{585}$$

$$\sum_{k=1}^{2} \sum_{j=1}^{3} \sum_{i=1}^{4} \det\{[(\mathbf{H}_{4\mapsto 1})_i]_j\}_k = 24s + 24 \ . \tag{586}$$

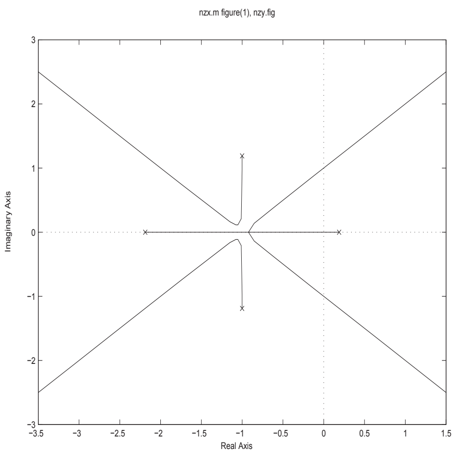

Figure 5: Root locus with break-away point close to -1 with `rlocus`([1], [1 4 6.001 4 -1])

17 Derivative of Square Performance Expressions

17.1 Energy of the State Variable

Consider the weighted energy of the transient $\mathbf{x}(t) = e^{\mathbf{A}_{cl}t}\mathbf{x}_0$ in a closed-loop control system with $\mathbf{A}_{cl} \triangleq \mathbf{A}+\mathbf{BK}$ where $\mathbf{A} \in \mathcal{R}^{n\times n}$, $\mathbf{B} \in \mathcal{R}^{n\times m}$, $\mathbf{K} \in \mathcal{R}^{m\times n}$, and the index of performance I_x based on the state variable $\mathbf{x}(t)$

$$I_x = \int_0^\infty \mathbf{x}(t)^T \mathbf{Q}\mathbf{x}(t)dt = \int_0^\infty \mathbf{x}_0^T e^{\mathbf{A}_{cl}^T t}\mathbf{Q}e^{\mathbf{A}_{cl}t}\mathbf{x}_0 dt \triangleq \mathbf{x}_0^T \mathbf{P}_Q \mathbf{x}_0 \ , \tag{587}$$

where $\mathbf{A}_{cl}^T\mathbf{P}_Q + \mathbf{P}_Q\mathbf{A}_{cl} = -\mathbf{Q} \in \mathcal{R}^{n\times n}$ and $\mathbf{P}_Q$ is symmetric (e.g., *Weinmann, A., 1995*).

The derivation is based on

$$\int_0^t e^{a\tau}d\tau \;=\; \frac{e^{a\tau}}{a}\Big|_0^t = \frac{e^{at}-1}{a} \quad (\text{and for } a<0,\ t\to\infty) \;=\; -1/a \tag{588}$$

$$\int_0^t e^{\mathbf{A}\tau}d\tau \;=\; \mathbf{A}^{-1}e^{\mathbf{A}\tau}\Big|_0^t = \mathbf{A}^{-1}(e^{\mathbf{A}t}-\mathbf{I}) \quad (\text{and for stable } \mathbf{A},\ t\to\infty) \;=\; -\mathbf{A}^{-1} \tag{589}$$

and partial integration

$$d(\mathbf{UV}) = \mathbf{U}d\mathbf{V} + (d\mathbf{U})\ \mathbf{V} \quad \rightsquigarrow \quad \int_0^\infty \mathbf{U}d\mathbf{V} = \mathbf{UV}\Big|_0^\infty - \int_0^\infty (d\mathbf{U}\cdot\mathbf{V}) \ , \tag{590}$$

where (presupposing stable $\mathbf{A}_{cl}$)

$$d\mathbf{V} \triangleq \mathbf{Q}e^{\mathbf{A}_{cl}t}dt \ , \quad \mathbf{V} = \mathbf{Q}\mathbf{A}_{cl}^{-1}e^{\mathbf{A}_{cl}t} \ , \quad \mathbf{U} \triangleq e^{\mathbf{A}_{cl}^T t} \ , \quad \mathbf{UV}\Big|_0^\infty = -\mathbf{Q}\mathbf{A}_{cl}^{-1} \ . \tag{591}$$

The integral

$$\mathbf{P}_Q = \int_0^\infty e^{\mathbf{A}_{cl}^T t}\mathbf{Q}e^{\mathbf{A}_{cl}t}dt \in \mathcal{R}^{n\times n} \tag{592}$$

cannot be utilized for $\frac{\partial \mathbf{P}_Q}{\partial \mathbf{K}}$ or $\frac{\partial I_x}{\partial \mathbf{K}}$ because $e^{\mathbf{A}+\mathbf{BK}} \neq e^{\mathbf{A}}e^{\mathbf{BK}}$, even for $\mathbf{K}+\Delta\mathbf{K}$ and $\Delta\mathbf{K}$ small in norm sense, see Eq.(18). However, resolving

$$\mathbf{A}_{cl}^T\mathbf{P}_Q + \mathbf{P}_Q\mathbf{A}_{cl} + \mathbf{Q} = \mathbf{0} \tag{593}$$

with respect to $\mathbf{P}_Q$, its derivation with respect to $\mathbf{K}$ can be obtained employing the first-order Taylor expansion

$$(\mathbf{A}_{cl}^T + \Delta\mathbf{K}^T\mathbf{B}^T)(\mathbf{P}_Q + \Delta\mathbf{P}_Q) + (\mathbf{P}_Q + \Delta\mathbf{P}_Q)(\mathbf{A}_{cl} + \mathbf{B}\Delta\mathbf{K}) + \mathbf{Q} = \mathbf{0}\ . \tag{594}$$

$$\mathbf{E}_a \ \stackrel{\triangle}{=}\ (\mathbf{A}^T \oplus \mathbf{A}^T)^{-1} = (\mathbf{I}_n \otimes \mathbf{A}^T + \mathbf{A}^T \otimes \mathbf{I}_n)^{-1} \in \mathcal{R}^{n^2\times n^2}\ . \tag{595}$$

With the additional abbreviation

$$\mathbf{Z}_0 \ \stackrel{\triangle}{=}\ (\mathbf{P}_Q\mathbf{B}) \otimes \mathbf{I}_n + [\mathbf{I}_n \otimes (\mathbf{P}_Q\mathbf{B})]\mathbf{U}_{nm} \ \in \mathcal{R}^{n^2\times mn} \tag{596}$$

one finds the derivative

$$\frac{\partial I_x}{\partial \mathbf{K}} \ = \ -\{\mathbf{I}_m \otimes [(\mathbf{x}_0^T \otimes \mathbf{x}_0^T)\ \mathbf{E}_a\mathbf{Z}_0]\}[(\text{col}\ \mathbf{I}_m) \otimes \mathbf{I}_n]\ , \tag{597}$$

$$\stackrel{(70)}{=} \ \{\text{loc}_n[\mathbf{Z}_0^T\mathbf{E}_a^T(\mathbf{x}_0 \otimes \mathbf{x}_0)]\}^T\ . \tag{598}$$

17.2 Time-Weighted Energy of the State Variable

Now, based on Lyapunov equations, the sensitivity of time-weighted and time-to-powers-weighted energy associated with the state variable is derived for an arbitrary power i

$$\mathbf{P}_{Qi} \stackrel{\triangle}{=} \int_0^\infty e^{\mathbf{A}^T t}\mathbf{Q}e^{\mathbf{A}t}\ \ t^i\ \ dt \qquad I_{xi} = \mathbf{x}_o^T\mathbf{P}_{Qi}\mathbf{x}_o \tag{599}$$

$$\mathbf{A}_{cl}^T\mathbf{P}_{Qi} + \mathbf{P}_{Qi}\mathbf{A}_{cl} + i\ \mathbf{P}_{Q,i-1} = \mathbf{0}\ , \tag{600}$$

where $\mathbf{P}_{Qo} = \mathbf{P}_Q$. Replacing $\mathbf{Z}_0$ by $\mathbf{Z}_i$ and $\mathbf{P}_Q$ by $\mathbf{P}_{Qi}$ in Eq.(596), one has new matrices $\mathbf{S}_i \in \mathcal{R}^{n^2\times mn}$

$$\mathbf{S}_0 \ = \ -\mathbf{E}_a\mathbf{Z}_0 \tag{601}$$

$$\mathbf{S}_1 \ = \ -\mathbf{E}_a\mathbf{Z}_1 + \mathbf{E}_a^2\mathbf{Z}_0 \tag{602}$$

$$\mathbf{S}_2 \ = \ -\mathbf{E}_a\mathbf{Z}_2 + 2\mathbf{E}_a^2\mathbf{Z}_1 - 2\mathbf{E}_a^3\mathbf{Z}_0 \tag{603}$$

$$\mathbf{S}_3 \ = \ -\mathbf{E}_a\mathbf{Z}_3 + 3\mathbf{E}_a^2\mathbf{Z}_2 - 6\mathbf{E}_a^3\mathbf{Z}_1 + 6\mathbf{E}_a^4\mathbf{Z}_0 \tag{604}$$

$$\mathbf{S}_4 \ = \ -\mathbf{E}_a\mathbf{Z}_4 + 4\mathbf{E}_a^2\mathbf{Z}_3 - 12\mathbf{E}_a^3\mathbf{Z}_2 + 24\mathbf{E}_a^4\mathbf{Z}_1 - 24\mathbf{E}_a^5\mathbf{Z}_0 \tag{605}$$

$$\vdots$$

$$\mathbf{S}_i \ = \ -\mathbf{E}_a\mathbf{Z}_i - i\ \mathbf{E}_a\mathbf{S}_{i-1}\ . \tag{606}$$

Using Eq.(606), the increment ΔI_{xi} weighted with the i-th power of t is (*Weinmann, A., 2003*)

$$\Delta I_{xi} = (\mathbf{x}_0^T \otimes \mathbf{x}_0^T)\mathrm{col}\Delta\mathbf{P}_{Qi} \ . \tag{607}$$

The matricial gradient results as

$$\begin{aligned} \frac{\partial I_{xi}}{\partial \mathbf{K}} &= \Big\{\mathbf{I}_m \otimes [(\mathbf{x}_0^T \otimes \mathbf{x}_0^T)\mathbf{S}_i]\Big\}[(\mathrm{col}\ \mathbf{I}_m) \otimes \mathbf{I}_n] & (608)\\ &\overset{(70)}{=} \{\mathrm{loc}_n[\mathbf{S}_i^T(\mathbf{x}_0 \otimes \mathbf{x}_0)]\}^T \ . & (609) \end{aligned}$$

17.3 Actuating Energy

From *Weinmann, A., 2003*, with $\mathbf{K}_y \in \mathcal{R}^{m\times r}$, $\mathbf{C} \in \mathcal{R}^{r\times n}$

$$\begin{aligned} I_u &\triangleq \int_0^\infty \mathbf{u}^T(t)\mathbf{R}\mathbf{u}(t)dt = \int_0^\infty \mathbf{x}^T(t)\mathbf{C}^T\mathbf{K}_y^T\mathbf{R}\mathbf{K}_y\mathbf{C}\mathbf{x}(t)dt & (610)\\ &= \mathbf{x}_0^T \int_0^\infty e^{\mathbf{A}_{cl}^T t}\mathbf{C}^T\mathbf{K}_y^T\mathbf{R}\mathbf{K}_y\mathbf{C}e^{\mathbf{A}_{cl}t}dt\ \mathbf{x}_0 \triangleq \mathbf{x}_0^T\mathbf{P}_{Ry}\mathbf{x}_0 \ , & (611) \end{aligned}$$

where

$$\mathbf{A}_{cl}^T\mathbf{P}_{Ry} + \mathbf{P}_{Ry}\mathbf{A}_{cl} + \mathbf{C}^T\mathbf{K}_y^T\mathbf{R}\mathbf{K}_y\mathbf{C} = \mathbf{0} \ . \tag{612}$$

Defining

$$\begin{aligned} \mathbf{E}_1 &\triangleq (\mathbf{I}_n \otimes \mathbf{A}_{cl}^T + \mathbf{A}_{cl}^T \otimes \mathbf{I}_n)^{-1} \in \mathcal{R}^{n^2\times n^2} & (613)\\ \mathbf{E}_2 &\triangleq \mathbf{P}_{Ry}\mathbf{B} + \mathbf{C}^T\mathbf{K}_y^T\mathbf{R} \in \mathcal{R}^{n\times m} & (614) \end{aligned}$$

$$\frac{\partial I_u}{\partial \mathbf{K}_y} = -\mathbf{E}_2^T\{\mathrm{loc}_n[\mathrm{col}\mathbf{X} + \mathbf{U}_{nn}\mathrm{col}\mathbf{X}]\}^T\mathbf{C}^T = -2\mathbf{E}_2^T\mathbf{X}\mathbf{C}^T \ , \tag{615}$$

where

$$\mathbf{X} \triangleq \mathtt{lyap(Acl, Acl', -x0 * x0')} \tag{616}$$

$$\mathbf{A}_{cl}\mathbf{X} + \mathbf{X}\mathbf{A}_{cl}^T = \mathbf{x}_0\mathbf{x}_0^T \tag{617}$$

$$(\mathbf{I}_n \otimes \mathbf{A}_{cl} + \mathbf{A}_{cl} \otimes \mathbf{I}_n)\mathrm{col}\,\mathbf{X} = \mathrm{col}(\mathbf{x}_0\mathbf{x}_0^T) \tag{618}$$

$$\mathrm{col}\,\mathbf{X} = (\mathbf{A}_{cl} \oplus \mathbf{A}_{cl})^{-1}(\mathbf{x}_0 \otimes \mathbf{x}_0) = \mathbf{E}_1^T(\mathbf{x}_0 \otimes \mathbf{x}_0) \tag{619}$$

$$\mathrm{loc}_r(\mathbf{v}^{[rm\times 1]}) \triangleq \mathbf{V}^{[r\times m]} \tag{620}$$

and r indicates the number of rows produced by loc_r.

Remark 1: There is an equivalent result although it looks quite different, hence the corresponding derivation is carried out: First, suppose

$$\mathbf{y} \triangleq [\mathbf{I}_{n^2} + \mathbf{U}_{nn}]\mathbf{E}_1^T(\mathbf{x}_0 \otimes \mathbf{x}_0) \quad \in \mathcal{R}^{n^2\times 1} \tag{621}$$

$$\mathbf{Y} \stackrel{\triangle}{=} \mathrm{loc}_n[\mathbf{y}] \ , \qquad \mathbf{y} = \mathrm{col}\mathbf{Y} \in \mathcal{R}^{n^2\times 1} \tag{622}$$

$$\{\mathrm{loc}_r[\mathrm{col}\,\mathbf{M}^{[r\times m]}\}^T = \mathbf{M}^T \ . \tag{623}$$

Second, from Eq.(612) $\Delta\mathbf{K}_y \mapsto \Delta\mathbf{P}_{Ry}$, from Eq.(611) $\Delta\mathbf{P}_{Ry} \mapsto \Delta I_u$, together $\Delta\mathbf{K}_y \mapsto \Delta I_u$. Third, using Eq.(351) etc. $\rightsquigarrow$ $\frac{\partial I_u}{\partial \mathbf{K}_y}$

$$\begin{aligned}
\frac{\partial I_u}{\partial \mathbf{K}_y} &= -\{\mathbf{I}_m \otimes [(\mathbf{x}_0 \otimes \mathbf{x}_0)^T\ \mathbf{E}_1(\mathbf{I}_{n^2} + \mathbf{U}_{nn})(\mathbf{E}_2 \otimes \mathbf{C}^T)]\}[(\mathrm{col}\mathbf{I}_m) \otimes \mathbf{I}_r] && (624)\\
&\stackrel{(70)}{\equiv} -\Big(\mathrm{loc}_r[(\mathbf{E}_2^T \otimes \mathbf{C})(\mathbf{I}_{n^2} + \mathbf{U}_{nn})\mathbf{E}_1^T(\mathbf{x}_0 \otimes \mathbf{x}_0)]\Big)^T && (625)\\
&= -\{\mathrm{loc}_r[(\mathbf{E}_2^T \otimes \mathbf{C})\mathbf{y}]\}^T = -\{\mathrm{loc}_r[(\mathbf{E}_2^T \otimes \mathbf{C})\mathrm{col}\mathbf{Y}]\}^T && (626)\\
&= -\{\mathrm{loc}_r[\mathrm{col}(\mathbf{C}\mathbf{Y}\mathbf{E}_2)]\}^T \stackrel{(623)}{=} -(\mathbf{C}\mathbf{Y}\mathbf{E}_2)^T \quad [\equiv (630),\ (615)] && (627)\\
&\equiv -\mathbf{E}_2^T\{\mathrm{loc}_n[(\mathbf{I}_{n^2} + \mathbf{U}_{nn})\mathbf{E}_1^T(\mathbf{x}_0 \otimes \mathbf{x}_0)]\}^T\mathbf{C}^T && (628)\\
&\stackrel{(621)}{=} -\mathbf{E}_2^T\{\mathrm{loc}_n[\mathbf{y}]\}^T\mathbf{C}^T = -\mathbf{E}_2^T\mathbf{Y}^T\mathbf{C}^T && (629)\\
&\stackrel{(619)}{=} -\mathbf{E}_2^T\{\mathrm{loc}_n[\mathrm{col}\,\mathbf{X} + \mathbf{U}_{nn}\mathrm{col}\,\mathbf{X}]\}^T\mathbf{C}^T = -2\mathbf{E}_2^T\mathbf{X}\mathbf{C}^T \ . && (630)
\end{aligned}$$

(Note that $\mathbf{Y} = 2\mathbf{X}$. The matrix $\mathbf{X}$ is symmetric, hence $\mathbf{U}_{nn}\mathrm{col}\mathbf{X} = \mathrm{col}\mathbf{X}$. But note: For $\mathbf{B} \in \mathcal{R}^{n\times m}$ one has $\mathrm{loc}_m(\mathrm{col}\mathbf{B}) \neq \mathbf{B}^T$.)

Remark 2: The Lyapunov equation

$$\begin{aligned}
\mathbf{A}^{(n\times n)}\mathbf{X}^{(n\times m)} + \mathbf{X}\mathbf{B}^{(m\times m)} &= \mathbf{C}^{(n\times m)} && (631)\\
\underbrace{(\mathbf{I}_m \otimes \mathbf{A} + \mathbf{B}^T \otimes \mathbf{I}_n)}_{\mathbf{R}}\mathrm{col}\mathbf{X} &= \mathrm{col}\ \mathbf{C} && (632)
\end{aligned}$$

is only solvable if $\mathbf{R}$ is nonsingular, i.e., iff $\mathbf{R}$ has no zero eigenvalues. The eigenvalues are $\lambda_i + \mu_j$ where $\lambda_i \in \Lambda(\mathbf{A})$ and $\mu_j \in \Lambda(\mathbf{B})$, $i \in \{1, 2, \ldots, n\}$, $j \in \{1, 2, \ldots, m\}$, i.e., $\mathbf{A}$ and $-\mathbf{B}$ must not have eigenvalues in common. If $\mathbf{A}$ and $\mathbf{B}$ are stable the solution is also given by

$$\mathbf{X} = -\int_0^\infty e^{\mathbf{A}t}\mathbf{C}e^{\mathbf{B}t}dt \ . \tag{633}$$

For eliminating double solutions in Lyapunov equations see *Weinmann, A., 2007d*.

Closed loop versus output controller matrix

$$\frac{\partial(\mathbf{A} + \mathbf{B}\mathbf{K}_y\mathbf{C})}{\partial \mathbf{K}_y^{(m\times r)}} = (\mathbf{I}_m \otimes \mathbf{B}^{(n\times m)})\bar{\mathbf{U}}_{mr}(\mathbf{I}_r \otimes \mathbf{C}^{(r\times n)}) \quad \in \mathcal{R}^{mn\times nr} \ . \tag{634}$$

17.4 Lyapunov Determinant with Respect to the Controller Matrix K

For the Lyapunov equation

$$\mathbf{A}_{cl}^T\mathbf{P}_n + \mathbf{P}_n\mathbf{A}_{cl} = -\mathbf{Q} \ , \quad \mathbf{A}_{cl} \stackrel{\triangle}{=} \mathbf{A} + \mathbf{B}\mathbf{K} \ , \tag{635}$$

one finds

$$\frac{\partial \det \mathbf{P}_n}{\partial \mathbf{K}} = [\mathbf{I}_m \otimes (\mathbf{h}_n^T \mathbf{U}_{n,m})][(\mathrm{col}\ \mathbf{I}_m) \otimes \mathbf{I}_n] \ , \tag{636}$$

where

$$\mathbf{h}_n^T \triangleq (\det \mathbf{P}_n)[\mathrm{col}\,(\mathbf{P}_n^{-T})]^T \mathbf{H}_n \tag{637}$$

and

$$\mathbf{H}_n \ = -[\mathbf{I}_n \otimes \mathbf{A}_{cl}^T + \mathbf{A}_{cl}^T \otimes \mathbf{I}_n]^{-1}\{[(\mathbf{P}_n\mathbf{B}) \otimes \mathbf{I}_n]\mathbf{U}_{m,n} + [\mathbf{I}_n \otimes (\mathbf{P}_n\mathbf{B})]\} \tag{638}$$

(*Weinmann, A., 2008d*).

17.5 Lyapunov-Based Efficiency Gradients

The efficiency measure r should be as low as possible

$$r \triangleq -\frac{\mathbf{x}^T\mathbf{Q}\mathbf{x}}{\mathbf{x}^T\mathbf{P}\mathbf{x}} \quad \to \quad \max_i\{\frac{-1}{\lambda_{Qi}[\mathbf{P}]}\} \ , \tag{639}$$

where $\mathbf{P}$ and $\mathbf{Q}$ are positive definite matrices in $\mathbf{A}^T\mathbf{P} + \mathbf{P}\mathbf{A} + \mathbf{Q} = \mathbf{0}$, and $\lambda_{Qi}[\mathbf{P}]$ is a generalized eigenvalue (see Section 4). One of the eigenvectors is the optimal solution $\mathbf{x}$.

Now, searching

$$\min_{\mathbf{Q}} \max_i \frac{-1}{\lambda_{Qi}[\mathbf{P}]} \ , \quad \text{the gradient results from} \tag{640}$$

$$\frac{\partial r}{\partial \mathbf{q}} = \frac{1}{(\mathbf{q}^T\mathbf{H}^{-1}\mathbf{y})^2}[\mathbf{H}^{-1}\mathbf{y}\mathbf{q}^T - \mathbf{y}\mathbf{q}^T\mathbf{H}^{-1}]\mathbf{y} \ , \quad \text{where} \tag{641}$$

$$\mathbf{q} \triangleq \mathrm{col}\mathbf{Q}, \quad \mathbf{y} \triangleq \mathbf{v}_{i_M} \otimes \mathbf{v}_{i_M} \text{ and } \ [\mathbf{V}, \mathbf{D}] = \mathrm{eig}[\mathbf{P}, \mathbf{Q}], \quad i_M \triangleq \arg_i \max D(i) \tag{642}$$

$$\mathbf{H} \triangleq \mathbf{I}_n \otimes \mathbf{A} + \mathbf{A} \otimes \mathbf{I}_n \ . \tag{643}$$

For discrete-time systems,

$$\mathbf{H} \triangleq (\mathbf{I}_n - \mathbf{\Phi} \otimes \mathbf{\Phi}) \quad \text{and} \quad \mathbf{\Phi}^T\mathbf{P}\mathbf{\Phi} - \mathbf{P} + \mathbf{Q} = \mathbf{0} \ . \tag{644}$$

18 Control with Prespecified Index of Performance

For the output feedback controlled system, $\mathbf{u}(t) = \mathbf{K}\mathbf{y}(t) = \mathbf{K}\mathbf{C}\mathbf{x}(t)$, abbreviating $\mathbf{F} \triangleq (\mathbf{A} + \mathbf{B}\mathbf{K}\mathbf{C})$, and for

$$I = \int_0^\infty (\mathbf{x}^T\mathbf{Q}\mathbf{x} + \mathbf{u}^T\mathbf{R}\mathbf{u})d\tau = \int_0^\infty \mathbf{x}_o^T e^{\mathbf{F}^T\tau}(\mathbf{Q} + \mathbf{C}^T\mathbf{K}^T\mathbf{R}\mathbf{K}\mathbf{C})e^{\mathbf{F}\tau}\ \mathbf{x}_o d\tau \ , \tag{645}$$

the performance limit is given $I \triangleq \mathbf{x}_o^T \mathbf{M} \mathbf{x}_o$. Then, we have the limit condition

$$\mathbf{F}^T\mathbf{M} + \mathbf{M}\mathbf{F} = -(\mathbf{Q} + \mathbf{C}^T\mathbf{K}^T\mathbf{R}\mathbf{K}\mathbf{C}) \ , \tag{646}$$

which is included via a Langrange multiplier $\mathbf{\Lambda}$. Carrying out the minimization, $\mathbf{\Lambda}$ results from

$$\mathbf{x}_o\mathbf{x}_o^T + \mathbf{F}\mathbf{\Lambda} + \mathbf{\Lambda}\mathbf{F}^T = \mathbf{0} \tag{647}$$

(*Yahagi, T., 1973; 1977*). The derivatives of I with respect to $\mathbf{A}, \mathbf{B}, \mathbf{C}$ and $\mathbf{K}$ are

$$\begin{aligned}
\frac{\partial I}{\partial \mathbf{A}} &= 2\mathbf{M}\mathbf{\Lambda} && (648)\\
\frac{\partial I}{\partial \mathbf{B}} &= 2\mathbf{M}\mathbf{\Lambda}\mathbf{C}^T\mathbf{K}^T && (649)\\
\frac{\partial I}{\partial \mathbf{C}} &= 2\mathbf{K}^T(\mathbf{R}\mathbf{K}\mathbf{C} + \mathbf{B}^T\mathbf{M})\mathbf{\Lambda} && (650)\\
\frac{\partial I}{\partial \mathbf{K}} &= 2(\mathbf{R}\mathbf{K}\mathbf{C} + \mathbf{B}^T\mathbf{M})\mathbf{\Lambda}\mathbf{C}^T \ . && (651)
\end{aligned}$$

$$\frac{\partial I}{\partial \mathbf{K}} = \mathbf{0} \quad \hookrightarrow \quad \mathbf{K} = -\mathbf{R}^{-1}\mathbf{B}^T\mathbf{M}\mathbf{\Lambda}\mathbf{C}^T(\mathbf{C}\mathbf{\Lambda}\mathbf{C}^T)^{-1} \ . \tag{652}$$

19 Derivatives for Flat Systems

The flat output $z \triangleq y_f$ is defined such that — under some specific $\mathbf{c}_f$ — the result $z = \mathbf{c}_f^T\mathbf{x}$ has the relative degree n; thus, one has $\mathbf{c}_f\dot{\mathbf{x}} \triangleq 0$, $\mathbf{c}_f\ddot{\mathbf{x}} \triangleq 0$ etc. up to $\mathbf{c}_f\mathbf{x}^{(n-1)} \triangleq 0$. Finally, $\mathbf{c}_f\mathbf{x}^{(n)} \triangleq 1$ (*Rothfuß, R., et al., 1997; Sira-Ramirez, H., and Agrawal, S.K., 2004; Zeitz, M., 2009*). Concatenating these conditions,

$$\begin{aligned}
\mathbf{c}_f^T(\mathbf{b} \vdots \mathbf{A}\mathbf{b} \vdots \mathbf{A}^2\mathbf{b} \vdots \ \dots \ \vdots \mathbf{A}^{n-1}\mathbf{b}) &= (0 \ \ 0 \ \ 0 \ \ 1) && (653)\\
\mathbf{c}_f^T\mathbf{L}_c &= \mathbf{e}_n^T && (654)\\
\mathbf{c}_f &= (\mathbf{e}_n^T\mathbf{L}_c^{-1})^T = \mathbf{L}_c^{-1,T}\mathbf{e}_n \ . && (655)
\end{aligned}$$

For subsequent use, the controllability matrix (*Franklin, G.F., et al., 2002*) and corresponding products

$$\mathbf{L}_c = [\mathbf{b} \vdots \mathbf{A}\mathbf{b} \vdots \ \dots \ \vdots \mathbf{A}^{i-1}\mathbf{b}] = \sum_{i=1}^{i=n} \mathbf{e}_i^T \otimes (\mathbf{A}^{i-1}\mathbf{b}) \quad \in \mathcal{R}^{n\times n} \tag{656}$$

$$\begin{aligned}
\mathbf{L}_c^T\mathbf{L}_c &= \sum_{i=1}^{i=n}\sum_{k=1}^{k=n}(\mathbf{e}_i\mathbf{e}_k^T) \otimes (\mathbf{b}^T\mathbf{A}^{T,i-1}\mathbf{A}^{k-1}\mathbf{b}) && (657)\\
\mathbf{L}_c\mathbf{L}_c^T &= \sum_{i=1}^{i=n}\mathbf{A}^{i-1}\mathbf{b}\mathbf{b}^T\mathbf{A}^{T,i-1} && (658)
\end{aligned}$$

are considered, especially for $\mathbf{A} = \mathbf{A}_o + \mathbf{b}\mathbf{k}_r^T$.

Derivative of the controllability matrix with respect to an element of the coefficient matrix $A_{\nu\mu}$:

$$\frac{\partial \mathbf{L}_c}{\partial A_{\nu\mu}} = \sum_{k=1}^{n} \frac{\partial}{\partial A_{\nu\mu}}[\mathbf{e}_k^T \otimes (\mathbf{A}^{k-1}\mathbf{b})] \tag{659}$$

$$\stackrel{(379)}{=} \sum_{k=1}^{n}(\mathbf{I}_1 \otimes \mathbf{U}_{1,n})(\frac{\partial \mathbf{A}^{k-1}\mathbf{b}}{\partial A_{\nu\mu}} \otimes \mathbf{e}_k^T) \cdot 1 \tag{660}$$

$$\stackrel{(351)}{=} \sum_{k=1}^{n} 1 \cdot [\frac{\partial \mathbf{A}^{k-1}}{\partial A_{\nu\mu}}(\mathbf{I}_1 \otimes \mathbf{b})] \otimes \mathbf{e}_k^T) \tag{661}$$

$$\stackrel{(665)}{=} \sum_{k=1}^{n}\sum_{i=1}^{k-1}(\mathbf{A}^{i-1}\frac{\partial \mathbf{A}}{\partial A_{\nu\mu}}\mathbf{A}^{k-i-1}\mathbf{b}) \otimes \mathbf{e}_k^T \tag{662}$$

$$\frac{\partial \mathbf{L}_c}{\partial A_{\nu\mu}} = \sum_{k=2}^{n}\sum_{i=1}^{k-1}(\mathbf{A}^{i-1}\mathbf{E}_{\nu,\mu}\mathbf{A}^{k-i-1}\mathbf{b}) \otimes \mathbf{e}_k^T \tag{663}$$

$$\frac{\partial \mathbf{L}_c}{\partial \mathbf{A}} = \sum_{\nu=1}^{n}\sum_{\mu=1}^{n}\mathbf{E}_{\nu,\mu} \otimes \left[\sum_{k=2}^{n}\sum_{i=1}^{k-1}\mathbf{A}^{i-1}\mathbf{E}_{\nu,\mu}\mathbf{A}^{k-i-1}\mathbf{b}) \otimes \mathbf{e}_k^T\right] , \tag{664}$$

$$\frac{\partial \mathbf{A}^{k-1}}{\partial \mathbf{k}_r} = \sum_{i=1}^{k-1}(\mathbf{I}_r \otimes \mathbf{A}^{i-1})\frac{\partial \mathbf{A}}{\partial \mathbf{k}_r}\mathbf{A}^{k-i-1} , \quad k \geq 2 . \tag{665}$$

$$\frac{\partial \mathbf{L}_c^T}{\partial A_{\nu\mu}} = \frac{\partial}{\partial A_{\nu\mu}}\sum_{i=1}^{n}\mathbf{e}_i \otimes (\mathbf{A}^{i-1}\mathbf{b})^T \stackrel{(381)}{=} \sum_{i=1}^{n}\frac{\partial \mathbf{b}^T\mathbf{A}^{T,i-1}}{\partial A_{\nu\mu}} \otimes \mathbf{e}_i , \tag{666}$$

$$\frac{\partial \mathbf{L}_c^T}{\partial A_{\nu\mu}} \stackrel{(371)}{=} \sum_{i=2}^{n}(\sum_{v=1}^{v=i-1}\mathbf{b}^T\mathbf{A}^{T,v-1}\mathbf{E}_{\nu,\mu}\mathbf{A}^{T,i-v-1}) \otimes \mathbf{e}_i . \tag{667}$$

$$\frac{\partial \mathrm{tr}[\mathbf{L}_c\mathbf{L}_c^T]}{\partial \mathbf{k}_r} = \mathrm{trh}[\frac{\partial(\mathbf{L}_c\mathbf{L}_c^T)}{\partial \mathbf{k}_r}] , \quad \text{see Eq.669} . \tag{668}$$

Select a vector of subtraces from a rectangular matrix $\mathbf{Z} \in \mathcal{R}^{n^2 \times n}$:

$$\mathrm{trh}[\mathbf{Z}] \stackrel{\triangle}{=} \sum_{i=1}^{n}\{\mathbf{e}_i \cdot \mathrm{tr}[(\mathbf{e}_i^T \otimes \mathbf{I}_n)\mathbf{Z}]\} . \tag{669}$$

19.1 Modal Flatness Matrix

Defining a flatness vector $\mathbf{z}$

$$\mathbf{z} \stackrel{\triangle}{=} \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{pmatrix} = \begin{pmatrix} z \\ \dot{z} \\ \vdots \\ z^{(n-1)} \end{pmatrix} = \begin{pmatrix} \mathbf{c}_f^T\mathbf{x} \\ \mathbf{c}_f^T\mathbf{A}\mathbf{x} \\ \vdots \\ \mathbf{c}_f^T\mathbf{A}^{n-1}\mathbf{x} \end{pmatrix} = \begin{pmatrix} \mathbf{c}_f^T \\ \mathbf{c}_f^T\mathbf{A} \\ \vdots \\ \mathbf{c}_f^T\mathbf{A}^{n-1} \end{pmatrix}\mathbf{x} \stackrel{\triangle}{=} \mathbf{T}_C\mathbf{x} \tag{670}$$

$$\dot{\mathbf{z}} = \begin{pmatrix} \mathbf{0}_{n-1,1} & \mathbf{I}_{n-1} \\ & \mathbf{a}_L^T \end{pmatrix} \mathbf{z} + \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} u \; . \tag{671}$$

The last row in the previous expression is

$$z^{(n)} = \mathbf{a}_L^T \mathbf{z} + u = \mathbf{c}_f^T \mathbf{A}^n \mathbf{x} + u = \mathbf{c}_f^T \mathbf{A}^n \mathbf{T}_C^{-1} \mathbf{z} + u \; . \tag{672}$$

The modal matrix for flatness coordinates and its derivative are

$$\mathbf{T}_C \overset{(670)}{=} \sum_{i=1}^{n} \mathbf{e}_i \otimes (\mathbf{c}_f \mathbf{A}^{i-1}) \overset{(655)}{=} \sum_{i=1}^{n} \mathbf{e}_i \otimes \{\mathbf{e}_n^T \mathbf{L}_c^{-1} \mathbf{A}^{i-1}\} \tag{673}$$

$$\overset{(656)}{=} \sum_{i=1}^{n} \mathbf{e}_i \otimes \{\mathbf{e}_n^T [\sum_{k=1}^{k=n} \mathbf{e}_k^T \otimes (\mathbf{A}^{k-1}\mathbf{b})]^{-1} \mathbf{A}^{k-1}\} \; , \tag{674}$$

$$\frac{\partial \mathbf{T}_C}{\partial A_{\nu\mu}} \overset{(381)}{=} \sum_{i=1}^{n} (\frac{\partial}{\partial A_{\nu\mu}} \{\mathbf{e}_n^T \mathbf{L}_c^{-1} \mathbf{A}^{i-1}\}) \otimes \mathbf{e}_i \qquad (\mathbf{U}_{n,1} = \mathbf{I}_n, \; \mathbf{U}_{1,n} = \mathbf{I}_n) \tag{675}$$

$$\overset{(307)}{=} \sum_{i=1}^{n} \left\{ \mathbf{e}_n^T \left(\frac{\partial}{\partial A_{\nu\mu}} \left[\sum_{k=1}^{n} \mathbf{e}_k^T \otimes (\mathbf{A}^{k-1}\mathbf{b}) \right]^{-1} \right) \mathbf{A}^{i-1} + \mathbf{e}_n^T \mathbf{L}_c^{-1} \frac{\partial \mathbf{A}^{i-1}}{\partial A_{\nu\mu}} \right\} \otimes \mathbf{e}_i$$

$$\overset{(381),(363)}{=} \sum_{i=1}^{n} \left\{ \mathbf{e}_n^T (-\mathbf{L}_c^{-1}) [\sum_{k=1}^{n} \frac{\partial \mathbf{A}^{k-1}\mathbf{b}}{\partial A_{\nu\mu}} \otimes \mathbf{e}_k^T] \mathbf{L}_c^{-1} \mathbf{A}^{i-1} + \mathbf{e}_n^T \mathbf{L}_c^{-1} \frac{\partial \mathbf{A}^{i-1}}{\partial A_{\nu\mu}} \right\} \otimes \mathbf{e}_i$$

$$= \sum_{i=1}^{n} \left\{ \mathbf{e}_n^T (-\mathbf{L}_c^{-1}) [\sum_{k=1}^{n} (\frac{\partial \mathbf{A}^{k-1}}{\partial A_{\nu\mu}} \mathbf{b}) \otimes \mathbf{e}_k^T] \mathbf{L}_c^{-1} \mathbf{A}^{i-1} + \mathbf{e}_n^T \mathbf{L}_c^{-1} \frac{\partial \mathbf{A}^{i-1}}{\partial A_{\nu\mu}} \right\} \otimes \mathbf{e}_i$$

$$\overset{(363)}{=} \sum_{i=1}^{n} \Big\{ \mathbf{e}_n^T \mathbf{L}_c^{-1} [-\sum_{k=2}^{n} \sum_{v=1}^{k-1} (\mathbf{A}^{v-1} \mathbf{E}_{\nu,\mu} \mathbf{A}^{k-v-1} \mathbf{b}) \otimes \mathbf{e}_k^T] \mathbf{L}_c^{-1} \mathbf{A}^{i-1}$$

$$+ \left(\sum_{v=1}^{i-1} \mathbf{A}^{v-1} \mathbf{E}_{\nu,\mu} \mathbf{A}^{i-v-1} \right)_{exept \;\, i=1} \Big\} \otimes \mathbf{e}_i \; , \tag{676}$$

where $\mathbf{L}_c$ could have been inserted from Eq.(656).

19.2 Second-Order Sensitivity Matrices

Suppose that some of of the entries of $\mathbf{A}$ can be altered by plant design activities. From Eq.(663), by additional differentiation with respect to p, q

$$\frac{\partial^2 \mathbf{L}_c}{\partial A_{pq} \partial A_{\nu\mu}} \overset{(382)}{=} \sum_{k=2}^{k=n} \sum_{i=1}^{i=k-1} \left(\frac{\partial \mathbf{A}^{i-1} \mathbf{E}_{\nu,\mu} \mathbf{A}^{k-i-1} \mathbf{b}}{\partial A_{pq}} \right) \otimes \mathbf{e}_k^T \tag{677}$$

$$\overset{(307)}{=} \sum_{k=2}^{k=n} \sum_{i=1}^{i=k-1} \Big[\frac{\partial \mathbf{A}^{i-1}}{\partial A_{pq}} \mathbf{E}_{\nu,\mu} \mathbf{A}^{k-i-1} \mathbf{b} + \mathbf{A}^{i-1} \mathbf{E}_{\nu,\mu} \frac{\partial \mathbf{A}^{k-i-1}}{\partial A_{pq}} \mathbf{b} \Big] \otimes \mathbf{e}_k^T$$

$$\overset{(371)}{=} \sum_{k=2}^{k=n} \sum_{i=1}^{i=k-1} \Big[\Big(\sum_{v=1}^{v=i-1} \mathbf{A}^{v-1}\mathbf{E}_{p,q}\mathbf{A}^{i-v-1} \Big)_{i\geq 2} \mathbf{E}_{\nu,\mu}\mathbf{A}^{k-i-1}\mathbf{b}$$

$$+\mathbf{A}^{i-1}\mathbf{E}_{\nu,\mu} \Big(\sum_{v=1}^{v=k-i-1} \mathbf{A}^{v-1}\mathbf{E}_{p,q}\mathbf{A}^{k-i-v-1} \Big)_{k-i\geq 2} \Big] \otimes \mathbf{e}_k^T \ . \quad (678)$$

Referring to Eq.(687), the sensitivity of the squared norm of the flat output with respect to A_{pq} is

$$S_{pq} \triangleq \frac{\partial^2 \|\mathbf{c}_f\|_F^2}{\partial A_{pq} \partial A_{\nu\mu}} \overset{(687)}{=} \frac{\partial}{\partial A_{pq}} [-2\mathbf{e}_n^T \mathbf{L}_c^{-1} \frac{\partial \mathbf{L}_c}{\partial A_{\nu\mu}} \mathbf{L}_c^{-1} \mathbf{L}_c^{-T} \mathbf{e}_n] \quad (679)$$

$$\overset{(307)}{=} -2\mathbf{e}_n^T \Big(\frac{\partial \mathbf{L}_c^{-1}}{\partial A_{pq}} \frac{\partial \mathbf{L}_c}{\partial A_{\nu\mu}} \mathbf{L}_c^{-1} \mathbf{L}_c^{-T} + \mathbf{L}_c^{-1} \frac{\partial^2 \mathbf{L}_c}{\partial A_{pq} \partial A_{\nu\mu}} \mathbf{L}_c^{-1} \mathbf{L}_c^{-T}$$

$$+ \mathbf{L}_c^{-1} \frac{\partial \mathbf{L}_c}{\partial A_{\nu\mu}} \frac{\partial \mathbf{L}_c^{-1}}{\partial A_{pq}} \mathbf{L}_c^{-T} + \mathbf{L}_c^{-1} \frac{\partial \mathbf{L}_c}{\partial A_{\nu\mu}} \mathbf{L}_c^{-1} \frac{\partial \mathbf{L}_c^{-T}}{\partial A_{pq}} \Big) \mathbf{e}_n \quad (680)$$

$$\overset{(363)}{=} -2\mathbf{e}_n^T \mathbf{L}_c^{-1} \Big(- \frac{\partial \mathbf{L}_c}{\partial A_{pq}} \mathbf{L}_c^{-1} \frac{\partial \mathbf{L}_c}{\partial A_{\nu\mu}} \mathbf{L}_c^{-1} + \frac{\partial^2 \mathbf{L}_c}{\partial A_{pq} \partial A_{\nu\mu}} \mathbf{L}_c^{-1}$$

$$- \frac{\partial \mathbf{L}_c}{\partial A_{\nu\mu}} \mathbf{L}_c^{-1} \frac{\partial \mathbf{L}_c}{\partial A_{pq}} \mathbf{L}_c^{-1} - \frac{\partial \mathbf{L}_c}{\partial A_{\nu\mu}} \mathbf{L}_c^{-1} \mathbf{L}_c^{-T} \frac{\partial \mathbf{L}_c^T}{\partial A_{pq}} \Big) \mathbf{L}_c^{-T} \mathbf{e}_n \ . \quad (681)$$

For a preselected (ν, μ), the matrix $\mathbf{S}(p, q)$ provides the combined sensitivity

$$\mathbf{S}(p, q) \triangleq \frac{\partial}{\partial A_{pq}} \frac{\partial \|\mathbf{c}_f\|_F^2}{\partial A_{\nu\mu}} \Big|_{\nu\mu} \mathbf{E}_{pq} \ , \quad (682)$$

in order to indicate which influence the (p, q) element of $\mathbf{A}$ has on $\frac{\partial \|\mathbf{c}_f\|_F^2}{\partial A_{\nu\mu}}$. This is a measure how some of the residual elements of $\mathbf{A}$ can change the sensitivity of the $\mathbf{c}_f$ norm with respect to an uncertainty. Lowering the sensitivity corresponds to a robustification of $\mathbf{c}_f$. Considering a single uncertainty at $A_{\nu\mu}$, it is desired to keep its influence on the flat output low, i.e., $\frac{\partial \|\mathbf{c}_f\|_F^2}{\partial A_{\nu\mu}}$ low but the corresponding $\mathbf{S}$ should be high. Influencing via $\mathbf{b}\mathbf{k}_r^T$ is impossible due to the orthogonal properties of $\mathbf{c}_f$.

The sensitivity of the Frobenius norm of the flat output is

$$\|\mathbf{c}_f\|_F^2 = \mathrm{tr}[\mathbf{c}_f \mathbf{c}_f^T] = \mathrm{tr}[\mathbf{L}_c^{-T} \mathbf{e}_n \mathbf{e}_n^T \mathbf{L}_c^{-1}] \quad (683)$$

$$\frac{\partial \|\mathbf{c}_f\|_F^2}{\partial \mathbf{k}} \overset{(669)}{=} \mathrm{trh}[\frac{\partial \mathbf{L}_c^{-T}}{\partial \mathbf{k}} \mathbf{e}_n \mathbf{e}_n^T \mathbf{L}_c^{-1} + [\mathbf{I}_n \otimes (\mathbf{L}_c^{-T} \mathbf{e}_n \mathbf{e}_n^T)] \frac{\partial \mathbf{L}_c^{-1}}{\partial \mathbf{k}}] \quad (684)$$

$$\frac{\partial \|\mathbf{c}_f\|_F^2}{\partial \mathbf{k}} = -\mathrm{trh} \Big[(\mathbf{I}_n \otimes \mathbf{L}_c^{-T}) \{ \frac{\partial \mathbf{L}_c^T}{\partial \mathbf{k}} \mathbf{L}_c^{-T} \mathbf{E}_{n,n} + [\mathbf{I}_n \otimes (\mathbf{E}_{n,n} \mathbf{L}_c^{-1})] \frac{\partial \mathbf{L}_c}{\partial \mathbf{k}} \} \mathbf{L}_c^{-1} \Big] \quad (685)$$

The derivative of the norm of the flat output coefficients with respect to an element of the coefficient matrix is

$$g(\nu, \mu) \triangleq \frac{\partial \|\mathbf{c}_f\|_F^2}{\partial A_{\nu\mu}} \overset{(685)}{=} -\mathrm{tr}[\mathbf{L}_c^{-T} \{ \frac{\partial \mathbf{L}_c^T}{\partial A_{\nu\mu}} \mathbf{L}_c^{-T} \mathbf{E}_{n,n} + \mathbf{E}_{n,n} \mathbf{L}_c^{-1} \frac{\partial \mathbf{L}_c}{\partial A_{\nu\mu}} \} \mathbf{L}_c^{-1}] \quad (686)$$

$$= -2[\mathbf{L}_c^{-1} \frac{\partial \mathbf{L}_c}{\partial A_{\nu\mu}} \mathbf{L}_c^{-1} \mathbf{L}_c^{-T}]_{nn} \ . \quad (687)$$

19.3 Incremental Representation

Utilize

$$(\mathbf{A}+\Delta\mathbf{A})^v \doteq \mathbf{A}^v + \sum_{i=1}^{i=v} \mathbf{A}^{v-i} \cdot \Delta\mathbf{A} \cdot \mathbf{A}^{i-1}, \quad v \geq 1 \ . \tag{688}$$

For an increment $\Delta\mathbf{A}$, we get the following increment of $\Delta\mathbf{L}_c$

$$\Delta\mathbf{L}_c \overset{(656)}{=} \sum_{i=2}^{n} \mathbf{e}_i^T \otimes [(\mathbf{A}+\Delta\mathbf{A})^{i-1}\mathbf{b} - \mathbf{A}^{i-1}\mathbf{b}] \tag{689}$$

$$\overset{\cdot(688)}{=} \sum_{i=2}^{n} \mathbf{e}_i^T \otimes \sum_{k=1}^{k=i-1} \mathbf{A}^{i-k-1}\Delta\mathbf{A} \cdot \mathbf{A}^{k-1}\mathbf{b} \ . \tag{690}$$

Referring to Eq.(490), the increment of the flat output parameter vector reads as

$$\Delta\mathbf{c}_f = \left(\sum_{i=1}^{n} \mathbf{e}_i^T \otimes [(\mathbf{A}+\Delta\mathbf{A})^{i-1}\mathbf{b}]\right)^{-T} \mathbf{e}_n - \left(\sum_{i=1}^{n} \mathbf{e}_i^T \otimes [\mathbf{A}^{i-1}\mathbf{b}]\right)^{-T} \mathbf{e}_n \tag{691}$$

$$\doteq -\mathbf{L}_c^{-T}\Big(\sum_{i=1}^{n} \mathbf{e}_i \otimes \sum_{k=1}^{k=i-1} \mathbf{b}^T\mathbf{A}^{T,i-k-1}\Delta\mathbf{A}^T\mathbf{A}^{T,k-1}\Big)\mathbf{L}_c^{-T}\mathbf{e}_n \ , \tag{692}$$

$$\|\Delta\mathbf{c}_f\|_F^2 = \mathbf{e}_n^T\mathbf{L}_c^{-1}\Delta\mathbf{L}_c \cdot \mathbf{L}_c^{-1}\mathbf{L}_c^{-T}\Delta\mathbf{L}_c^T \cdot \mathbf{L}_c^{-T}\mathbf{e}_n \ . \tag{693}$$

20 Stability Margin, Continuous-Time Systems, State Space

20.1 Minimum Hodograph Distance

Presupposing a stable control system, the closest point of the closed-loop hodograph (Mikhailov hodograph) to the origin follows from *Weinmann, A., 2005a*

$$\min_\omega |p_{cl}(j\omega)|^2 = \min_\omega p_{cl}^*(j\omega)p_{cl}(j\omega) \triangleq h_0^2 \ . \tag{694}$$

Using

$$p_{cl}(j\omega) = \det(j\omega\mathbf{I}_n - \mathbf{A}_{cl}) \tag{695}$$

and differentiating (*Brewer, J.W., 1978*), the optimum frequency ω_0 is determined by

$$f(\omega_0, \mathbf{A}_{cl}) = \mathrm{tr}[\mathbf{adj}(\omega_0^2\mathbf{I}_n + \mathbf{A}_{cl}^2)] = 0 \ , \tag{696}$$

and the minimum distance h_0 squared is

$$h_0^2 = \det(\omega_0^2\mathbf{I}_n + \mathbf{A}_{cl}^2) \ . \tag{697}$$

The differential sensitivity of h_0^2 with respect to $\mathbf{K}_y$ results as

$$\frac{\partial \det(\omega_0^2 \mathbf{I}_n + \mathbf{A}_{cl}^2)}{\partial \mathbf{K}_y} = \frac{\partial (h_0^2)}{\partial \mathbf{K}_y} \overset{(17)}{=} 2[\mathbf{C}\mathbf{A}_{cl}(\mathbf{adj}\mathbf{U})\mathbf{B}]^T \ . \tag{698}$$

In a reasonable range the minimum h_0 is obtained for $\omega_0 = 0$ in Eq.(696). Then, from Eq.(698)

$$\frac{\partial (h_0^2)}{\partial \mathbf{K}_y} = 2[\mathbf{C}\mathbf{A}_{cl}^{-1}\mathbf{B}]^T \det(\mathbf{A}_{cl}^2) \ . \tag{699}$$

This result shows a similarity to the gradient of holistic controllers, which minimize $\mathrm{tr}[\mathbf{A}_{cl}^{-1}]$. Their gradient is $\partial \mathrm{tr}[\mathbf{A}_{cl}^{-1}]/\partial \mathbf{K}_y = [\mathbf{C}\mathbf{A}_{cl}^{-2}\mathbf{B}]^T$ (*Weinmann, A., 2005*) .

20.2 Resonant Frequency ω_0 with Respect to the Output State Controller

We evaluate the increment of ω_0 when increasing $\mathbf{K}_y$ with an increment $\Delta\mathbf{K}_y$. From Eq.(696) the total differential expression is (*Weinmann, A., 2005a*)

$$d\,f = \frac{\partial f}{\partial \omega_0} d\omega_0 + \sum_{ij} \frac{\partial f}{\partial K_{y,ij}} dK_{y,ij} = \frac{\partial f}{\partial \omega_0} d\omega_0 + \mathrm{tr}[(\frac{\partial f}{\partial \mathbf{K}_y})^T d\mathbf{K}_y] = 0 \tag{700}$$

$$\mathbf{U} \ \triangleq \ \omega_0^2 \mathbf{I}_n + \mathbf{A}_{cl}^2 \overset{(543)}{=} \mathbf{H}(j\omega)\mathbf{H}^\star(j\omega), \tag{701}$$

$$\frac{d\omega_0}{d\mathbf{K}_y} = -\frac{\frac{\partial f}{\partial \mathbf{K}_y}}{\frac{\partial f}{\partial \omega_0}} = \frac{\{\mathbf{C}[\mathbf{U}^{-1} - \mathbf{I}_n \mathrm{tr}\mathbf{U}^{-1}]\mathbf{U}^{-1}\mathbf{A}_{cl}\mathbf{B}\}^T}{\omega_0\{(\mathrm{tr}[\mathbf{U}^{-1}])^2 - \mathrm{tr}[\mathbf{U}^{-2}]\}} \ . \tag{702}$$

20.3 Descriptor Systems

Often the process equations are composed not only of a linear combination of state variables but also a linear combination of the first derivatives. Then,

$$\mathbf{E}\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u} \tag{703}$$

is a generalization of $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u}$. A block diagram can be set up transmitting the linear combination $\mathbf{E}\dot{\mathbf{x}}$ to $\mathbf{x}$ by $\int\!\!\!\!E$, see Fig. 6.

A corresponding coefficient matrix $\mathbf{E}^{-1}\mathbf{A}$ cannot be used since $\mathbf{E}$ is singular in general. Defining

$$\mathbf{A}\mathbf{v} = \lambda_g \mathbf{E}\mathbf{v} \tag{704}$$

constitutes the generalized eigenvalue λ_g. Then, $\mathbf{A} - \lambda_g\mathbf{E}$ is termed the matrix pencil of the matrices $\mathbf{A}$ and $\mathbf{E}$ (*Laub, A.J., 2005*). The characteristic polynomial of the matrix pair $(\mathbf{A}, \mathbf{E})$ is $p(s) = \det(s\mathbf{E} - \mathbf{A})$. If, e.g., one of the equations of (703) is purely algebraic, one of the eigenvalues λ_g is $-\infty$, thus expressing the lost dynamics in this equation. From $\mathbf{A}\mathbf{a}_g = \lambda_g\mathbf{E}\mathbf{a}_g$, $\mathbf{a}_g$ is the right generalized eigenvector.

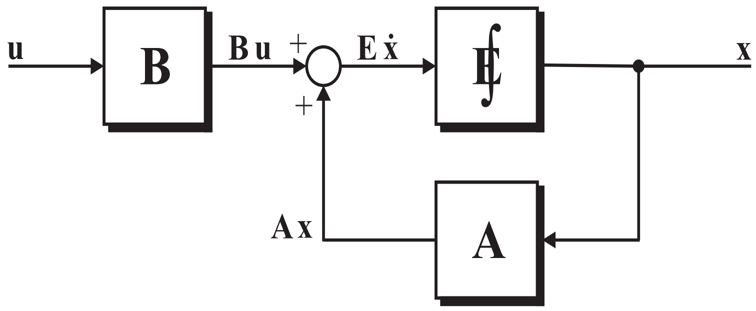

Figure 6: Block diagramm for a descriptor system

21 Stability Margin, Discrete-Time Systems, State Space

Consider the nth order discrete-time plant with sampling time T_s

$$\mathbf{x}(k+1) = \mathbf{\Phi}(T_s)\mathbf{x}(k) + \mathbf{\Psi}\mathbf{u}(k) \qquad \in \mathcal{R}^n \tag{705}$$

$$\mathbf{y}(k) = \mathbf{C}\mathbf{x}(k) + \mathbf{D}\mathbf{u}(k), \qquad \mathbf{D} = \mathbf{0}\ , \tag{706}$$

and a controller $\mathbf{u}(k) = \mathbf{K}_y\mathbf{y}(k) = \mathbf{K}_y\mathbf{C}\mathbf{x}(k)$. The input matrix $\mathbf{\Psi}$ is set constant although there is a slight dependence on T_s.

Then, the characteristic equation of the closed-loop system for $z = e^{j\omega T_s}$ is

$$p_{cl}(z)|_{z=e^{j\omega T_s}} = h(z)|_{z=e^{j\omega T_s}} \triangleq \det(e^{j\omega T_s}\mathbf{I}_n - \mathbf{\Phi}_{cl}) = \det(e^{j\omega T_s}\mathbf{I}_n - \mathbf{\Phi} - \mathbf{\Psi}\mathbf{K}_y\mathbf{C}) = 0\ . \tag{707}$$

Selecting the closed-loop polynomial and replacing $z = e^{j\omega T_s}$, the hodograph (Mikhailov plot) is obtained (*Weinmann, A., 2006b*). The closest point to the origin at frequency ω_0 results from

$$|\det(e^{j\omega T_s}\mathbf{I}_n - \mathbf{\Phi}_{cl})| \to \min_{\omega}\ , \tag{708}$$

where $\mathbf{\Phi}_{cl} = \mathbf{\Phi}(T_s) + \mathbf{\Psi}\mathbf{K}_y\mathbf{C}$. The minimum distance squared is

$$h_0^2 \triangleq \min_{\omega} \det(e^{j\omega T_s}\mathbf{I}_n - \mathbf{\Phi}_{cl})\det(e^{j\omega T_s}\mathbf{I}_n - \mathbf{\Phi}_{cl})^* \tag{709}$$

$$= \min_{\omega} \det(\mathbf{I}_n - 2\,\mathbf{\Phi}_{cl}\cos\omega T_s + \mathbf{\Phi}_{cl}^2)\ . \tag{710}$$

Differentiating Eq.(710) with respect to ω yields the solution ω_0

$$\mathrm{tr}[\mathbf{\Phi}_{cl}\mathbf{adj}(\mathbf{I}_n - 2\mathbf{\Phi}_{cl}\cos\omega_0 T_s + \mathbf{\Phi}_{cl}^2)] = 0 \tag{711}$$

$$\frac{\partial(h_o^2)}{\partial\mathbf{K}_y} = 2\mathbf{\Psi}\mathbf{U}_{d0}^{-T}[\mathbf{\Phi}_{cl}^T - \mathbf{I}_n\cos\omega_0 T_s]\mathbf{C}^T\det\mathbf{U}_{d0}\ , \tag{712}$$

$$\text{where} \quad \mathbf{U}_{d0} \triangleq \mathbf{I}_n - 2\mathbf{\Phi}_{cl}\cos\omega_0 T_s + \mathbf{\Phi}_{cl}^2. \tag{713}$$

Maximizing the *integral* of the stability margin in a discrete-time system

$$I = \int_0^{\omega_T} |p_{cl}(j\omega)|^2 d\omega = \int_0^{\omega_T} (\det \mathbf{U}_d)\ d\omega \to \max_{\mathbf{K}_y} \ , \tag{714}$$

where Eqs.(707) and $\mathbf{U}_d = \mathbf{I}_n - 2\mathbf{\Phi}_{cl}\cos\omega T_s + \mathbf{\Phi}_{cl}^2$ are used. For $\omega_T = 2\pi/T_s$, one finds

$$\frac{\partial I}{\partial \mathbf{K}_y} = 2\mathbf{\Psi}^T \int_0^{\omega_T/2} (\mathbf{\Phi}_{cl}^T - \mathbf{I}_n \cos\omega T_s)(\mathbf{adj}^T \mathbf{U}_d) d\omega \cdot \mathbf{C}^T \ . \tag{715}$$

22 Stability Margin, Continuous-Time Transfer Functions

22.1 Single-Input Single-Output Sytems

The closed-loop characteristic polynomial results from the return difference $1 + G(s)K(s)$ as

$$h(s) = (\mathbf{n}_k^T\ \mathbf{s})(\mathbf{n}_G^T\ \mathbf{s}) + (\mathbf{z}_k^T\ \mathbf{s})(\mathbf{z}_G^T\ \mathbf{s}) \ , \tag{716}$$

where

$$\mathbf{s} \stackrel{\triangle}{=} (1 \quad s \quad s^2 \ \ldots \ s^n)^T \stackrel{(4)}{=} \sum_0^n s^i \mathbf{e}_{i+1}^{[n+1]} \quad \in \mathcal{C}^{n+1}. \tag{717}$$

By replacing $s = j\omega$, the distance (squared) from the origin to the so-called Mikhailov plot is (*Weinmann, A., 2006*)

$$\begin{aligned} |h(j\omega)|^2 &= [n_K(s) n_G(s) + z_K(s) z_G(s)] \\ &\quad \times [n_K(s)\ n_G(s) + z_K(s)\ z_G(s)]^* \Big|_{s=j\omega} && (718) \\ &= \mathbf{n}_K^T\ \mathbf{s} \cdot \mathbf{n}_K^T\ \mathbf{s}^* \cdot \mathbf{n}_G^T\ \mathbf{s} \cdot \mathbf{n}_G^T\ \mathbf{s}^* + \ldots \Big|_{s=j\omega} && (719) \\ &= \mathbf{n}_K^T \mathbf{S} \mathbf{n}_K \cdot \mathbf{n}_G^T \mathbf{S} \mathbf{n}_G + \ldots \Big|_{s=j\omega} \ , && (720) \end{aligned}$$

where

$$\mathbf{s} \stackrel{\triangle}{=} (1 \quad s \quad s^2 \ \ldots \ s^n)^T \ , \quad \mathbf{S} \ \stackrel{\triangle}{=} \ \mathbf{s}\mathbf{s}^H \in \mathcal{C}^{(n+1)\times(n+1)} \ , \tag{721}$$

is the dyadic product and $\mathbf{s}^H \ \stackrel{\triangle}{=} \ (\mathbf{s}^*)^T$ is the "Hermite" conjugate transpose.

The coefficients of the polynomials $\mathbf{z}_K$ in the numerator are collected in a vector $\mathbf{z}_K$, which is the first half of the parameter vector $\mathbf{p}_{ij}$; the denominator polynomial is $n_K(s)$, which corresponds to the second half of the parameter vector $\mathbf{p}_{ij}$. Hence, one has

$$\mathbf{p}_K \stackrel{\triangle}{=} \begin{pmatrix} \mathbf{z}_K \\ \mathbf{n}_K \end{pmatrix}, \quad \mathbf{z}_K = \mathbf{P}_z \mathbf{p}_K, \quad \mathbf{n}_K = \mathbf{P}_n \mathbf{p}_K, \tag{722}$$

$$\mathbf{p}_G \stackrel{\triangle}{=} \begin{pmatrix} \mathbf{z}_G \\ \mathbf{n}_G \end{pmatrix}, \quad \mathbf{z}_G = \mathbf{P}_z \mathbf{p}_G, \quad \mathbf{n}_G = \mathbf{P}_n \mathbf{p}_G, \tag{723}$$

$$\text{where } \mathbf{P}_z \stackrel{\triangle}{=} (\mathbf{I}_n \ \vdots \ \mathbf{0}_n), \qquad \mathbf{P}_n \stackrel{\triangle}{=} (\mathbf{0}_n \ \vdots \ \mathbf{I}_n) \in \mathcal{R}^{n\times 2n} \ . \tag{724}$$

The derivative of the squared absolute value $|v|^2$ with respect to a real-valued parameter is

$$\frac{\partial |v|^2}{\partial \mathbf{p}} = \frac{\partial (vv^*)}{\partial \mathbf{p}} = \frac{\partial v}{\partial \mathbf{p}} v^* + v \frac{\partial v^*}{\partial \mathbf{p}} \tag{725}$$

$$\begin{aligned} &= \frac{\partial (\Re e\ v + j\Im m\ v)}{\partial \mathbf{p}} (\ \Re e\ v - j\Im m\ v) \\ &\quad + (\Re e\ v + j\Im m\ v) \frac{\partial (\Re e\ v - j\Im m\ v)}{\partial \mathbf{p}} \end{aligned} \tag{726}$$

$$\frac{\partial |v|^2}{\partial \mathbf{p}} = 2\left[(\Re e\ v)\ \frac{\partial (\Re e\ v)}{\partial \mathbf{p}} + (\Im m\ v) \frac{\partial (\Im m\ v)}{\partial \mathbf{p}} \right] = 2\ \Re e\ [v^* \frac{\partial v}{\partial \mathbf{p}}] \ . \tag{727}$$

Continuing from Eq.(720), one has

$$|h(j\omega)|^2 = [n_K n_K^* n_G n_G^* + z_K n_K^* z_G n_G^* + n_K z_K^* n_G z_G^* + z_K z_K^* z_G z_G^*]\Big|_{s=j\omega} \tag{728}$$

$$\begin{aligned} &\stackrel{(721)}{=} [\mathbf{p}_K^T \mathbf{P}_n^T \mathbf{S} \mathbf{P}_n \mathbf{p}_K \mathbf{n}_G^T \mathbf{S} \mathbf{n}_G + \mathbf{p}_K^T \mathbf{P}_z^T \mathbf{S} \mathbf{P}_n \mathbf{p}_K \mathbf{z}_G^T \mathbf{S} \mathbf{n}_G \\ &\quad + \mathbf{p}_K^T \mathbf{P}_n^T \mathbf{S} \mathbf{P}_z \mathbf{p}_K \mathbf{n}_G^T \mathbf{S} \mathbf{z}_G + \mathbf{p}_K^T \mathbf{P}_z^T \mathbf{S} \mathbf{P}_z \mathbf{p}_K \mathbf{z}_G^T \mathbf{S} \mathbf{z}_G]\Big|_{s=j\omega} \ . \end{aligned} \tag{729}$$

Because $\mathbf{n}_G$ is fixed it is kept unchanged in Eq.(730) in what follows; $\mathbf{n}_K$ and $\mathbf{z}_K$ are separated from $\mathbf{P}_K$ in order to be operated by the gradient method. Using the definitions

$$\mathbf{A}_G(s) \stackrel{\triangle}{=} \begin{pmatrix} \mathbf{n}_G^T \\ \mathbf{z}_G^T \end{pmatrix} \mathbf{S} (\mathbf{n}_G \ \vdots \ \mathbf{z}_G) = \begin{pmatrix} \mathbf{n}_G^T \mathbf{S} \mathbf{n}_G & \mathbf{n}_G^T \mathbf{S} \mathbf{z}_G \\ \mathbf{z}_G^T \mathbf{S} \mathbf{n}_G & \mathbf{z}_G^T \mathbf{S} \mathbf{z}_G \end{pmatrix} \tag{730}$$

$$\mathbf{B}_K(s) \stackrel{\triangle}{=} \begin{pmatrix} \mathbf{P}_n^T \\ \mathbf{P}_z^T \end{pmatrix} \mathbf{S} (\mathbf{P}_n \vdots \mathbf{P}_z) = \begin{pmatrix} \mathbf{P}_n^T \mathbf{S} \mathbf{P}_n & \mathbf{P}_n^T \mathbf{S} \mathbf{P}_z \\ \mathbf{P}_z^T \mathbf{S} \mathbf{P}_n & \mathbf{P}_z^T \mathbf{S} \mathbf{P}_z \end{pmatrix} \tag{731}$$

and an elementwise (Hadamard) product of matrices

$$\begin{aligned} \mathbf{F} &\stackrel{\triangle}{=} A_{G11}(s) \mathbf{B}_{K11}(s) + A_{G21}(s) \mathbf{B}_{K21}(s) + A_{G12}(s) \mathbf{B}_{K12}(s) + A_{G22}(s) \mathbf{B}_{K22}(s) \\ &\stackrel{\triangle}{=} \mathbf{A}_G(s)\ \cdot * \ \mathbf{B}_K(s) = \mathrm{btr}[\mathbf{A}_G^T(s) \mathbf{B}_K(s)], \end{aligned} \tag{732}$$

where btr$[\cdot]$ is a blocktrace. Applying block transposition in $\mathbf{F}$ does not influence the result. $\mathbf{S}$ is Hermite, $\mathbf{S} = \mathbf{S}^H$. Hence, one can state $\mathbf{F} = \mathbf{F}^H$ in this SISO case. Then,

$$|h(j\omega)|^2 = \mathbf{p}_K^T \mathbf{F}\Big|_{s=j\omega} \mathbf{p}_K \ . \tag{733}$$

The differential quotient is

$$\begin{aligned}
\frac{\partial |h(j\omega)|^2}{\partial \mathbf{p}_K} &= 2\Big(\mathbf{P}_n^T\mathbf{S}\mathbf{P}_n\mathbf{n}_G^T\mathbf{S}\mathbf{n}_G + \mathbf{P}_z^T\mathbf{S}\mathbf{P}_n\mathbf{z}_G^T\mathbf{S}\mathbf{n}_G \\
&\quad +\mathbf{P}_n^T\mathbf{S}\mathbf{P}_z\mathbf{n}_G^T\mathbf{S}\mathbf{z}_G + \mathbf{P}_z^T\mathbf{S}\mathbf{P}_z\mathbf{z}_G^T\mathbf{S}\mathbf{z}_G\Big)\Big|_{s=j\omega}\mathbf{p}_K && (734)\\
&= \frac{\partial \mathbf{p}_K^T\mathbf{F}\mathbf{p}_K}{\partial \mathbf{p}} = (\mathbf{F}+\mathbf{F}^T)\mathbf{p} && (735)\\
\frac{\partial |h(j\omega)|^2}{\partial \mathbf{p}_K}\Big|_{\omega=\omega_0} &= 2\Re e\,\mathbf{F}\Big|_{s=j\omega_0}\mathbf{p}_K = 2\Re e \sum_{i=1,j=1}^{i=2,j=2}[A_{Gij}\mathbf{B}_{Kij}]\Big|_{j\omega_0}\mathbf{p}_K\ . && (736)
\end{aligned}$$

22.2 Derivative of a Transfer Function

Suppose

$$K_{ij} = \frac{\mathbf{z}^T\mathbf{s}}{\mathbf{n}^T\mathbf{s}} = \frac{\mathbf{s}^T\mathbf{P}_z\mathbf{p}}{\mathbf{s}^T\mathbf{P}_n\mathbf{p}}\ . \qquad (737)$$

Differentiating with respect to the νth element of $\mathbf{p}$ yields (with $\mathbf{s}$ from Eq.(717) and $\mathbf{S} \stackrel{\triangle}{=} \mathbf{s}\mathbf{s}^H$)

$$\begin{aligned}
\frac{\partial K_{ij}(s)}{\partial p_\nu} &= \frac{1}{\mathbf{s}^T\mathbf{P}_n\mathbf{p}}\mathbf{s}^T\mathbf{P}_z\mathbf{e}_\nu + \frac{\mathbf{s}^T\mathbf{P}_z\mathbf{p}}{(\mathbf{s}^T\mathbf{P}_n\mathbf{p})^2}(-1)\mathbf{s}^T\mathbf{P}_n\mathbf{e}_\nu && (738)\\
&= \frac{1}{(\mathbf{s}^T\mathbf{P}_n\mathbf{p})^2}\mathbf{e}_\nu^T[\mathbf{P}_z^T\mathbf{S}\mathbf{P}_n - \mathbf{P}_n^T\mathbf{S}\mathbf{P}_z]\mathbf{p} && (739)\\
\frac{\partial K_{ij}(s)}{\partial \mathbf{p}} &= \frac{1}{(\mathbf{s}^T\mathbf{P}_n\mathbf{p})^2}[\mathbf{P}_z^T\mathbf{S}\mathbf{P}_n - \mathbf{P}_n^T\mathbf{S}\mathbf{P}_z]\mathbf{p}\ . && (740)
\end{aligned}$$

23 Stability Margin, Polynomial Matrices

Consider a multiple-input multiple-output controller and plant given by polynomial matrices $\in \mathcal{C}^{n\times n}$

$$\mathbf{K}(s) = \mathbf{N}_K^{-1}(s)\mathbf{Z}_K(s)\ , \qquad \mathbf{G}(s) = \mathbf{Z}_G(s)\mathbf{N}_G^{-1}(s). \qquad (741)$$

Then, the characteristic equation of the closed-loop system is

$$\det\mathbf{Q} \stackrel{\triangle}{=} \det[\mathbf{Z}_K(s)\mathbf{Z}_G(s) + \mathbf{N}_K(s)\mathbf{N}_G(s)] = 0. \qquad (742)$$

The Mikhailov hodograph results from the characteristic polynomial by replacing s by $j\omega$ and the results is (*Weinmann, A., 2006*)

$$\det[\mathbf{Z}_K(j\omega)\mathbf{Z}_G(j\omega) + \mathbf{N}_K(j\omega)\mathbf{N}_G(j\omega)] = \det\mathbf{Q} \qquad (743)$$

$$\frac{\partial h_0^2}{\partial p_{ijk}} = \frac{\partial|\det \mathbf{Q}|^2}{\partial p_{ijk}} = 2|\det \mathbf{Q}|^2 \quad \times \quad \Big[\Re e\,(\mathbf{Z}_G\mathbf{Q}^{-1})_{ji}\frac{\partial \Re e\, Z_K(i,j)}{\partial p_{ijk}} \\ -\Im m\,(\mathbf{Z}_G\mathbf{Q}^{-1})_{ji}\frac{\partial \Im m\, Z_K(i,j)}{\partial p_{ijk}} \\ +\Re e\,(\mathbf{N}_G\mathbf{Q}^{-1})_{ji}\frac{\partial \Re e\, N_K(i,j)}{\partial p_{ijk}} \\ -\Im m\,(\mathbf{N}_G\mathbf{Q}^{-1})_{ji}\frac{\partial \Re e\, N_K(i,j)}{\partial p_{ijk}}\Big]. \quad (744)$$

23.1 Inverse sensitivity

The maximum absolute value of the sensitivity function $S(j\omega)$ yields the worst steady-state oscillation and its corresponding frequency, the closed-loop system reacts with the maximum deviation. Referring to

$$\omega_0 = \arg\max_{\omega} |S(j\omega)| = \arg\min_{\omega} |S^{-1}(j\omega)|\ , \quad (745)$$

$$f^2 \triangleq |\det(\mathbf{I}+\mathbf{GK})|^2 = \det(\mathbf{I}+\mathbf{GK})\det(\mathbf{I}+\mathbf{G}^*\mathbf{K}^*)\ , \quad (746)$$

$$f_0 \triangleq \min_{\omega} |S^{-1}(j\omega)| = \min_{\omega} |\det(\mathbf{I}+\mathbf{GK})|\ . \quad (747)$$

The gradient of f_0^2 with respect to the matrix elements K_{ij} and its determining parameter $\mathbf{p}_{ij}$ is (using the reciprocal return difference)

$$\frac{\partial f_0^2}{\partial \mathbf{p}_{ij}} = 2|\det(\mathbf{I}+\mathbf{GK})|^2 \Big\{[\Re e\,(\mathbf{I}+\mathbf{GK})^{-1}\mathbf{G}]_{ji}\frac{\partial \Re e\, K_{ij}}{\partial \mathbf{p}_{ij}} \\ - [\Im m\,(\mathbf{I}+\mathbf{GK})^{-1}\mathbf{G}]_{ji}\frac{\partial \Im m\, K_{ij}}{\partial \mathbf{p}_{ij}}\Big\}\ . \quad (748)$$

23.2 Stability Margin, Power-Oriented Matrices

Suppose now, the polynomial matrix $\mathbf{N}(s)$ is split into separate coefficient matrices $\mathbf{N}_i$

$$\mathbf{N}(s) = \sum_0^n \mathbf{N}_i s^i \ \in \mathcal{C}^{m\times m}. \quad (749)$$

Concatenating these coefficient matrices in a hyper parameter matrix $\mathbf{P}_N$

$$\mathbf{P}_N \triangleq (\mathbf{N}_0 \vdots \mathbf{N}_1 \vdots \ldots \vdots \mathbf{N}_i \vdots \ldots \vdots \mathbf{N}_n) \in \mathcal{R}^{m\times m(n+1)}\ . \quad (750)$$

Consider both the plant and controller given by fractions of polynomial matrices $\mathbf{G}(s) \triangleq \mathbf{Z}_G\mathbf{N}_G^{-1}$ and $\mathbf{K}(s) \triangleq \mathbf{N}_K^{-1}\mathbf{Z}_K$, respectively; in this specific order. Moreover,

$$\mathbf{N} \triangleq \mathbf{N}_K\mathbf{N}_G, \quad \mathbf{Z} \triangleq \mathbf{Z}_K\mathbf{Z}_G \quad (751)$$

$$\mathbf{N}_G \triangleq \mathbf{P}_{NG}\mathbf{S}_\sigma, \quad \mathbf{N}_K \triangleq \mathbf{P}_{NK}\mathbf{S}_\sigma, \quad \mathbf{Z}_G \triangleq \mathbf{P}_{ZG}\mathbf{S}_\sigma, \quad \mathbf{Z}_K \triangleq \mathbf{P}_{ZK}\mathbf{S}_\sigma, \tag{752}$$

where $\mathbf{P}_{ZK},\ \mathbf{P}_{NK},\ \mathbf{P}_{NG},\ \mathbf{P}_{ZG}\ \in \mathcal{R}^{m\times m(n+1)}$ and

$$\mathbf{S}_\sigma \triangleq \begin{pmatrix} \mathbf{I}_m \\ \mathbf{I}_m s \\ \vdots \\ \mathbf{I}_m s^i \\ \vdots \\ \mathbf{I}_m s^n \end{pmatrix} \overset{(717)}{\equiv} \mathbf{s} \otimes \mathbf{I}_m \quad \in \mathcal{C}^{m(n+1)\times m}, \qquad \begin{matrix} \text{e.g., for} \\ m = 3 \\ n = 2 : \end{matrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ s & 0 & 0 \\ 0 & s & 0 \\ 0 & 0 & s \\ s^2 & 0 & 0 \\ 0 & s^2 & 0 \\ 0 & 0 & s^2 \end{pmatrix}. \tag{753}$$

Then, the open-loop transfer function of the multivariable system is $\mathbf{N}^{-1}\mathbf{Z}$. From the return difference $\mathbf{I}+\mathbf{K}(s)\mathbf{G}(s) = \mathbf{I}+\mathbf{N}_K^{-1}\mathbf{Z}_K\mathbf{Z}_G\mathbf{N}_G^{-1}$, by pre- and postmultiplying there results $\mathbf{N}+\mathbf{Z}$. Keeping plant and controller separated, the characteristic polynomial of the closed-loop system is

$$h(s) = \det[\mathbf{N}(s) + \mathbf{Z}(s)] = \det[\mathbf{N}_K(s)\mathbf{N}_G(s) + \mathbf{Z}_K(s)\mathbf{Z}_G(s)] \ . \tag{754}$$

From the frequency hodograph (Mikhailov hodograph) the stability margin is defined $h_0 \triangleq \min_\omega |h(j\omega)|^2$. Changing $h_0^2 = |h(j\omega_0)|^2$ stepwise with respect to $\mathbf{P}_{ZK}$ and $\mathbf{P}_{NK}$, one finds (*Weinmann, A., 2007a*)

$$0.5\frac{\partial(h_0^2)}{\partial\binom{\mathbf{P}_{ZK}}{\mathbf{P}_{NK}}} = |\det(\mathbf{N}+\mathbf{Z})|^2\ \Re e\ [\mathbf{S}_\sigma(\mathbf{Z}_G \ \vdots\ \mathbf{N}_G)\mathrm{bd}\{(\mathbf{N}+\mathbf{Z})^{-1}\}]^T \ . \tag{755}$$

The operator "bd" terms block diagonal, i.e., $\mathrm{bd}(\mathbf{J}) \triangleq \begin{pmatrix} \mathbf{J} & \mathbf{0} \\ \mathbf{0} & \mathbf{J} \end{pmatrix}$. The problem subject to the condition $||\Delta\mathbf{P}_K||_F = n_{Ko}$, see *Weinmann, A., 2007a*.

24 Nyquist Encirclements, Transfer Functions

The encirclements of $1 + F_o(j\omega)$ around the origin, corresponding to the Nyquist criterion as derived from the open-loop transfer function (not coinciding with the encirclements of the Mikhailov hodograph), are determined by (*Weinmann, A., 2007b*)

$$1 + F_o(j\omega) \triangleq \frac{z(s)}{p(s)}|_{s=j\omega} \triangleq \frac{\mathbf{z}^T\mathbf{s}}{\mathbf{n}^T\mathbf{s}}|_{s=j\omega} \quad \triangleq \quad f\ e^{j\beta} \tag{756}$$

$$\Big[\ln(\mathbf{z}^T\mathbf{s}) - \ln(\mathbf{n}^T\mathbf{s})\Big]_{s=j\omega} \quad \triangleq \quad \ln f + j\beta. \tag{757}$$

The abbreviations $\mathbf{s} \stackrel{\triangle}{=} (1 \quad s^1 \quad s^2 \quad \dots \quad s^n)$ and $\frac{\partial \mathbf{s}}{\partial s} = (0 \quad 1 \quad 2s \quad \dots \quad ns^{n-1})$ and $\frac{\partial^2 \mathbf{s}}{\partial s^2} = (0 \ \vdots \ 0 \ \vdots \ 2 \ \vdots \ 6s \ \vdots \ 12s^2 \ \vdots \dots \vdots \ n(n-1)s^{n-2})^T$ are used, different from $\mathbf{s}$ in Eqs.(717) and (721). Then,

$$\beta = [-j\ln(\mathbf{z}^T\mathbf{s}) + j\ln(\mathbf{n}^T\mathbf{s}) + \ln f]_{s=j\omega} = \Im m\ [\ln(\mathbf{z}^T\mathbf{s}) - \ln(\mathbf{n}^T\mathbf{s})]_{s=j\omega} \tag{758}$$

$$\frac{\partial \beta}{\partial \omega} = \frac{\partial \beta}{\partial s}\frac{\partial s}{\partial \omega} = \frac{\partial \beta}{\partial s} \cdot j = \left[(\frac{1}{\mathbf{z}^T\mathbf{s}}\mathbf{z}^T - \frac{1}{\mathbf{n}^T\mathbf{s}}\mathbf{n}^T)\frac{\partial \mathbf{s}}{\partial s} \right]_{s=j\omega} + j\frac{\partial f}{\partial \omega} \tag{759}$$

$$\frac{\partial \beta}{\partial \omega} = \Re e \left[(\frac{\mathbf{z}^T}{\mathbf{z}^T\mathbf{s}} - \frac{\mathbf{n}^T}{\mathbf{n}^T\mathbf{s}})\frac{\partial \mathbf{s}}{\partial s} \right]_{s=j\omega} \tag{760}$$

$$\frac{\partial^2 \beta}{\partial \omega^2} = \Im m \left[\frac{1}{(\mathbf{z}^T\mathbf{s})^2}(\mathbf{z}^T\frac{\partial \mathbf{s}}{\partial s})^2 - \frac{1}{(\mathbf{n}^T\mathbf{s})^2}(\mathbf{n}^T\frac{\partial \mathbf{s}}{\partial s})^2 - (\frac{\mathbf{z}^T}{\mathbf{z}^T\mathbf{s}} - \frac{\mathbf{n}^T}{\mathbf{n}^T\mathbf{s}})\frac{\partial^2 \mathbf{s}}{\partial s^2} \right]_{s=j\omega}. \tag{761}$$

24.1 Derivative of the Closed-Loop Encirclements

For a linear continuous-time dynamic system with the coefficient matrix $\mathbf{A}_{cl}$, stability requires that arg $h(j\omega)$ increases monotonically with rising ω, where

$$h(j\omega) \stackrel{\triangle}{=} \det(j\omega\mathbf{I}_n - \mathbf{A}_{cl}) \tag{762}$$

$$\mathbf{A}_{cl} \stackrel{\triangle}{=} \mathbf{A} + \mathbf{B}\mathbf{K}_y\mathbf{C} \tag{763}$$

and $\mathbf{A} \in \mathcal{R}^{n\times n}$; $\mathbf{B} \in \mathcal{R}^{n\times m}$; $\mathbf{C} \in \mathcal{R}^{r\times n}$, $\mathbf{K}_y \in \mathcal{R}^{m\times r}$. The open-loop coefficient matrix is $\mathbf{A}$, the closed-loop one $\mathbf{A}_{cl}$; $\mathbf{K}_y$ is the output state controller matrix. Focus is put on the phase of $h(j\omega) = h_0 e^{j\alpha(\omega)}$ and its differential quotients with respect to ω (*Weinmann, A., 2007b*)

$$\frac{\partial \alpha(\omega)}{\partial \omega} = \Re e\ \mathrm{tr}[(j\omega\mathbf{I}_n - \mathbf{A}_{cl})^{-1}] \tag{764}$$

$$\frac{\partial^2 \alpha(\omega)}{\partial \omega^2} \stackrel{(764)}{=} \Re e\ \mathrm{tr}[(j\omega\mathbf{I}_n - \mathbf{A}_{cl})^{-2} \cdot j] = -\Im m\ \mathrm{tr}[(j\omega\mathbf{I}_n - \mathbf{A}_{cl})^{-2}]\ . \tag{765}$$

25 Phase and Gain Margin Gradient

Consider a nominal plant $G(s) = \frac{z_G(s)}{n_G(s)}$ and a nominal controller $K(s) = \frac{z_K(s)}{p_K(s)} = \frac{\mathbf{p}_{Kz}\mathbf{s}}{\mathbf{p}_{Kn}\mathbf{s}}$ where $\mathbf{s} \stackrel{\triangle}{=} (1 \ \ s \ \ s^2 \ \ \dots \ \ s^n)^T \ \in \mathcal{C}^{n+1}$ from Eq.(717). The order of the plant and the controller is n_G and n_K, respectively. The coefficients of the polynomials $z_K(s)$ in the numerator are collected in a vector $\mathbf{p}_{Kz}$, which is the first half of the parameter vector $\mathbf{p}_K$; the denominator polynomial is $p_K(s)$ corresponding to the second half of the parameter vector $\mathbf{p}_K$. Hence, one has

$$\mathbf{p}_K \triangleq \begin{pmatrix}\mathbf{p}_{Kz}\\ \mathbf{p}_{Kn}\end{pmatrix}, \quad \mathbf{p}_{Kz} = \mathbf{P}_z\mathbf{p}_K, \quad \mathbf{p}_{Kn} = \mathbf{P}_n\mathbf{p}_K \quad \in \mathcal{R}^{n_K+1}, \tag{766}$$

$$\mathbf{p}_G \triangleq \begin{pmatrix}\mathbf{p}_{Gz}\\ \mathbf{p}_{Gn}\end{pmatrix}, \quad \mathbf{p}_{Gz} = \mathbf{P}_z\mathbf{p}_G, \quad \mathbf{p}_{Gn} = \mathbf{P}_n\mathbf{p}_G \quad \in \mathcal{R}^{n_G+1}, \tag{767}$$

$$\text{where } \mathbf{P}_z \triangleq (\mathbf{I}_{n+1} \vdots \mathbf{0}_{n+1}), \qquad \mathbf{P}_n \triangleq (\mathbf{0}_{n+1} \vdots \mathbf{I}_{n+1}) \in \mathcal{R}^{(n+1)\times 2(n+1)} \tag{768}$$

for $n = n_K$ or $n = n_G$, respectively. The open-loop transfer function is

$$F_0(s) = G(s)K(s) = G(s)\frac{\mathbf{p}_{Kz}\mathbf{s}}{\mathbf{p}_{Kn}\mathbf{s}}\Big|_{s=j\omega} \triangleq f_0 e^{j\beta_0} \tag{769}$$

$$\ln G + \ln(\mathbf{p}_{Kz}\mathbf{s}) - \ln(\mathbf{p}_{Kn}\mathbf{s}) = \ln f_0 + j\beta_0 \ . \tag{770}$$

$$\beta_0 = \Im m\ [\ln G(s) + \ln(\mathbf{p}_{Kz}\mathbf{s}) - \ln(\mathbf{p}_{Kn}\mathbf{s})]_{s=j\omega_D} = \beta_0(\omega_D)\ . \tag{771}$$

The classical phase margin α_R results from

$$\alpha_R = \pi + \arg G(s)K(s)\Big|_{s=j\omega_D} = \pi + \beta_0(\omega_D)\ , \qquad |G(j\omega_D)K(j\omega_D)| = 1\ , \tag{772}$$

$$\frac{\partial \alpha_R}{\partial \mathbf{p}_K} = \Im m\ [(\frac{\mathbf{P}_z^T}{\mathbf{p}_K^T\mathbf{P}_z^T\mathbf{s}} - \frac{\mathbf{P}_n^T}{\mathbf{p}_K^T\mathbf{P}_n^T\mathbf{s}})\mathbf{s}]_{s=j\omega_D} \tag{773}$$

and in state space

$$\frac{\partial \alpha_R}{\partial \mathbf{k}} = \Im m\ [\frac{(s\mathbf{I}-\mathbf{A})^{-1}\mathbf{b}}{\mathbf{k}^T(s\mathbf{I}-\mathbf{A})^{-1}\mathbf{b}}\Big|_{s=j\omega_D}]\ . \tag{774}$$

The gain margin gradient is

$$\frac{\partial \ln 1/A_R}{\partial \mathbf{p}_K} = -\frac{\partial \ln A_R}{\partial \mathbf{p}_K} = \frac{\partial f_0}{\partial \mathbf{p}_K} = \Re e\ [\frac{1}{\mathbf{p}_K^T\mathbf{P}_z^T\mathbf{s}}\mathbf{P}_z^T - \frac{1}{\mathbf{p}_K^T\mathbf{P}_n^T\mathbf{s}}\mathbf{P}_n^T]_{j\omega_R}\ , \tag{775}$$

where $\arg F_0(j\omega_R) = -\pi$.

Based on frequency domain functions, for the crossover frequency gradient one finds

$$\frac{\partial \omega_D}{\partial \mathbf{p}_K} = -\frac{|G|^2\ \Re e\ [K^*\frac{\partial K}{\partial \mathbf{p}_K}]}{|K|^2\ \Re e\ [G^*\frac{\partial G}{\partial \omega_D}] + |G|^2\ \Re e\ [K^*\frac{\partial K}{\partial \omega_D}]}\ . \tag{776}$$

For the plant $(\mathbf{A}, \mathbf{b})$ and the state space controller $\mathbf{k}$

$$\frac{\partial \omega_D}{\partial \mathbf{k}} = \frac{\Re e\ [\mathbf{k}^T\mathbf{R}(-j\omega)\mathbf{b}\cdot\mathbf{R}(j\omega)\mathbf{b}]}{\Im m\ [\mathbf{k}^T\mathbf{R}(-j\omega)\mathbf{b}\cdot\mathbf{k}^T\mathbf{R}^2(j\omega)\mathbf{b}]}, \quad \text{where } \mathbf{R}(j\omega) \triangleq j\omega\mathbf{I} - \mathbf{A} \tag{777}$$

(*Weinmann, A., 2008*).

26 Limit Cycle Frequency

Consider a single-loop system with a series interconnection of a nonlinear element N, the plant G and a linear controller K. Presuppose the existence of a limit cycle with control error e_r and frequency ω_r. Then, for the gradient of the limit cycle error amplitude and frequency, we utilize $1 + N(e_r)G(j\omega_r)K(j\omega_r) = 0$. Defining

$$a \triangleq \frac{\partial N}{\partial e_r} GK, \quad b \triangleq N(\frac{\partial G}{\partial \omega_r} K, \quad c \triangleq NG, \tag{778}$$

$$\begin{pmatrix} \frac{\partial e_r}{\partial \mathbf{p}_K} \\ \frac{\partial \omega_r}{\partial \mathbf{p}_K} \end{pmatrix} = - \begin{pmatrix} \Re e[a]\ \mathbf{I} & \Re e[b]\ \mathbf{I} \\ \Im m[a]\ \mathbf{I} & \Im m[b]\ \mathbf{I} \end{pmatrix}^{-1} \begin{pmatrix} \Re e[c]\ \mathbf{I} & -\Im m[c]\ \mathbf{I} \\ \Im m[c]\ \mathbf{I} & \Re e[c]\ \mathbf{I} \end{pmatrix} \begin{pmatrix} \Re e\ \frac{\partial K}{\partial \mathbf{p}_K} \\ \Im m\ \frac{\partial K}{\partial \mathbf{p}_K} \end{pmatrix} \tag{779}$$

is achieved (*Weinmann, A., 2008a*).

27 Interaction Energy Gradient, Continuous Time

For the closed-loop system of Fig. 7 with output controller $\mathbf{K}_y$ and prefilter $\mathbf{V}_F$, one has

$$\mathbf{y}(s) = \mathbf{C}(s\mathbf{I}_n - \mathbf{A} - \mathbf{B}\mathbf{K}_y\mathbf{C})^{-1}\mathbf{B}\mathbf{V}_F\mathbf{y}_{ref}(s)\ . \tag{780}$$

From the input $y_{ref,k} = 1$ to the output $y_i(s)$, the impulse response is

$$y_i(s) = \mathbf{c}_i^T(s\mathbf{I}_n - \mathbf{A}_{cl})^{-1}\mathbf{w}_k \tag{781}$$

where $\mathbf{w}_k$ is the corresponding column of $\mathbf{W} \triangleq \mathbf{B}\mathbf{V}_F$.

The controllability Gramian $\mathbf{L}_{ck}$ for the open-loop single input system with input matrix $\mathbf{b}_k$ results from

$$\mathbf{A}\mathbf{L}_{ck} + \mathbf{L}_{ck}\mathbf{A}^T + \mathbf{b}_k\mathbf{b}_k^T = \mathbf{0}\ . \tag{782}$$

In the Lyapunov Eq.(782), $\mathbf{A}$ has to be replaced by $\mathbf{A}_{cl} \triangleq \mathbf{A} + \mathbf{B}\mathbf{K}_y\mathbf{C}$ for the closed-loop system of Fig. 7, and $\mathbf{b}_k$ by $\mathbf{w}_k$.

To design a system with low interaction, the output y_i should be influenced as low as possible by $y_{ref,k}$ for $k \neq i$. Hence, for unequally enumerated reference and output signals the entire energy associated with the interaction should obey to

$$\sum_{i=1,\ k=1;\ i\neq k}^{r} \mathbf{c}_i^T\mathbf{L}_{ck}\mathbf{c}_i \quad \rightarrow \quad \min_{\mathbf{K}_y}\ . \tag{783}$$

The controller $\mathbf{K}_y$ is used to drive the system to a low interaction level.

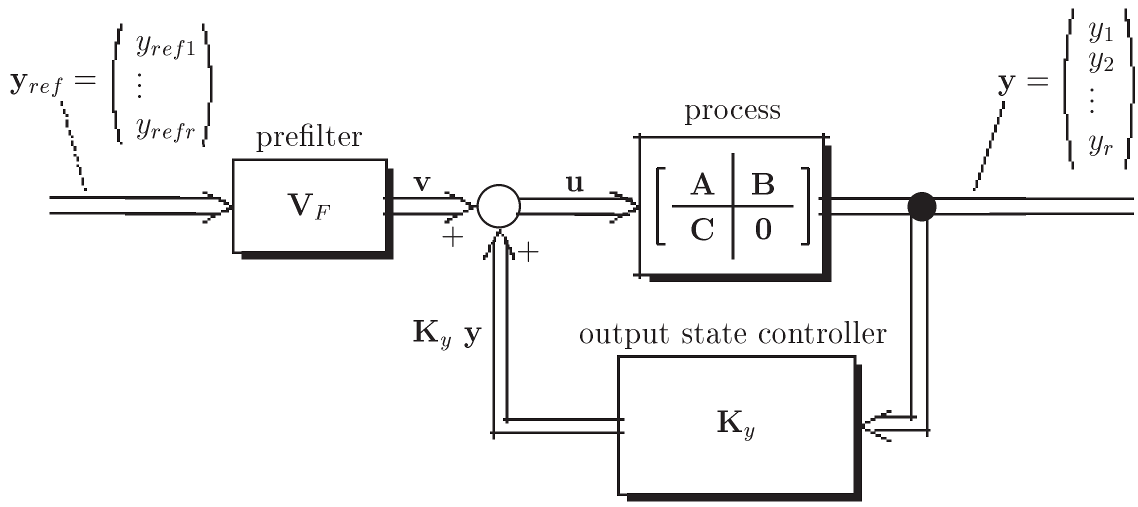

Figure 7: Output state controller

Investigating the increments $\Delta\mathbf{L}_{ck}$ caused by $\Delta\mathbf{K}_y$ when referring to the Lyapunov equation Eq.(782) and abbreviating

$$\mathbf{E}_1 \triangleq (\mathbf{I}_n \otimes \mathbf{A}_{cl} + \mathbf{A}_{cl} \otimes \mathbf{I}_n)^{-1} \in \mathcal{R}^{n^2 \times n^2} , \tag{784}$$

one finds (*Weinmann, A., 2005b*)

$$\frac{\partial I_{G,ik}}{\partial \mathbf{K}_y} \overset{(351),(324)}{=} -\{\mathbf{I}_m \otimes [(\mathbf{c}_i^T \otimes \mathbf{c}_i^T)\mathbf{E}_1(\mathbf{I}_{n^2} + \mathbf{U}_{nn})[\mathbf{B} \otimes (\mathbf{C}\mathbf{L}_{ck})^T]\} \times \\ \times[(\mathrm{col}\,\mathbf{I}_m) \otimes \mathbf{I}_r] \in \mathcal{R}^{m \times r} \tag{785}$$

$$\frac{\partial I_{G,ik}}{\partial \mathbf{K}_y} \overset{(70)}{=} -\{\mathrm{loc}_r[\mathbf{B}^T \otimes (\mathbf{C}\mathbf{L}_{ck})](\mathbf{I}_{n^2} + \mathbf{U}_{nn})\mathbf{E}_1^T(\mathbf{c}_i \otimes \mathbf{c}_i)\}^T , \tag{786}$$

where loc_r is a command for decolumnizing a column such that a matrix with r rows is produced.

Defining $\mathbf{X}_1$ via

$$\mathbf{E}_1^T(\mathbf{c}_i \otimes \mathbf{c}_i) \triangleq \mathrm{col}\ \mathbf{X}_1 \tag{787}$$

$$(\mathbf{I}_n \otimes \mathbf{A}_{cl}^T + \mathbf{A}_{cl}^T \otimes \mathbf{I}_n)^{-1}(\mathbf{c}_i \otimes \mathbf{c}_i) = \mathrm{col}\ \mathbf{X}_1 \tag{788}$$

$$(\mathbf{c}_i \otimes \mathbf{c}_i) = \mathrm{col}(\mathbf{c}_i\mathbf{c}_i^T) = (\mathbf{I}_n \otimes \mathbf{A}_{cl}^T + \mathbf{A}_{cl}^T \otimes \mathbf{I}_n)\mathrm{col}\ \mathbf{X}_1 \tag{789}$$

$$\mathbf{A}_{cl}^T\mathbf{X}_1 + \mathbf{X}_1\mathbf{A}_{cl} = \mathbf{c}_i\mathbf{c}_i^T , \tag{790}$$

one finally has

$$\frac{\partial I_{G,ik}}{\partial \mathbf{K}_y} = -\{\mathrm{loc}_r[\mathbf{B}^T \otimes (\mathbf{C}\mathbf{L}_{ck})]\ 2\ \mathrm{col}\,\mathbf{X}_1\}^T \tag{791}$$

$$= -2\{\mathrm{loc}_r\mathrm{col}(\mathbf{C}\mathbf{L}_{ck})\mathbf{X}_1\mathbf{B}\}^T = -2\mathbf{B}^T\mathbf{X}_1\mathbf{L}_{ck}\mathbf{C}^T , \tag{792}$$

where $\mathbf{X}_1$ and $\mathbf{L}_{ck}$ result from Eqs.(790) and (782), respectively.

27.1 Control and Output Energy, Hankel Norm $\|\mathbf{G}(s)\|_H$

Assume $\mathbf{G}(s) \in H_\infty$ and $\mathbf{G}(s) = \mathbf{C}(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{B} + \mathbf{D}$. Then, solve the Lyapunov equations

$$\mathbf{A}\mathbf{L}_c + \mathbf{L}_c\mathbf{A}^T + \mathbf{B}\mathbf{B}^T = \mathbf{0} \tag{793}$$

$$\mathbf{A}^T\mathbf{L}_o + \mathbf{L}_o\mathbf{A} + \mathbf{C}^T\mathbf{C} = \mathbf{0} \ . \tag{794}$$

The solutions $\mathbf{L}_c$ and $\mathbf{L}_o$, i.e, the controllability and observability gramians, respectively, are real symmetric matrices. Then, the Hankel norm is given by

$$\|\mathbf{G}(s)\|_H \triangleq \sigma_{H\max}[\mathbf{G}(s)] \triangleq \sigma_{\max}[(\mathbf{L}_c\mathbf{L}_o)^{1/2}] = \|(\mathbf{L}_c\mathbf{L}_o)^{1/2}\|_s = \sqrt{\lambda_{\max}[\mathbf{L}_c\mathbf{L}_o]} \tag{795}$$

The matrix $\mathbf{D}$ does not influence $\|\mathbf{G}\|_H$.

If $\mathbf{A}$ is stable the Hankel singular value is given by

$$\sigma_{Hi}[\mathbf{G}(s)] \triangleq \sqrt{\lambda_i[\mathbf{L}_c\mathbf{L}_o]} \ . \tag{796}$$

The matrix $\mathbf{L}_c$ is proved to be the unique solution of the Lyapunov equation Eq.(793).

The controllability gramian is useful for determining the minimum control energy

$$\min_{\mathbf{u}} \int_{-\infty}^{0} \mathbf{u}^T(t)\mathbf{u}(t)dt \quad \text{subject to} \quad \mathbf{x}(0) = \mathbf{x}_o \tag{797}$$

Then, see *Glover, K., 1984* ,

$$\arg\min_{\mathbf{u}} \int_{-\infty}^{0} \mathbf{u}^T(t)\mathbf{u}(t)dt = \mathbf{u}^\star(t) = \mathbf{B}^T e^{-\mathbf{A}^T t}\mathbf{L}_c^{-1}\mathbf{x}_o \tag{798}$$

$$\text{and} \quad \min_{\mathbf{u}} \int_{-\infty}^{0} \mathbf{u}^T(t)\mathbf{u}(t)dt = \mathbf{x}_o^T\mathbf{L}_c^{-1}\mathbf{x}_o \ . \tag{799}$$

If $\mathbf{u}(t) = \mathbf{0}$ for $t \geq 0$ and the system is released from $\mathbf{x}_o$, the output energy is $\int_0^\infty \mathbf{y}^T(t)\mathbf{y}(t)dt = \mathbf{x}_o^T\mathbf{L}_o\mathbf{x}_o$. The observability gramian relates the output energy to the initial conditions.

28 Interaction Energy Gradient, Discrete Time

Based on the continuous-time single-input single-output system

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{b}u(t) \ , \qquad y(t) = \mathbf{c}^T\mathbf{x}(t) \ , \tag{800}$$

employing a sampler with sampling time T_s and a zero-order hold, one has a discrete-time system obeying the sampling instants kT_s

$$\mathbf{x}(k+1) = \mathbf{\Phi}\mathbf{x}(k) + \boldsymbol{\psi}u(k)\ , \qquad y(k) = \mathbf{c}^T\mathbf{x}(k)\ , \tag{801}$$

where

$$\mathbf{\Phi} = \mathbf{\Phi}(T_s) = e^{\mathbf{A}T_s}\ , \qquad \boldsymbol{\psi}(T_s) = \mathbf{A}^{-1}[\mathbf{\Phi}(T_s) - \mathbf{I}]\mathbf{b}\ . \tag{802}$$

If the stable discrete-time single-input single-output system is excited by a single input $u(0) = 1, u(k) = 0\ \forall\ k \geq 1$, $\mathbf{x}(0) = \mathbf{0}$, then

$$y(k+1) = \mathbf{c}^T\mathbf{x}(k+1) = \mathbf{c}^T\mathbf{\Phi}^k\boldsymbol{\psi} \tag{803}$$

$$\sum_{k=0}^{\infty} y^2(k+1) = \mathbf{c}^T(\sum_{k=0}^{\infty}\mathbf{\Phi}^k\boldsymbol{\psi}\boldsymbol{\psi}^T\mathbf{\Phi}^{k,T})\mathbf{c} \ \stackrel{\triangle}{=}\ \mathbf{c}^T\mathbf{L}_c\mathbf{c}\ . \tag{804}$$

Simply splitting the controllability Gramian $\mathbf{L}_c$ leads to

$$\mathbf{L}_c \ \stackrel{\triangle}{=}\ \sum_{k=0}^{\infty}\mathbf{\Phi}^k\boldsymbol{\psi}\boldsymbol{\psi}^T\mathbf{\Phi}^{k,T} = \boldsymbol{\psi}\boldsymbol{\psi}^T + \sum_{k=1}^{\infty}\mathbf{\Phi}^k\boldsymbol{\psi}\boldsymbol{\psi}^T\mathbf{\Phi}^{k,T} \tag{805}$$

$$\mathbf{L}_c = \boldsymbol{\psi}\boldsymbol{\psi}^T + \sum_{k=0}^{\infty}\mathbf{\Phi}^{k+1}\boldsymbol{\psi}\boldsymbol{\psi}^T\mathbf{\Phi}^{k+1,T} = \boldsymbol{\psi}\boldsymbol{\psi}^T + \mathbf{\Phi}\mathbf{L}_c\mathbf{\Phi}^T \tag{806}$$

$$\mathrm{col}\mathbf{L}_c = (\mathbf{I} - \mathbf{\Phi}\otimes\mathbf{\Phi})^{-1}\mathrm{col}(\boldsymbol{\psi}\boldsymbol{\psi}^T)\ , \tag{807}$$

where "col" is an operator columnizing the matrix. Eq.(806) is the appropriate Lyapunov equation.

Consider the output energy of a discrete-time single-input single-output system excited by a single input $u(0) = 1$. Referring to Eq.(804) and $\mathbf{\Phi}_{cl} \stackrel{\triangle}{=} \mathbf{\Phi} + \mathbf{\Psi}\mathbf{K}_y\mathbf{C}$, the output energy is

$$I_y = \sum_{0}^{\infty} y^2(k+1) = \mathbf{c}^T(\sum_{k=0}^{\infty}\mathbf{\Phi}^k\boldsymbol{\psi}\boldsymbol{\psi}^T\mathbf{\Phi}^{k,T})\mathbf{c} \ =\ \mathbf{c}^T\mathbf{L}_c\mathbf{c}\ , \tag{808}$$

where $\mathbf{L}_c$ follows from Eq.(806). The changes in $\Delta\mathbf{L}_c$ in terms of the increments $\Delta\mathbf{K}_y$ result from

$$\mathrm{col}\Delta\mathbf{L}_c = \mathbf{E}_3[(\mathbf{E}_4\otimes\mathbf{\Psi})\mathrm{col}\Delta\mathbf{K}_y + (\mathbf{\Psi}\otimes\mathbf{E}_4)\mathrm{col}\Delta\mathbf{K}_y^T], \tag{809}$$

where $\mathbf{E}_4 \ \stackrel{\triangle}{=}\ \mathbf{\Phi}_{cl}\mathbf{L}_c\mathbf{C}^T$ and $\mathbf{E}_3 \stackrel{\triangle}{=} (\mathbf{I} - \mathbf{\Phi}_{cl}\otimes\mathbf{\Phi}_{cl})^{-1}$.

Stepping forward to multiple-input multiple-output systems (*Weinmann, A., 2006a*), denote single-input single-output partitions from input k to output i. Referring to Eq.(804), the index of performance for y_i is

$$I_{yi} = (\mathbf{c}_i^T\otimes\mathbf{c}_i^T)\mathrm{col}\,\mathbf{L}_{ck}\ , \tag{810}$$

where $\mathbf{c}_i$ belongs to y_i and $\mathbf{L}_{ck}$ follows from Eq.(806) with columns $\boldsymbol{\psi} = \boldsymbol{\psi}_k$. The index I_{yi} should decrease to reduce the interaction. (When a prefilter $\mathbf{V}$ is switched in front of the system, then the kth column of $\mathbf{\Psi V}$ is used.)

The MATLAB command `dlyap` is introduced. Then, $(\mathbf{\Psi}^T \otimes \mathbf{E}_4^T)\text{col}\mathbf{X}_3$ is rephrased as $\text{col}[\mathbf{E}_4^T \mathbf{X}_3 \mathbf{\Psi}] = \text{col}[\mathbf{C}\mathbf{L}_c \mathbf{\Phi}_{cl}^T \mathbf{X}_3 \mathbf{\Psi}]$, and the intersection gradient is

$$\frac{\partial I_{yi}}{\partial \mathbf{K}_y} = 2\mathbf{\Psi}^T \mathbf{X}_3 \mathbf{\Phi}_{cl} \mathbf{L}_{ck} \mathbf{C}^T \ , \tag{811}$$

$$\text{where} \qquad \mathbf{X}_3 = \texttt{dlyap}(\mathbf{\Phi}_{cl}^T,\ \mathbf{c}_i \mathbf{c}_i^T) \tag{812}$$

$$\mathbf{L}_{ck} = \texttt{dlyap}(\mathbf{\Phi}_{cl}, \boldsymbol{\psi}_k \boldsymbol{\psi}_k^T) \tag{813}$$

$$\boldsymbol{\psi}_k = k\text{-th column of} \mathbf{\Psi V} \ . \tag{814}$$

The influence of uncertainties in the plant is carried out in *Weinmann, A., 2007*.

29 State Variable Matrices in Optimal Control

The state variable *matrix* is defined as the dyadic product of the classic state vector, $\mathbf{X}(t) \stackrel{\triangle}{=} \mathbf{x}(t)\mathbf{x}^T(t)$. For a linear time-varying system $\mathbf{A}(t)$ and output feedback $\mathbf{u}(t) = \mathbf{K}(t)\mathbf{y}(t)$, one finds

$$\dot{\mathbf{X}}(t) = \dot{\mathbf{x}}\mathbf{x}^T + \mathbf{x}\dot{\mathbf{x}}^T = (\mathbf{A} + \mathbf{BKC})\mathbf{X} + \mathbf{X}(\mathbf{A} + \mathbf{BKC})^T \tag{815}$$

with the initial condition $\mathbf{X}(0) = \mathbf{x}_o \mathbf{x}_o^T$. Minimizing the performance criterion yields

$$I = \mathbf{x}^T(t_f)\mathbf{F}_f \mathbf{x}(t_f) + \int_{t_o}^{t_f} [\mathbf{x}^T(t)\mathbf{Q}\mathbf{x}(t) + \mathbf{u}^T(t)\mathbf{R}\mathbf{u}(t)]dt \tag{816}$$

$$I = \text{tr}\ \mathbf{F}_f \mathbf{X} + \int_{t_o}^{t_f} \text{tr}\ (\mathbf{Q} + \mathbf{C}^T \mathbf{K}^T \mathbf{RKC})\mathbf{X}\ dt \ . \tag{817}$$

Using the Hamiltonian function

$$H \stackrel{\triangle}{=} \text{tr}\ (\mathbf{Q} + \mathbf{C}^T \mathbf{K}^T \mathbf{RKC})\mathbf{X} + \text{tr}\ \dot{\mathbf{X}}\mathbf{\Lambda}^T \ , \tag{818}$$

the canonical equations are

$$\dot{\mathbf{X}}(t) = \frac{\partial H}{\partial \mathbf{\Lambda}(t)} \ . \tag{819}$$

$$\dot{\mathbf{\Lambda}}(t) = -\frac{\partial H}{\partial \mathbf{X}(t)} \ , \tag{820}$$

with the terminal condition (transversality condition)

$$\mathbf{\Lambda}(t_f) = \frac{\partial}{\partial \mathbf{X}(t_f)} \mathbf{F}_f[\mathbf{X}(t_f)] \ . \tag{821}$$

The matrix $\mathbf{\Lambda}$ denotes the costate matrix. In state vector form, the canonical equations are collected in the Hamilton matrix $\begin{pmatrix} \mathbf{A} & \mathbf{BR}^{-1}\mathbf{B}^T \\ \mathbf{Q} & -\mathbf{A}^T \end{pmatrix}$, see also Eq.(37),

$$\begin{pmatrix} \dot{\mathbf{x}}(t) \\ \dot{\boldsymbol{\lambda}}(t) \end{pmatrix} = \begin{pmatrix} \mathbf{A} & \mathbf{BR}^{-1}\mathbf{B}^T \\ \mathbf{Q} & -\mathbf{A}^T \end{pmatrix} \begin{pmatrix} \mathbf{x}(t) \\ \boldsymbol{\lambda}(t) \end{pmatrix} \qquad \mathbf{x}(0) = \mathbf{x}_o;\ \boldsymbol{\lambda}(t_f) = \mathbf{0}\ . \tag{822}$$

The matricial gradients then result as (*Bernussou, J., and Titli, A., 1982*)

$$\frac{\partial H}{\partial \mathbf{K}} = 2\ \mathbf{RKCXC}^T + 2\ \mathbf{B}^T\mathbf{\Lambda XC}^T\ , \tag{823}$$

$$\frac{\partial I(\mathbf{K})}{\partial \mathbf{K}} = \int_{t_o}^{t_f} \frac{\partial}{\partial \mathbf{K}}[f + \mathrm{tr}\ \mathbf{\Lambda}^T\mathbf{F}]\mathrm{dt}\ = \int_{t_o}^{t_f} \frac{\partial}{\partial \mathbf{K}} H(\mathbf{X}, \mathbf{\Lambda}, \mathbf{K})\mathrm{dt} \tag{824}$$

$$\frac{\partial I(\mathbf{K})}{\partial \mathbf{K}} = 2\int_{t_o}^{t_f} (\mathbf{RKC} + \mathbf{B}^T\mathbf{\Lambda})\mathbf{XC}^T\ \mathrm{dt}\ , \tag{825}$$

$$\dot{\mathbf{X}} = (\mathbf{A} + \mathbf{BKC})\mathbf{X} + \mathbf{X}(\mathbf{A} + \mathbf{BKC})^T, \qquad \mathbf{X}(0) = \mathbf{X}_o \tag{826}$$

$$\dot{\mathbf{\Lambda}} = -\mathbf{Q} - \mathbf{C}^T\mathbf{K}^T\mathbf{RKC} - (\mathbf{A} + \mathbf{BKC})^T\mathbf{\Lambda} - \mathbf{\Lambda}(\mathbf{A} + \mathbf{BKC}), \tag{827}$$

$$\mathbf{\Lambda}(t_f) = \frac{\partial}{\partial \mathbf{X}(t_f)} \mathbf{F}_f[\mathbf{X}(t_f)]\ . \tag{828}$$

In the time-invariant infinite horizon case $\mathbf{\Lambda}$ is constant. Defining

$$\mathbf{V} = \int_o^{\infty} \mathbf{X}(t)\ \mathrm{dt} \tag{829}$$

and integrating $\int_o^\infty \dot{\mathbf{X}}(t)\mathrm{dt}\ = \mathbf{X}(\infty) - \mathbf{X}(0) = \mathbf{0} - \mathbf{X}_o = -\mathbf{X}_o$ yields

$$\frac{\partial I(\mathbf{K})}{\partial \mathbf{K}} = 2(\mathbf{RKC} + \mathbf{B}^T\mathbf{\Lambda})\mathbf{VC}^T \tag{830}$$

$$(\mathbf{A} + \mathbf{BKC})\mathbf{V} + \mathbf{V}(\mathbf{A} + \mathbf{BKC})^T + \mathbf{X}_o = \mathbf{0} \tag{831}$$

$$\dot{\mathbf{\Lambda}} = \mathbf{0} = \mathbf{Q} + \mathbf{C}^T\mathbf{K}^T\mathbf{RKC} + (\mathbf{A} + \mathbf{BKC})^T\mathbf{\Lambda} + \mathbf{\Lambda}(\mathbf{A} + \mathbf{BKC}) = \mathbf{0}\ . \tag{832}$$

30 Minimizing Controller Matrix Norm vs Pole Assignment. Desensitization

An important problem is minimizing the norm of the state controller $\mathbf{K}$

$$\|\mathbf{K}\| \to \min_{\mathbf{K}} \tag{833}$$

$$\text{subject to}\ \ \lambda_i[\mathbf{A} + \mathbf{BK}] = \lambda_{di}\ \ \forall i = \{1, 2 \ldots n\}\ . \tag{834}$$

The condition of desired closed-loop eigenvalues λ_{di} can be stated as a vector-valued condition

$$\mathbf{f} \stackrel{\triangle}{=} \mathbf{vec}_i[\det(\lambda_{di}\mathbf{I}_n - \mathbf{A} - \mathbf{B}\mathbf{K})] = \mathbf{0}\ . \tag{835}$$

Starting from Eq.(835) and using Eq.(351), one has

$$2\mathbf{K} + (\mathbf{I}_m \otimes \boldsymbol{\mu}^T)\frac{\partial \mathbf{f}}{\partial \mathbf{K}} = \mathbf{0} \in \mathcal{R}^{m\times n}\ , \tag{836}$$

where

$$\frac{\partial \mathbf{f}}{\partial \mathbf{K}} = \begin{pmatrix} \frac{\partial \mathbf{f}}{\partial K_{11}} & \frac{\partial \mathbf{f}}{\partial K_{12}} & \cdots & \frac{\partial \mathbf{f}}{\partial K_{1n}} \\ \vdots & \ddots & & \vdots \\ \frac{\partial \mathbf{f}}{\partial K_{m1}} & \cdots & & \frac{\partial \mathbf{f}}{\partial K_{mn}} \end{pmatrix} \in \mathcal{R}^{nm\times n}\ . \tag{837}$$

Applying

$$\frac{\partial}{\partial \mathbf{K}}\det(s_i\mathbf{I}_n - \mathbf{A} - \mathbf{B}\mathbf{K}) = -2\mathbf{B}^T\mathbf{adj}^T(s_i\mathbf{I}_n - \mathbf{A} - \mathbf{B}\mathbf{K})\ , \tag{838}$$

a matrix is defined in what follows

$$\mathbf{P}_P \stackrel{\triangle}{=} \begin{pmatrix} -2\mathbf{B}^T\mathbf{adj}^T(\lambda_{d1}\mathbf{I}_n - \mathbf{A} - \mathbf{B}\mathbf{K}) \\ -2\mathbf{B}^T\mathbf{adj}^T(\lambda_{d2}\mathbf{I}_n - \mathbf{A} - \mathbf{B}\mathbf{K}) \\ \vdots \\ -2\mathbf{B}^T\mathbf{adj}^T(\lambda_{dn}\mathbf{I}_n - \mathbf{A} - \mathbf{B}\mathbf{K}) \end{pmatrix} \in \mathcal{R}^{nm\times n}\ . \tag{839}$$

Afterwards a fixed-valued permutation matrix $\mathbf{M}_P \in \mathcal{R}^{mn\times mn}$ is used to change the row order and to precisely achieve the order of Eq.(837), where $\mathbf{M}_P = \mathbf{U}_{k,l} = \mathbf{U}_{m,n}$, see Eq.(25).

Finally, the entire result is given by

$$\begin{aligned} 2\mathbf{K} + (\mathbf{I}_m \otimes \boldsymbol{\mu}^T)\ \mathbf{M}_P\ \mathbf{P}_P(\mathbf{K}) &= \mathbf{0} \quad \in \mathcal{R}^{m\times n} & (840)\\ \mathbf{f}(\mathbf{K}) &= 0 \quad \in \mathcal{R}^{n}\ . & (841) \end{aligned}$$

This is a nonlinear system of mn unknowns in the matrix $\mathbf{K}$ and n unknowns in the Lagrange multiplier $\boldsymbol{\mu}$.

Including desensitization, the problem is

$$||\mathbf{K}||_F + \alpha||\frac{\partial \boldsymbol{\lambda}}{\partial \mathbf{K}}||_F \to \min_{\mathbf{K}} \tag{842}$$

$$\text{subject to} \quad \lambda_i[\mathbf{A} + \mathbf{B}\mathbf{K}] = \lambda_{di} \quad \forall i = \{1, 2 \ldots n\}\ . \tag{843}$$

The gradient of λ_i with respect to $\mathbf{K}$ is

$$\frac{\partial \lambda_i[\mathbf{A}_{cl}]}{\partial \mathbf{K}} = (\mathbf{I}_m \otimes \mathbf{v}_i^{\triangleleft * T})\frac{\partial \mathbf{A}_{cl}}{\partial \mathbf{K}}(\mathbf{I}_n \otimes \mathbf{v}_i)\ , \tag{844}$$

where $\mathbf{v}_i$ is the right eigenvector of the matrix $\mathbf{A}_{cl} \triangleq \mathbf{A} + \mathbf{BK}$ and $\mathbf{v}_i^{\triangleleft * T}$ denotes the conjugate transpose of the left eigenvector.

Eq.(836) turns out as

$$2\,\mathbf{K} + \alpha \frac{\partial}{\partial \mathbf{K}} \| \frac{\partial \boldsymbol{\lambda}}{\partial \mathbf{K}} \|_F + (\mathbf{I}_m \otimes \boldsymbol{\mu}^T) \frac{\partial \mathbf{f}}{\partial \mathbf{K}} = \mathbf{0} \in \mathcal{R}^{m \times n} \ . \tag{845}$$

For examples see *Weinmann, A., 2003a.*

31 Complex Minimization Condition for Reduced Function Observer

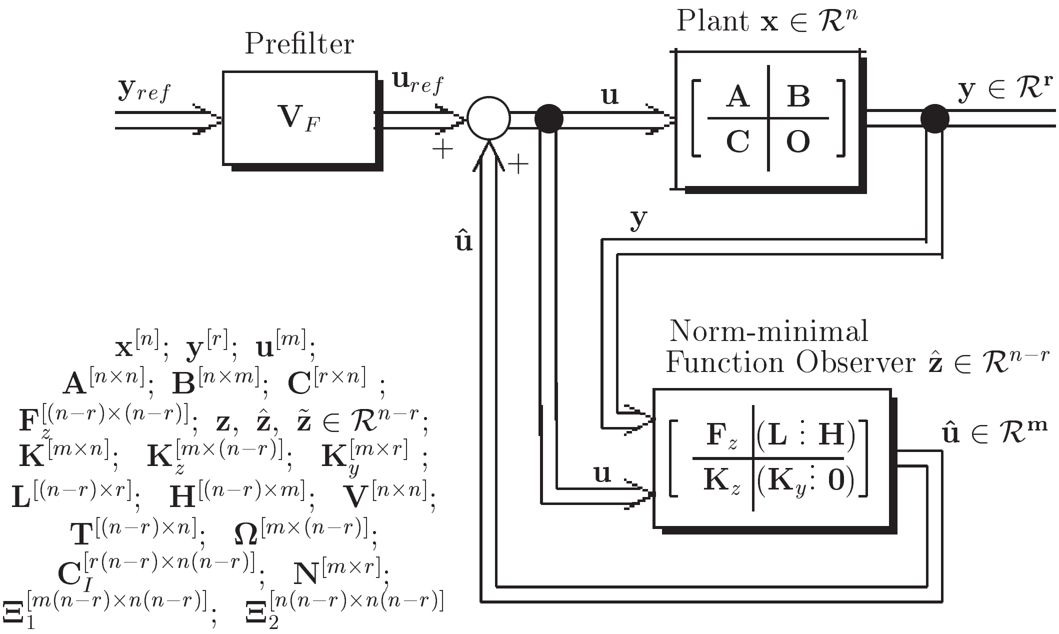

Figure 8: Control system with function observer

Consider $\mathbf{F}_z$ and $\mathbf{K}$ given. Following Fig. 8 and $\dot{\tilde{\mathbf{z}}} = \mathbf{F}_z \tilde{\mathbf{z}}$,

$$\hat{\mathbf{u}} = \mathbf{K}_z \hat{\mathbf{z}} + \mathbf{K}_y \mathbf{y} \ ; \quad \dot{\hat{\mathbf{z}}} = \mathbf{F}_z \hat{\mathbf{z}} + \mathbf{L}\mathbf{y} + \mathbf{H}\mathbf{u} \ ; \quad \mathbf{z} = \mathbf{T}\mathbf{x} = \hat{\mathbf{z}} + \tilde{\mathbf{z}} \ . \tag{846}$$

The observer obeys

$$\mathbf{F}_z \mathbf{T} - \mathbf{T}\mathbf{A} + \mathbf{L}\mathbf{C} = \mathbf{0} \tag{847}$$

$$\mathbf{H} = \mathbf{T}\mathbf{B} \tag{848}$$

$$(\mathbf{K}_z \vdots \mathbf{K}_y) = \mathbf{K} \begin{pmatrix} \mathbf{T} \\ \mathbf{C} \end{pmatrix}^{-1} \ . \tag{849}$$

Find $\mathbf{L}$ in Eq.(847), in order to minimize the norm sum

$$f_N \triangleq \|\mathbf{L}\|_F^2 + \|\mathbf{H}\|_F^2 + \|\mathbf{K}_z\|_F^2 + \|\mathbf{K}_y\|_F^2 = \|\mathbf{L}\|_F^2 + \|\mathbf{H}\|_F^2 + \|\mathbf{K}_z \vdots \mathbf{K}_y\|_F^2 \triangleq f_1 + f_2 + f_3 \quad (850)$$

Denoting $\mathbf{L}$ the unknown variable matrix or an initial assumption, $\mathbf{T}$ is

$$\mathbf{T} \overset{(847)}{=} \mathrm{lyap}(\mathbf{F}_z,\ -\mathbf{A},\ \mathbf{LC}) \quad \in \mathcal{R}^{(n-r)\times n}\ ; \quad \text{or} \quad (851)$$

$$\mathrm{col}\mathbf{T} \overset{(847)}{=} (\mathbf{A}^T \otimes \mathbf{I}_{n-r} - \mathbf{I}_n \otimes \mathbf{F}_z)^{-1} \mathrm{col}(\mathbf{LC}) = \boldsymbol{\Xi}_2 \mathbf{C}_I^T \mathrm{col}\mathbf{L} = \boldsymbol{\Xi}_2 (\mathbf{C}^T \otimes \mathbf{I}_{n-r}) \mathrm{col}\,\mathbf{L}, \quad (852)$$

$$\text{where} \quad \mathbf{C}_I \triangleq \mathbf{C} \otimes \mathbf{I}_{n-r} \in \mathcal{R}^{r(n-r)\times n(n-r)}\ ; \quad (853)$$

$$\boldsymbol{\Xi}_2 \triangleq (\mathbf{A}^T \otimes \mathbf{I}_{n-r} - \mathbf{I}_n \otimes \mathbf{F}_z)^{-1} \quad \in \mathcal{R}^{n(n-r)\times n(n-r)}\ ; \quad (854)$$

$$\boldsymbol{\Xi}_1 \triangleq (\mathbf{B}^T \otimes \mathbf{I}_{n-r}) \boldsymbol{\Xi}_2 \quad \in \mathcal{R}^{m(n-r)\times n(n-r)}\ ; \quad (855)$$

$$\mathbf{V} \triangleq \begin{pmatrix} \mathbf{T} \\ \mathbf{C} \end{pmatrix}^{-1} = \begin{bmatrix} \mathrm{loc}_{n-r}(\boldsymbol{\Xi}_2 \mathbf{C}_I^T\ \mathrm{col}\mathbf{L}) \\ \mathbf{C} \end{bmatrix}^{-1} = \mathbf{V}(\mathbf{L}) \quad \in \mathcal{R}^{n\times n}\ . \quad (856)$$

Partitioning the matrix $\mathbf{KV}$ including the dependence on the unknown $\mathbf{L}$, one has

$$(\boldsymbol{\Omega} \vdots \mathbf{N}) \triangleq \mathbf{KV} \ \in \mathcal{R}^{m\times n}; \qquad \left({}_m \overset{n-r}{\boldsymbol{\Omega}} \ \vdots {}_m \overset{r}{\mathbf{N}} \right) = \left({}_m \overset{n}{\mathbf{K}} \right) \times \left({}_n \overset{n}{\mathbf{V}} \right) = \left({}_m \overset{n}{\mathbf{KV}} \right) \quad (857)$$

$$\boldsymbol{\Omega} = \mathbf{K} \begin{pmatrix} \mathbf{T} \\ \mathbf{C} \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{I}_{n-r} \\ \mathbf{0}_{r,n-r} \end{pmatrix} = \mathbf{KV} \begin{pmatrix} \mathbf{I}_{n-r} \\ \mathbf{0}_{r,n-r} \end{pmatrix} \in \mathcal{R}^{m\times(n-r)} \quad (858)$$

$$\mathbf{N} = \mathbf{KV} \begin{pmatrix} \mathbf{0}_{n-r,r} \\ \mathbf{I}_r \end{pmatrix} \in \mathcal{R}^{m\times r} \quad (859)$$

Now all the interdependencis $\mathbf{U}_3(\mathbf{L}),\ \mathbf{T}(\mathbf{L}),\ \mathbf{V}(\mathbf{L}), \mathbf{H}(\mathbf{L})$ are available.

$$f_1 \triangleq \|\mathbf{L}\|_F^2 \leadsto \frac{\partial f_1}{\partial \mathrm{col}\mathbf{L}} = 2\mathrm{col}\mathbf{L}\ . \quad (860)$$

$$f_2 \triangleq \|\mathbf{H}\|_F^2 = \|\mathrm{col}\mathbf{H}\|_F^2 \overset{(848)}{=} \|(\mathbf{B}^T \otimes \mathbf{I}_{n-r})\mathrm{col}\ \mathbf{T}\|_F^2 = \|\boldsymbol{\Xi}_1 \mathbf{C}_I^T \mathrm{col}\mathbf{L}\|_F^2 \quad (861)$$

$$\frac{\partial f_2}{\partial \mathrm{col}\ \mathbf{L}} \overset{(219)}{=} 2(\mathbf{C} \otimes \mathbf{I}_{n-r}) \boldsymbol{\Xi}_1^T \boldsymbol{\Xi}_1 (\mathbf{C}^T \otimes \mathbf{I}_{n-r}) \mathrm{col}\ \mathbf{L}\ . \quad (862)$$

For f_3 we find

$$
\begin{aligned}
f_3 &\stackrel{\triangle}{=} \|\mathbf{K}\begin{pmatrix}\mathbf{T}\\ \mathbf{C}\end{pmatrix}^{-1}\|_F^2 && (863)\\
\Delta f_3 &= \|\mathbf{K}\begin{pmatrix}\mathbf{T}+\Delta\mathbf{T}\\ \mathbf{C}\end{pmatrix}^{-1}\|_F^2 - \|\mathbf{K}\begin{pmatrix}\mathbf{T}\\ \mathbf{C}\end{pmatrix}^{-1}\|_F^2 && (864)\\
&= \|\mathbf{K}\left[\begin{pmatrix}\mathbf{T}\\ \mathbf{C}\end{pmatrix}+\begin{pmatrix}\Delta\mathbf{T}\\ \mathbf{0}\end{pmatrix}\right]^{-1}\|_F^2 - \|\mathbf{K}\begin{pmatrix}\mathbf{T}\\ \mathbf{C}\end{pmatrix}^{-1}\|_F^2 && (865)\\
&= \|\mathbf{K}\left[\begin{pmatrix}\mathbf{T}\\ \mathbf{C}\end{pmatrix}^{-1}-\begin{pmatrix}\mathbf{T}\\ \mathbf{C}\end{pmatrix}^{-1}\begin{pmatrix}\Delta\mathbf{T}\\ \mathbf{0}\end{pmatrix}\begin{pmatrix}\mathbf{T}\\ \mathbf{C}\end{pmatrix}^{-1}\right]\|_F^2 - \|\mathbf{K}\begin{pmatrix}\mathbf{T}\\ \mathbf{C}\end{pmatrix}^{-1}\|_F^2 && (866)\\
&= \|\mathbf{KV}-\mathbf{KV}\begin{pmatrix}\Delta\mathbf{T}\\ \mathbf{0}\end{pmatrix}\mathbf{V}\|_F - \|\mathbf{KV}\|_F^2 && (867)\\
&= \mathrm{tr}\left[\mathbf{V}^T\mathbf{K}^T-\mathbf{V}^T\begin{pmatrix}\Delta\mathbf{T}\\ \mathbf{0}\end{pmatrix}^T\mathbf{V}^T\mathbf{K}^T\right]\left[\mathbf{KV}-\mathbf{KV}\begin{pmatrix}\Delta\mathbf{T}\\ \mathbf{0}\end{pmatrix}^T\mathbf{V}\right] - \|\mathbf{KV}\|_F && (868)\\
&= -2\mathrm{tr}\left[\mathbf{V}^T\mathbf{K}^T\mathbf{KVV}^T\begin{pmatrix}\Delta\mathbf{T}\\ \mathbf{0}\end{pmatrix}^T\right] = -2\mathrm{tr}\left[\mathbf{VV}^T\mathbf{K}^T\boldsymbol{\Omega}\Delta\mathbf{T}\right] && (869)\\
&\stackrel{(49)}{=} -2(\mathrm{col}\ \mathbf{I}_n)^T\ \mathrm{col}(\mathbf{VV}^T\mathbf{K}^T\boldsymbol{\Omega}\Delta\mathbf{T}) \stackrel{(51),(454)}{=} -2(\mathrm{col}\ \mathbf{I}_n)^T\ \mathrm{col}[\mathbf{I}_n\otimes(\mathbf{VV}^T\mathbf{K}^T\boldsymbol{\Omega})]\mathrm{col}\Delta\mathbf{T}\\
&\stackrel{(51)}{=} -2(\mathrm{col}\ \mathbf{I}_n)^T[\mathbf{I}_n\otimes(\mathbf{VV}^T\mathbf{K}^T\boldsymbol{\Omega})]\boldsymbol{\Xi}_2\mathbf{C}_I^T\mathrm{col}\Delta\mathbf{L} && (870)\\
\frac{\partial f_3}{\partial \mathrm{col}\ \mathbf{L}} &\stackrel{(450)}{=} -2\mathbf{C}_I\boldsymbol{\Xi}_2^T\left[\mathbf{I}_n\otimes(\boldsymbol{\Omega}^T\mathbf{KVV}^T)\right]\mathrm{col}\ \mathbf{I}_n\ . && (871)
\end{aligned}
$$

Numerical check: For some specific setup and $\mathbf{L} = (2\ \ 4)$ one gets $f_3 \stackrel{(863)}{=} 1.898362$ and $\frac{\partial f_3}{\partial \mathrm{col}\ \mathbf{L}} \stackrel{(871)}{=} (0.388969 \quad -0.195592)^T$. For $\mathbf{L} = (2\ \ 4.01)$ the result is $f_3 = 1.896412$. The difference $\Delta f_3 = -0.001950$ divided by the increment 0.01 matches the second component of the gradient.

Adding the three terms f_i and equating with zero yields the necessary condition for the over-all optimum result. The prefilter $\mathbf{V}_F = -[\mathbf{C}(\mathbf{A} + \mathbf{BK}_z\mathbf{T} + \mathbf{BK}_y\mathbf{C})^{-1}\mathbf{B}]^{-1} \in \mathcal{R}^{m\times r}$ is independent of $\mathbf{L}$ because $\mathbf{K}_z\mathbf{T} + \mathbf{K}_y\mathbf{C} = (\mathbf{K}_z \vdots \mathbf{K}_y)\begin{pmatrix}\mathbf{T}\\ \mathbf{C}\end{pmatrix} = \mathbf{K}$. Arbitrary weighting matrices could be included everywhere (*Weinmann, A., 2003b*).

32 Matrix Element Interdependencies

For specific matrix interdependencies, e.g., matrix $\mathbf{M}$ symmetric or skew-symmetric (*Geering, H.P., 1976*) there are different correspondences, different from Eq.(312),

$$\frac{\partial \mathbf{M}}{\partial M_{ik}} = \mathbf{E}_{ik} + \mathbf{E}_{ki} - \mathbf{E}_{ii}\delta_{ik}\ ,\quad \mathbf{M}\ \text{symmetric} \qquad (872)$$

$$\frac{\partial \mathbf{M}}{\partial M_{ik}} = \mathbf{E}_{ik} - \mathbf{E}_{ki} \ , \quad \mathbf{M} \text{ skew symmetric} \ , \tag{873}$$

where $\mathbf{E}_{ik} = \mathbf{e}_i \mathbf{e}_k^T$ is the Kronecker matrix. E.g., Eq.(188) is not true if the matrix is symmetric. For further examples, see *Geering, H.P., 1976; Weinmann, A., 2001b*. For symmetric matrices, there is the general relation

$$\frac{\partial \det \mathbf{M}}{\partial \mathbf{M}} = (2\mathbf{M}^{-1} - \mathbf{diag}\ \{\mathbf{M}^{-1}\}) \det \mathbf{M} \ , \tag{874}$$

where $\mathbf{diag}\ \{\mathbf{M}^{-1}\}$ is a diagonal matrix made of main-diagonal elements of $\mathbf{M}^{-1}$.

For $f(\mathbf{M}) \triangleq \operatorname{tr}[\mathbf{LM}]$

$$\frac{\partial f(\mathbf{M})}{\partial \mathbf{M}} = \begin{cases} (\mathbf{L}^T + \mathbf{L} - \mathbf{diag}\ \mathbf{L}) & \text{for } \mathbf{M} \text{ symmetric} \\ (\mathbf{L}^T - \mathbf{L}) & \text{for } \mathbf{M} \text{ skew symmetric} \ . \end{cases} \tag{875}$$

For $\mathbf{M}$ symmetric the linear relation $\mathbf{L}_1\mathbf{M}\mathbf{L}_2$ yields (*Brewer, J.W., 1977*)

$$\frac{\partial}{\partial \mathbf{M}} \operatorname{tr}[\mathbf{L}_1\mathbf{M}\mathbf{L}_2] = \mathbf{L}_1^T\mathbf{L}_2^T + \mathbf{L}_2\mathbf{L}_1 - \mathbf{diag}\{\mathbf{L}_2\mathbf{L}_1\} \tag{876}$$

and for the relation of second order

$$\frac{\partial}{\partial \mathbf{M}} \operatorname{tr}[\mathbf{L}_1\mathbf{M}\mathbf{L}_2\mathbf{M}] = (\mathbf{L}_2\mathbf{M}\mathbf{L}_1)^T + \mathbf{L}_1\mathbf{M}\mathbf{L}_2 - \mathbf{diag}\{\mathbf{L}_2\mathbf{M}\mathbf{L}_1 + \mathbf{L}_1\mathbf{M}\mathbf{L}_2\} \ . \tag{877}$$

The reason for the strong differences in the differential quotients for occasionally existing matrix interdependencies is the fact that the partial derivative conditions are not fulfilled.

33 Gradients for Computer-Assisted Conditions for Positive Realness

The vector $\mathbf{s}$ is defined a vector of powers of the Laplace operator s. For $s \triangleq j\omega$ and $\omega = 1$ we get the matrices Ξ_v and Ξ_y, which are needed for determining the signs of polynomials partly replaced by their complex conjugate

$$\mathbf{s} \triangleq (1 \quad s \quad s^2 \quad \ldots)^T \ , \quad \Xi_v \triangleq \mathbf{s}|_{s=j}\mathbf{s}|_{s=j}^H \ , \quad \Xi_y \triangleq \mathbf{s}|_{s=j}\mathbf{s}|_{s=j}^T \ , \tag{878}$$

where ^{H} means Hermite, i.e., transposition and the complex conjugate, and * is the complex conjugate without transposition.

Multiplying two polynomials $a(s)$ and $b^\star(s)$ given by their polynomial coefficients $\mathbf{a}$ and $\mathbf{b}$, respectively, requires the Hadamard product (dot product) and summing up elements in some secondary diagonal referring to Eqs.(881) and (882) $(\mathbf{ab}^T)\ .*\ \Xi_v$.

The dot product assigns the appropriate sign and imaginary unit j. The conditions for positive realness in the case of controller $K(s)$ and plant $G(s)$, switched in series, result from

$$F_o(s) = \frac{Z_K(s)Z_G(s)}{P_K(s)P_G(s)} = \frac{(\sum_1^n z_{Ki}s^i)(\sum_1^n z_{Gi}s^i)}{(\sum_1^n p_{Ki}s^i)(\sum_1^n p_{Gi}s^i)} = \frac{(\mathbf{z}_K^T\mathbf{s})(\mathbf{z}_G^T\mathbf{s})(\mathbf{p}_K^T\mathbf{s}^*)(\mathbf{p}_G^T\mathbf{s}^*)}{(\mathbf{p}_K^T\mathbf{s})(\mathbf{p}_G^T\mathbf{s})(\mathbf{p}_K^T\mathbf{s}^*)(\mathbf{p}_G^T\mathbf{s}^*)} \quad (879)$$

$$= \frac{\mathbf{s}^T(\mathbf{z}_K\mathbf{p}_K^T)\mathbf{S}(\mathbf{z}_G\mathbf{p}_G^T)\mathbf{s}^*}{\mathbf{s}^T(\mathbf{p}_K\mathbf{p}_K^T)\mathbf{S}(\mathbf{p}_G\mathbf{p}_G^T)\mathbf{s}^*} = \frac{\mathbf{s}^T\bar{\mathbf{K}}\mathbf{S}\bar{\mathbf{G}}\mathbf{s}^*}{\mathbf{s}^T\mathbf{P}_K\mathbf{S}\mathbf{P}_G\mathbf{s}^*} , \quad (880)$$

where the following abbreviations have been used $\bar{\mathbf{K}} \triangleq \mathbf{z}_K\mathbf{p}_K^T$, $\bar{\mathbf{G}} \triangleq \mathbf{z}_G\mathbf{p}_G^T$, $\mathbf{P}_K \triangleq \mathbf{p}_K\mathbf{p}_K^T$, $\mathbf{P}_G \triangleq \mathbf{p}_G\mathbf{p}_G^T$, $\mathbf{s} \triangleq (1 \quad s \quad s^2 \quad \ldots)^T$, $\mathbf{s}|_{s=j\omega} = (1 \quad j\omega \quad -\omega^2 \quad \ldots)^T$, $\mathbf{s}^*|_{s=j\omega} = (1 \quad -j\omega \quad -\omega^2 \quad \ldots)^T$, $\mathbf{S} \triangleq \mathbf{s}^*|_{s=j\omega}\mathbf{s}^T|_{s=j\omega}$. The scalar expression $\mathbf{s}^T\bar{\mathbf{K}}\mathbf{S}\bar{\mathbf{G}}\mathbf{s}^*|_{s=j\omega} \triangleq A$ is a polynomial of degree $4(n-1)$. The positive realness conditions of the polynomial coefficients can be separated by the following MATLAB algorithm

```
A(1)=A
for i=1:2*n
    A(i+1)=simplify((A(i)-subs(A(i),om,0))/om); end
for i=1:2:2*n
    B(i)=subs(A(i),om,0); end
```

$B(i)$ for $i = 1, 3, 4 \ldots$ are the coefficients of the numerator polynomial of $G(j\omega)$ after having made the denominator real-valued.

Multiplying two general polynomials[2] $a(s)b^*(s)$ (with equal degree up to $n-1$), the coefficients of the powers of ω of the product are given in what follows. The coefficient of ω^{i-1} is

$$\mathrm{tr}[\mathbf{I}_{i\times n}\{(\mathbf{a}\mathbf{b}^T)\ .*\ \Xi_v\}\mathbf{I}_{i\times n}^T\mathbf{I}_{Ri}], \qquad 1 \le i \le n \quad (881)$$

$$\mathrm{tr}[\mathbf{I}_{(2n-i)\times n}\mathbf{I}_{Rn}\{(\mathbf{a}\mathbf{b}^T)\ .*\ \Xi_v\}\mathbf{I}_{Rn}^T\mathbf{I}_{(2n-i)\times n}^T\mathbf{I}_{R(2n-i)}], \qquad n+1 \le i \le 2n-1 , \quad (882)$$

where Ξ_v and $\mathbf{I}_{i\times n}$ are defined by Eqs.(878) and (883), respectively, and

$$\mathbf{I}_{i\times n} \triangleq \text{e.g.,} \mathbf{I}_{2\times 4} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} . \quad (883)$$

Defining a controller $\mathbf{k}$ with two elements, only, e.g.,

$$\mathbf{k} \triangleq (c \quad e)^T \quad (884)$$

$$\mathbf{h} \triangleq \mathrm{vec}_i[B(i)] , \quad (885)$$

[2] equal to the convolution operation if available in symbolic operation

the affordable changes Δc, Δe of the controller $\mathbf{k}$ result from the gradient, the weighting vector $\mathbf{v} \stackrel{\triangle}{=} 0.5\mathrm{vec}_i\{[1 - B(i)/|B(i)|]w_i\}$, applied to each gradient element, and the stepsize Δr (*Weinmann, A., 2011*)

$$\Delta\mathbf{k} = \begin{pmatrix} \Delta c \\ \Delta e \end{pmatrix} = (\frac{\partial \mathbf{h}^T}{\partial \mathbf{k}})\ \mathbf{v}\ \Delta r = \begin{pmatrix} \frac{\partial h_1}{\partial c} & \frac{\partial h_1}{\partial e} \\ \frac{\partial h_2}{\partial c} & \frac{\partial h_2}{\partial e} \\ \frac{\partial h_3}{\partial c} & \frac{\partial h_3}{\partial e} \end{pmatrix}^T \mathbf{v}\ \Delta r\ . \tag{886}$$

34 Concluding Remarks

Gradients and matricial gradients of several scalar functions of vectors and matrices have been presented as a collection. Applications in automatic control have been preferred.

Gradients are very useful for designing controllers in a stepwise version; optimizing a given criterion (or a combination of criteria) while observing some other conditions, e.g., for the actuating signal amplitude, interaction energy etc.

In the presence of uncertainties (e.g., *Weinmann, A., 2007e*), their influence on the control system performance has to be included into the design process. In these cases, the given performance matrix manipulations are also very helpful.

35 References

Athans, M., 1968, The matrix minimum principle, *Information and Control* **11**, pp. 592-606

Barmish, B.R., 1994, New Tools for Robustness of Linear Systems (NewYork: Macmillan Publ.Comp.)

Bellman, R., 1960, Introduction to Matrix Analysis (New York, McGraw-Hill)

Bernussou, J., and Titli, A., 1982, Interconnected Dynamical Systems: Stability, Decomposition and Decentralization (North-Holland, Amsterdam New York Oxford)

Brewer, J.W., 1977, The gradient with respect to a symmetric matrix, *IEEE-Trans.* **AC-22**, pp. 265-267

Brewer, J.W., 1977a, The derivative of the exponential matrix with respect to a matrix, *IEEE-Trans.* **AC-22**, pp. 656-657

Brewer, J.W., 1978, Kronecker products and matrix calculus in system theory, *IEEE-Trans.* **CAS-25**, pp. 772-781

Brewer, J.W., 1978a, Matrix calculus and the sensitivity analysis of linear dynamic systems, *IEEE-Trans.* **AC-23**, pp. 748-751

Cheok, K.C., Loh, N.K., Zohdy, M.A., 1986, Generalized optimal feedback controllers for discrete-time linear systems, *Journal of Dynamic Systems, Measurement, and Control* **108**, pp. 91-95

Doyle, J.C., Glover, K., Zhou, K.,1996, Robust and Optimal Control (Englewood Cliffs, Prentice Hall)

Feliachi, A., 1986, Decentralized stabilization of interconnected systems, *Int.J.Control* **44**, pp. 1499-1505

Franklin, G.F., Powell, J.D., Emami-Naeini, A., 2002, Feedback Control of Dynamic Systems, 4th ed., (Upper Saddle River, N.J., Prentice Hall)

Franklin, J.N., 1968, Matrix Theory (Prentice Hall, Englewood Cliffs)

Freudenberg, J.S., Looze, D.P., and Cruz, J.B., 1982, Robustness analysis using singular value sensitivities, *Int.J.Control* **35**, pp. 95-116

Friedland, S., 1983 Simultaneous similarity of matrices, *Adv.Math.* **50**, pp. 189-265

Geering, H.P., 1976, On calculating gradient matrices, *IEEE Trans.* **AC-21**, pp. 615-616

Glover, K., 1984, All optimal Hankel-norm approximations of linear multivariable systems and their L^∞-error bounds, *Int.J.Control* **39**, pp. 1115-1193

Hoerl, A.E., 1962, Application of ridge analysis to regression problems, *Chem. Eng. Progr.* **58**, pp. 54-59

Hoerl, A.E., and Kennard, R.W., 1970, Ridge regression: biased estimation for nonorthogonal problems, *Technometrics* **12**, No.1, pp. 55-67

Laub, A.J., 2005, Matrix Analysis for Scientists and Engineers (SIAM, Society for Industrial and Applied Mathematics)

Levine, W.S., and Athans, M., 1970, On the determination of the optimal constant output feedback gains for linear multivariable systems, *IEEE-Trans.* **AC-15**, pp. 44-48

Ludyk, G., 1995, Theoretische Regelungstechnik 1 und 2. Berlin: Springer-Verlag

Papageoriou, M., 1991, Optimierung (Muenchen Oldenbourg)

Rothfuß, R., Rudolph, J., and Zeitz, M., 1997, Flachheit: Ein neuer Zugang zur Steuerung und Regelung nichtlinearer Systeme, *Automatisierungstechnik* **45**, pp. 517-525

Sevaston, G.E., and Longman, R.W., 1988, Systematic design of robust control systems using small gain stability concepts, *Int.J.Systems Sci.* **19**, pp. 439-451

Sira-Ramirez, H., and Agrawal, S.K., 2004, Differentially flat systems, New York, Basel: M. Dekker

Vetter, W.J., 1970, Derivative operations on matrices, *IEEE-Trans.* **AC-15**, pp.241-244 and **AC-16**, p.113

Vetter, W.J., 1971, An extension to gradient matrices, *IEEE-Trans.* **SMC-1**, pp. 184-186

Vetter, W.J., 1973, Matrix calculus operations and Taylor expansions, *SIAM Review* **15**, pp.352-369

Vetter, W.J., Balchen, J.G., and Pohner, F., 1972, Near-optimal control of parameter-dependent linear plants, *Int.J.Control* **15**, pp. 683-691

Weinmann, A., 1991, Uncertain Models and Robust Control (Vienna and NewYork: Springer)

Weinmann, A., 1995, Direct and incremental calculation of the restricted total least squares problem for improved process identification, *Int. J. Automation Austria* **3**, No. 1, pp. 1-14

Weinmann, A., 1995a Parameter-Weighted Performance in Total Least Squares Iterative Solutions, *Mathematical Modelling of Systems (Swets and Zeitlinger, The Netherlands)* **1**, No.

3, pp. 167-183

Weinmann, A., 2001, Matrix assignment subject to Lyapunov and closed-loop norm conditions, *2001 American Control Conference, Arlington, Virginia*, MP12-5

Weinmann, A., 2001a, Control system design based on holistic eigenvalue allocation, e&i *Elektrotechnik und Informationstechnik* **118**, H. 4, S. 167-173

Weinmann, A., 2001b, Gradients of norms, traces and determinants for automatic control applications, *Int. J. Automation Austria* **9**, H.1/2, pp. 36-50

Weinmann, A., 2003, Control systems design driven by worst-initial energy sensitivity, *Cybernetics and Systems: An International Journal* **34**, No.8, pp. 591-615

Weinmann, A., 2003a, Multi-Lagrange-Multiplier Problems: Desensitized pole assignment with minimial controller matrix norm for networked control systems and constraints in high-order mechatronics, *Mathematical and Computer Modelling of Dynamical Systems* **9**, pp. 345-362, Swets & Zeitlinger Publishers

Weinmann, A., 2003b, Beobachter minimaler Norm, *Tagungsband des 13. Steirischen Seminars über Regelungstechnik und Prozeßautomatisierung in Retzhof*, S. 1-20

Weinmann, A., 2004, Optimal assignment of multiple control system eigenvalues with minimum controller norm, *e & i (Elektrotechnik und Informationstechnik)* **121**, pp.37-42

Weinmann, A., 2004a, Robust worst-case initialization for norm-limited output state controllers in embedded systems, *e & i (Elektrotechnik und Informationstechnik)* **121**, pp. 282-285

Weinmann, A., 2005, Holistic criteria for incrementally designed controller, *e & i (Elektrotechnik und Informationstechnik)* **122**, No. 10, pp. 361-368

Weinmann, A., 2005a, A dialog-oriented and gradient-based stability margin in uncertain systems, *Cybernetics and Systems: An International Journal* **36**, Number 7, pp.641-666

Weinmann, A., 2005b, Gradients for decoupling continuous-time control systems using the controllability Gramian, *Tagungsband des 14. Steirischen Seminars über Regelungstechnik und Prozeßautomatisierung in Retzhof, 21. Sept. 2005*, S. 99-116

Weinmann, A., 2005c, Stability margin gradients and robustification of control systems based on transfer matrices, *Int. J. Automation Austria* **13**, pp. 37-49

Weinmann, A., 2006, Stability margin gradients using the polynomial matrix representation, *e & i (Elektrotechnik und Informationstechnik)* **123**, H.5, S. 197-201

Weinmann, A., 2006a, Attenuating the interaction via discrete-time controllability gramian, *Int. J. Automation Austria* **14**, pp. 13-30

Weinmann, A., 2006b, Robustness gradients for discrete-time control systems, *Cybernetics and Systems: An International Journal* **37**, Number 4, pp. 283-292

Weinmann, A., 2006c, Stability margin and spherical uncertainty, *Cybernetics and Systems: An International Journal* **37**, Number 7, pp. 685-705

Weinmann, A., 2006d, Test- und Prüfungsaufgaben Regelungstechnik. 457 durchgerechnete Beispiele mit Lösungen. 2. Auflage (Wien, New York: Springer)

Weinmann, A., 2007, Robustness and dynamic interaction, *Cybernetics and Systems: An International Journal* **38**, No 3, pp. 229-248

Weinmann, A., 2007a, Duality-based robust polynomial matrix controller, *e & i (Elektrotechnik und Informationstechnik)* **124**, H.7/8, pp. 273-276

Weinmann, A., 2007b, The encirclements of the hodograph of perturbed high-order dynamic systems, *Int. J. Automation Austria* **15**, pp. 117-129

Weinmann, A., 2007c, Robust stability radius gradient. A duality approach for high-order systems, *Int. J. Automation Austria* **15**, pp. 86-97

Weinmann, A., 2007d, Eliminating double solutions in Lyapunov equations, *Int. J. Automation Austria* **15**, pp. 130-134

Weinmann, A., 2007e, Performance gradients under spherical uncertainties, *Tagungsband des 15. Steirischen Seminars über Regelungstechnik und Prozeßautomatisierung in Retzhof, 19. Sept. 2007*, S. 141-163

Weinmann, A., 2008, Robust improvement of the phase and gain margin and crossover frequency, *e & i (Elektrotechnik und Informationstechnik)* **125**, H. 1/2 S. 17-24

Weinmann, A., 2008a, Robust limit cycle gradients, *e & i (Elektrotechnik und Informationstechnik)* **125**, H. 1/2 S.12-16

Weinmann, A., 2008c, Performance gradients in automatic control analysis, a catalogue of correspondences. *Int. J. Automation Austria* **16**, pp. 1-59. See http://www.acin.tuwien.ac.at/Publikationen/IJAA

Weinmann, A., 2008d, Gradients of Lyapunov subdeterminants preparing state controller design, *Int. J. Automation Austria* **16**, pp. 126-136

Weinmann, A., 2011, Positive realness and positive definiteness: gradients and robustness. *Int. J. Automation Austria* **19**, pp. 71-101

Yahagi, T., 1973, On the design of optimal output feedback control systems, *Int.J.Control* **18**, pp. 839-848

Yahagi, T., 1977, Optimal output feedback control with reduced performance index sensitivity, *Int.J.Control* **25**, pp. 769-783

Zeitz, M., 2009, Flachheit – Eine nützliche Methodik auch für lineare SISO-Systeme, *16. Steirisches Seminar über Regelungstechnik und Prozessautomatisierung*, pp. 3-15

26. Österreichischer Automatisierungstag,

Fr. 21.10.2011, Innsbruck

Der 26. Österreichische Automatisierungstag fand am Freitag den 21.10.2011 in Innsbruck statt und wurde vom Institut für Automatisierungs- und Regelungstechnik der UMIT (Univ.Prof. Dr. Michael Hofbaur), Hall in Tirol hervorragend organisiert.

Der verstärkte Einsatz von Automatisierung, insbesondere von Robotertechnologie, für Fertigungsprozesse ist mittlerweile auch für KMUs ein wesentlicher Faktor zur Erhaltung der Wettbewerbsfähigkeit in einem kontinuierlich global ausgerichteten Markt. Dieser 26. Österreichische Automatisierungstag hatte sich daher den Themenkomplex Automatisierung in KMUs und Robotik als Schwerpunkt gesetzt und beleuchtete relevante und aktuelle automatisierungstechnische Fragestellungen durch Vorträge aus Industrie und Forschung.

Als Einstieg in den Themenbereich Automatisierungstechnik und KMUs wurden von Herrn Ing. Rainer Haag (EMATRIC) technische und organisatorische Problemstellungen im Vortrag „Anlagenautomatisierung: Probleme und Vorgehensweisen für KMUs“ eingehend behandelt. Im Vortrag „WEB-basierendes Service Diagnose Management“ wurde von Herrn Ing. Hubert Brunner (Bernecker+Rainer) eine für den Betrieb wesentliche Funktionalität moderner Automatisierungsanalgen beschrieben.

Das Themengebiet Robotik prägte den zweiten Teil der Fachveranstaltung. Im Forschungs-Beitrag wurde von Herrn DI Peter Staufer (Inst. f. Robotik der Johannes Keppler Universität Linz) über die „Modellierung und Regelung von strukturelastischen Robotern“ berichtet. Im Anschluss zu diesem aufschlussreichen Beitrag über innovative Robotik-Forschung in Österreich wurde von Herrn Christian Peer (KUKA) das Roboter Steuerungssystem KR C4 vorgestellt und damit ein Einblick in die aktuellen Themenschwerpunkte für Industrieroboter-Steuerungen gewährt. Abschließend zum Themenkomplex Robotik und als Brücke zum Schwerpunkt KMUs des Automatisierungstages wurde im Anwendungsbeitrag „Optische Positions- und Lagebestimmung für Pick & Place Operationen“ von Herrn DI Alfred Moser (WESTCAM) eine neuartige, Roboter-basierte Anwendung eingehend dargestellt.

Im Zuge des Automatisierungstages wurden auch die diesjährigen Fred Margulies Preise des IFAC-Beirates Österreich und der ÖGART wurden an 6 Preisträger verliehen.

Der Automatisierungstag wurde in Zusammenarbeit mit dem Cluster Mechatronik der Standortagentur Tirol durchgeführt. Neben dem bewährten Nachmittagsprogramm wurde erstmals ein zusätzliches Workshop Programm am Vormittag angeboten. Abgestimmt auf die Themenstellung des Automatisierungstages wurde im ersten Workshop das Thema „Mobile Greifersysteme und Roboterhände“ durch Herrn Dr. Dirk Osswald (Schunk) behandelt und damit für das anwesende Fachpublikum ein tiefer Einblick in diese, für die Kooperation zwischen Mensch, Roboter und Umwelt wesentliche Technologie gewährt. Der zweite Workshopbeitrag beschäftigte sich mit dem relevanten Thema der Sicherheitstechnik. Im Workshop-Vortrag „Integrated Safety Technology“ von Herrn DI Thomas Dicker (Bernecker+Rainer) wurden das Thema Sicherheitstechnik und deren Integration in ein gesamtheitliches Automatisierungssystem eingehend behandelt.

Dieser Automatisierungstag war mit ca 50 Teilnehmern sehr gut besucht. Der 27. Österreichische Automatisierungstag findet am 12. Oktober 2012 an der TU Wien statt und wird von der Abteilung „Regelungstechnik und Prozessautomatisierung“ des Institutes für „Mechanik und Mechatronik“ (Univ. Prof. Dr. Stefan Jakubek) organisiert.

M.Hofbaur, P.Kopacek

Buchbesprechung

Feedback Systems

an Introduction for Scientists and Engineers

Âström K.J. und Murray R. M.

Princeton University Press, 2008
396 Seiten. Gebunden
ISBN-10: 0691165762; ISBN-13: 978-0691165762

Das vorliegende Buch ist eine praxisorientierte Einführung in die Modellbildung, Analyse und den Entwurf von Regelungs- (Rückgekoppelten-) Systemen. Es soll Wissenschaftern und Ingenieuren aus verschiedenen Fachgebieten wie beispielsweise Physik, Biologie, Informationstechnik, Sozial- und Wirtschaftswissenschaften mit dem Wesen der Rückkopplung vertraut machen. Daher werden die mathematischen Grundlagen auf das Notwendigste beschränkt und am Ende jedes Kapitels Beispiele aus den verschiedensten Wissensdisziplinen eingebunden wobei die Lösungen, auf Anfrage, im Internet abrufbar sind.

Die erste Hälfte des Buches basiert ausschließlich auf der Zustandsraumdarstellung. Beginnend mit der Modellierung von Systemen mittels Differential- und Differenzengleichungen, über das dynamische Verhalten von linearen und nichtlinearen Systemen bis zum Reglerentwurf sowie Steuerbarkeit und Beobachtbarkeit werden die Grundlagen anschaulich, ohne mathematischen Ballast, erläutert. Für die Modellierung von Systemen verwenden die Autoren Methoden aus der Physik, der Informationstechnik und des Operations Research

Der zweite Teil des Buches ist dem „klassischen" Zugang zu Regelsystemen gewidmet. Die Werkzeuge werden im Frequenzbereich mittels der Übertragungsfunktionen, PID Regelung, Entwurf im Frequenzbereich und Robustheit von Regelungen erklärt. Abschließend werden die Erkenntnisse der beiden Teile des Buches zusammengeführt und damit die Wechselwirkungen zwischen Robustheit und Effizienz beleuchtet sowie ein Ausblick auf eine tiefgreifendere Befassung mit diesem Fachgebiet gegeben.

Das Buch behandelt sowohl Methoden des Zustandsraumes für Analyse und Entwurf einschließlich der Stabilität. Hier finden Lyapunovfunktionen, Erreichbarkeit, Beobachtbarkeit und Schätzverfahren Anwendung. Zentrale Rolle nimmt dabei die Matrizen-Exponentialfunktion ein, welche eine Erklärung der meisten bekannten Methoden erlaubt.

Das Buch wurde mit dem Textbook Preis der „International Federation auf Automatic Control IFAC" ausgezeichnet.

Zusammenfassend kann festgestellt werden, dass dieses Buch sowohl Studenten als auch Forschern und in der Praxis stehenden Ingenieuren und Nicht-Technikern wärmstens empfohlen werden kann.

P. Kopacek

Instruction to authors – presented as a pattern paper (18 pt)

A. Maier, F. Huber (12 pt)
Department ....., Vienna, Austria

Received April 8, 1999

Abstract
This paper shows …… (italics, 12 pt)

1 General (14 pt)

Authors should prepare their manuscript camera ready, format A 4, 12 typeface and must present their manuscript in good quality. At the left/right edge 2.5 cm, at the to/bottom edge 3 cm. Authors are invited to use papers of this journal as a sample. Please do not use an eraser or erasing fluid. Footnotes should be avoided if possible. Authors are expected to submit their paper to one of the publishers (O.Univ.Prof.Dr. Peter Kopacek, Intelligent Handling Devices and Robotics, Vienna University of Technology, Favoritenstrasse 9-11, A-1040 Vienna, Austria, Fax: +43 1 58801-31899 or O.Univ.Prof.Dr. Alexander Weinmann, Institute of Automation and Control, Vienna University of Technology, Gusshausstr. 27-29, A-1040 Vienna, Austria, Fax: +43 1 58801 37699).

Email address for submitting pdf-manuscripts: weinmann@acin.tuwien.ac.at

2 References (14 pt)

Within the paper references should appear in the following form:
(Mayer, H., 1990) or (*Mayer, H., 1990)* (12 pt);
Mayer, H., 1990, discovered that....

3 Figures and Tables (14 pt)

Figures and Tables should be integrated in the paper and have to be referred to as Fig. 4.1 or Tab. 5.2.

4 References

References are to be listed alphabetically according to first author. (11 pt)

5 Word Processing System/Editor

Microsoft Word 2000 or higher; TeX or LaTeX.

Wenn unzustellbar, retour an:

IFAC-Beirat Österreich (E318 / E376)
Favoritenstraße 9-11, A-1040 Wien /
Gußhausstraße 27-29, A-1040 Wien