Analytic Perturbation Theory and Its Applications

Konstantin E. Avrachenkov
Inria Sophia Antipolis
Sophia Antipolis, France

Jerzy A. Filar
Flinders University
Adelaide, Australia

Phil G. Howlett
University of South Australia
Adelaide, Australia

Analytic Perturbation Theory and Its Applications

siam
Society for Industrial and Applied Mathematics
Philadelphia

10 9 8 7 6 5 4 3 2 1

Library of Congress Cataloging-in-Publication Data

Avrachenkov, Konstantin, author.
Analytic perturbation theory and its applications / Konstantin E. Avrachenkov, Inria Sophia Antipolis, Sophia Antipolis, France, Jerzy A. Filar, Flinders University, Adelaide, Australia, Phil G. Howlett, University of South Australia, Adelaide, Australia.
pages cm
Includes bibliographical references and index.
ISBN 978-1-611973-13-6
1. Perturbation (Mathematics) I. Filar, Jerzy A., 1949- author. II. Howlett, P. G. (Philip G.), 1944- author. III. Title.
QA871.A97 2013
515’.392--dc23
2013033335

To our students, who, we believe,
will advance this topic far beyond
what is reported here.
Though they may not realize it,
we learned from them at least as much
as they learned from us.

Contents

Preface

We live in an era in which ever more complex phenomena (e.g., climate change dynamics, stock markets, complex logistics, and the Internet) are being described with the help of mathematical models, frequently referred to as systems. These systems typically depend on one or more parameters that are assigned nominal values based on the current understanding of the phenomena. Since, usually, these nominal values are only estimates, it is important to know how deviations from these values affect the solutions of the system and, in particular, whether for some of these parameters even small deviations from nominal values can have a big impact.

Naturally, it is crucially important to understand the underlying causes and nature of these big impacts and to do so for neighborhoods of multiparameter configurations. Unfortunately, in their most general settings, multiparameter deviations are still too complex to analyze fully, and even single-parameter deviations pose significant technical challenges. Nonetheless, the latter constitute a natural starting point, especially since in recent years much progress has been made in analyzing the asymptotic behavior of these single-parameter deviations in many special settings arising in the sciences, engineering, and economics.

Consequently, in this book we consider systems that can be disturbed, to a varying degree, by changing the value of a single perturbation parameter loosely referred to as the "perturbation." Since in most applications such a perturbation would be small but unknown, a fundamental issue that needs to be understood is the behavior of the solutions as the perturbation tends to zero. This issue is important because for many of the most interesting applications there is, roughly speaking, a discontinuity at the limit, which complicates the analysis. These are the so-called singularly perturbed problems.

Put a little more precisely, the book analyzes—in a unified way—the general linear and nonlinear systems of algebraic equations that depend on a small perturbation parameter. The perturbation is analytic; that is, left-hand sides of the perturbed equations can be expanded as a power series of the perturbation parameter. However, the solutions may have more complicated expansions such as Laurent or even Puiseux series. These series expansions form a basis for the asymptotic analysis (as the perturbation tends to zero). The analysis is then applied to a wide range of problems including Markov processes, constrained optimization, and linear operators on Hilbert and Banach spaces. The recurrent common themes in the analyses presented is the use of fundamental equations, series expansions, and the appropriate partitioning of the domain and range spaces.

We would like to gratefully acknowledge most valuable contributions from many colleagues and students including Amie Albrecht, Eitan Altman, Vladimir Ejov, Vladimir Gaitsgory, Moshe Haviv, Jean-Bernard Lasserre, Nelly Litvak, (the late) Charles Pearce, and Jago Korf. Similarly, the institutions where we have worked during the long period of writing, University of South Australia, Inria, and Flinders University, have also generously supported this effort. Finally, many of the analyses reported here were carried out as parts of Discovery and International Linkage grants from the Australian Research Council.

Konstantin E. Avrachenkov, Jerzy A. Filar, and Phil G. Howlett

Chapter 1

Introduction and Motivation

1.1 ▪ Background

In a vast majority of applications of mathematics, systems of governing equations include parameters that are assumed to have known values. Of course, in practice, these values may be known only up to a certain level of accuracy. Hence, it is essential to understand how deviations from their nominal values may affect solutions of these governing equations. Naturally, there is a desire to study the effect of all possible deviations. However, in its most general setting, this is a formidable challenge, and hence structural assumptions are usually required if strong, constructive results are to be explicitly derived.

Frequently, parameters of interest will be coefficients of a matrix. Therefore, it is natural to begin investigations by analyzing matrices with perturbed elements. Historically, there was a lot of interest in understanding how such perturbations affect key properties of the matrix. For instance, how will the eigenvalues and eigenvectors of this matrix be affected?

Perhaps the first comprehensive set of answers was supplied in the, now classical, treatise of Kato [99]. Indeed, Kato's treatment was more general and covered the analysis of linear operators as well as matrices. However, Kato [99] and a majority of other researchers have concentrated their effort on the perturbation analysis of the eigenvalue problem.

In this book we shall study a range of problems that is more general than spectral analysis. In particular, we will be interested in the behavior of solutions to perturbed linear and polynomial systems of equations, perturbed mathematical programming problems, perturbed Markov chains and Markov decision processes, and some corresponding extensions to operators in Hilbert and Banach spaces.

In the same spirit as Kato, we focus on the case of analytic perturbations. The latter have the structural form where the perturbed data specifying the problem can be expanded as a power series in terms of first, second, and higher orders of deviations multiplied by corresponding powers of an auxiliary perturbation variable. When that variable tends to zero the perturbation dissipates and the problem reduces to the original, unperturbed, problem. Nonetheless, the same need not be true of the solutions that are of most interest to the researchers studying the system. These can exhibit complex behaviors that involve discontinuities, singularities, and branching.

Indeed, since the 1960s researchers in various disciplines have studied particular manifestations of the complex behavior of solutions to many important problems.

For instance, perturbed mathematical programs were studied by Pervozvanski and Gaitsgori [126], and the study of perturbed Markov chains was, perhaps, formally initiated by Schweitzer [137]. It is this, not uncommon, complexity of the limiting behavior of solutions that stimulated the present book.

1.2 ▪ Raison d'Être and Exclusions

Imagine that the perturbed matrix mentioned in the previous section had the form

$$\tilde{A} = A + D, \tag{1.1}$$

where A is a matrix of nominal coefficient values, $\tilde{A}$ is a matrix of perturbed data, and D is the perturbation itself. There are numerous publications devoted to this subject (see, e.g., the books by Stewart and Sun [147] and Konstantinov et al. [103] and the survey by Higham [80]). However, without any further structural assumptions on D, asymptotic analysis as the norm of D tends to zero is typically only possible when the rank of the perturbed matrix $\tilde{A}$ is the same as the rank of A. Roughly speaking, this corresponds to the case of what we later define to be a *regular perturbation*. Generally, in such a case solutions of the perturbed problem tend to solutions of the original unperturbed problem.

In this book we wish to explain some of the complex asymptotic behavior of solutions such as discontinuity, singularity, and branching. Typically, this arises when the rank of the perturbed matrix $\tilde{A}$ is different from the rank of A. For instance, consider the simple system of linear equations

$$\tilde{A}x = \begin{bmatrix} 1 & 1 \\ 1+\varepsilon & 1+2\varepsilon \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}. \tag{1.2}$$

Clearly, $\tilde{A}$ is of the form (1.1) since we can write

$$\tilde{A} = A + D = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ \varepsilon & 2\varepsilon \end{bmatrix}.$$

Now, for any $\varepsilon \neq 0$, the inverse of $\tilde{A}$ exists and can be written as

$$\tilde{A}^{-1} = \frac{1}{\varepsilon} \begin{bmatrix} 1+2\varepsilon & -1 \\ -1-\varepsilon & 1 \end{bmatrix} = \frac{1}{\varepsilon} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} + \begin{bmatrix} 2 & 0 \\ -1 & 0 \end{bmatrix}.$$

Hence, the unique solution of (1.2) has the form of *Laurent series*

$$\tilde{x} = \frac{1}{\varepsilon} \begin{bmatrix} 1 \\ -1 \end{bmatrix} + \begin{bmatrix} 2 \\ -1 \end{bmatrix}.$$

Despite the fact that the norm of D tends to 0 as $\varepsilon \to 0$, we see that $\tilde{x}$ diverges. The singular part of the Laurent series indicates the direction along which $\tilde{x}$ diverges as $\varepsilon \to 0$.

The above example indicates that a singularity manifests itself in the series expansion of a solution. This phenomenon is common in a wide range of interesting mathematical and applied problems and lends itself to rigorous analysis if we impose the additional assumption that the perturbed matrix is of the form

$$A(\varepsilon) = A_0 + \varepsilon A_1 + \varepsilon^2 A_2 + \dots\ , \tag{1.3}$$

where the above power series is assumed to be convergent in some neighborhood of $\varepsilon = 0$. Hence it is natural to call this particular type of perturbation an *analytic perturbation*.

Consequently, it is also natural to consider a *singular perturbation* to be one where solutions to the perturbed problem are not analytic functions with respect to the perturbation parameter ε.

It will be seen that with the above analytic perturbation assumption, a unified treatment of both the regular and singular perturbations is possible. Indeed, the approach we propose has been inspired by Kato's systematic analysis of the perturbed spectrum problem but applied to a much wider class of problems. Thus, while Kato's motivating problem is captured by the eigenvalue equation

$$A(\varepsilon)x(\varepsilon) = \lambda(\varepsilon)x(\varepsilon), \tag{1.4}$$

our motivating problem is the asymptotic behavior of solutions to the perturbed system of equations

$$f(x,\varepsilon) = 0,$$

where $f(x,\varepsilon)$ can be a system of linear or polynomial equations. In the linear case this reduces to

$$L(\varepsilon)x(\varepsilon) = c(\varepsilon).$$

In particular, if $L(\varepsilon)$ has an inverse for $\varepsilon \neq 0$, and sufficiently small, then we investigate the properties of the perturbed inverse operator $L^{-1}(\varepsilon)$ (or matrix-valued function $A^{-1}(\varepsilon)$ in the finite dimensional case). For example, we rely on the fact that $A^{-1}(\varepsilon)$ can always be expanded as a Laurent series

$$A^{-1}(\varepsilon) = \frac{1}{\varepsilon^s}B_{-s} + \cdots + \frac{1}{\varepsilon}B_{-1} + B_0 + \varepsilon B_1 + \dots \,. \tag{1.5}$$

The preceding system equation $f(x,\varepsilon) = 0$ arises as a building block of solutions to many practical problems. In particular, there is an enormous number of problems that are formulated as either linear or nonlinear mathematical programs. Hence a fundamental question that arises concerns the stability (or instability) of a solution when the problem is slightly perturbed.

Perhaps surprisingly, this can be a very difficult question. Even in the simplest case of linear programming, standard Operations Research textbooks discuss only the most straightforward cases and scrupulously avoid the general issue of how to analyze the effect of a perturbation when the whole coefficient matrix is also affected.

The next example (taken from [126]) illustrates that even in the "trivial" case of linear programming the effect of a small perturbation can be "nontrivial." Consider the simple optimization problem in two variables

$$\max_{x_1,x_2} \; x_2$$

$$\begin{array}{lrcrcl} \text{s.t.} & x_1 & + & x_2 & = & 1, \\ & (1+\varepsilon)x_1 & + & (1+2\varepsilon)x_2 & = & 1+\varepsilon, \\ & & & x_1 \geq 0, \quad x_2 \geq 0. & & \end{array}$$

It is clear that for any $\varepsilon > 0$ there is a unique (and hence optimal) feasible solution at $x_1^* = 1$, $x_2^* = 0$. However, when $\varepsilon = 0$, the two equality constraints coincide, the set of feasible solutions becomes infinite, and the maximum is attained at $\hat{x}_1 = 0$, $\hat{x}_2 = 1$.

More generally, techniques developed in this book permit us to describe the asymptotic behavior of solutions[1] to a generic, perturbed, mathematical program:

[1] The word *solution* is used in a broad sense at this stage. In some cases the solution will, indeed, be a global optimum, while in other cases it will be only a local optimum or a stationary point.

$$\begin{aligned} &\max f(x,\varepsilon) \\ \text{s.t.} \quad & g_i(x,\varepsilon)=0,\ i=1,\dots,m, \\ & h_j(x,\varepsilon)\le 0,\ j=1,\dots,p, \end{aligned} \tag{MP(ε)}$$

where $x \in \mathbb{R}^n, \varepsilon \in [0,\infty)$, and f, g_i's, h_j's are functions on $\mathbb{R}^n \times [0,\infty)$. We will be especially concerned with characterizing solutions, $x^*(\varepsilon)$, of (MP(ε)) as functions of the perturbation parameter, ε. This class of problems is closely related to the well-established topics of sensitivity or postoptimality, or parametric analysis of mathematical programs (see Bonnans and Shapiro [29]). However, our approach covers both the regularly and singularly perturbed problems and thereby resolves instances such as that illustrated in the above simple linear programming example.

Other important applications treated here include perturbed Markov chains and decision processes and their applications to Google PageRank and the Hamiltonian cycle problems.

Let us give an idea of applicability of the perturbation theory to the example of Google PageRank. PageRank is one of the principal criteria according to which Google sorts answers to a user's query. It is a centrality ranking on the directed graph of web pages and hyperlinks. Let A be an adjacency matrix of this graph. Namely, $a_{ij}=1$ if there is a hyperlink from page i to page j, and $a_{ij}=0$ otherwise. Let D be a diagonal matrix whose diagonal elements are equal to the out-degrees of the vertices. The matrix $L=D-A$ is called the graph Laplacian. If a page does not have outgoing hyperlinks, it is assumed that it points to all pages. Also, let v^T be a probability distribution vector which defines preferences of some group of users, and let ε be some regularization parameter. Then, PageRank can be defined by the following equation:

$$\pi = \varepsilon v^T [L+\varepsilon A]^{-1} D.$$

Since the graph Laplacian L has at least one zero eigenvalue, $L+\varepsilon A$ is a singular perturbation of L, and its inverse can be expressed in the form of Laurent series (1.5). This application is studied in detail in Chapter 6.

Consequently, the book is intended to bridge at least some of the gap between the theoretical perturbation analysis and areas of applications where perturbations arise naturally and cause difficulties in the interpretation of "solutions" which require rigorous and yet pragmatic resolution. To achieve this goal, the book is organized as an advanced textbook rather than a research monograph. In particular, a lot of expository material has been included to make the book as self-contained as practicable. In the next section, we outline a number of possible courses that can be taught on the basis of the material covered. Nonetheless, the book also contains sufficiently many new, or very recent, results to be of interest to researchers involved in the study of perturbed systems.

Finally, it must be acknowledged that a number of, clearly relevant, topics have been excluded so as to limit the scope of this text. These include the theories of perturbed ordinary and partial differential equations, stochastic diffusions, and perturbations of the spectrum. Most of these are well covered by several existing books such as Kato [99], Baumgärtel [22], O'Malley [125], Vasileva et al. [153], Kevorkian and Cole [102], and Verhulst [156]. Singular perturbations of Markov processes in continuous time are well covered in the book of Yin and Zhang [162]. Elementwise regular perturbations of matrices are extensively treated in the books of Stewart and Sun [147] and Konstantinov et al. [103].

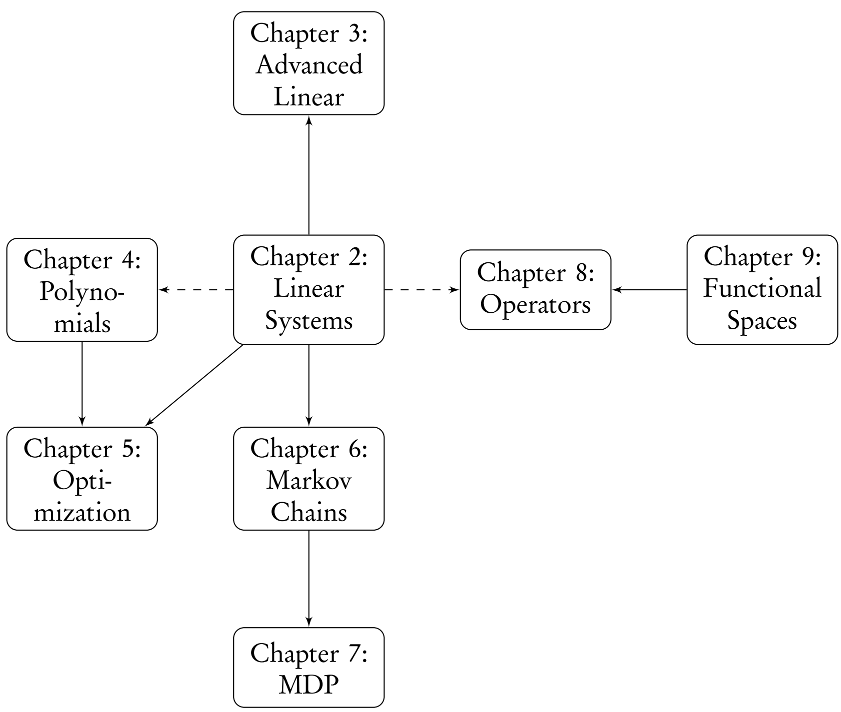
Although the question of numerical computation is an extremely important aspect of perturbation analysis, we shall not undertake per se a systematic study of this topic.

We are well aware that the difference between an exact solution and a numerically computed solution is a prima facie case where perturbation theory may be used to define suitable error bounds. Nevertheless we do recommend that best practice should be used for all relevant numerical computations. This applies particularly to the numerical solution of any collection of key equations.

1.3 ▪ Organization of the Material

Since problems induced by perturbations manifest themselves in a variety of settings, some of which already led to established lines of research, the parts and chapters of this book are arranged so as to facilitate quick access to a wide range of results. The three main parts group chapters containing material related to (I) finite dimensional perturbations, (II) application of results in Part I to optimization and Markov processes, and (III) infinite dimensional perturbations.

The figure below displays some of the logical connections among various chapters. The solid arrows in the figure indicate that a significant part of the material in the chapter at the tail of the arrow is required for understanding the material in the chapter at the head of the arrow. On the other hand, the broken arrows indicate a weaker connection between the corresponding chapters. Indeed, it is possible to follow the material in chapters connected by the solid arrows without prior knowledge of the material in the remaining chapters. Since some readers will already have the requisite knowledge of functional analysis and operator theory, we chose not to precede Chapter 8 with these prerequisites. Instead, we included the latter, presented in a manner best suited to the contents of this book, in the final Chapter 9, which can also serve as a brief, stand-alone introduction to elements of functional analysis.



1.4 ▪ Possible Courses with Prerequisites

As mentioned earlier, in addition to the book's research mission, it can also serve as a textbook for at least the following courses.

1. A one-semester introductory course on perturbation theory of finite dimensional linear systems intended for advanced undergraduates or first year graduate students. This course could be based on Sections 2.1–3.2, Section 5.2, and Section 6.2. The only prerequisites for this course are standard undergraduate linear algebra and calculus courses.

2. A one-semester continuation course on perturbation theory intended for graduate students. This course would take the material covered in the preceding introductory course as assumed knowledge and would cover Section 3.3, Chapter 4, Section 5.4, and Sections 8.1–8.6. Prerequisites for this course include complex analysis and very basic functional analysis. In fact, Chapter 9 and Section 8.5 contain accessible review of the necessary material from Fourier and functional analysis.

3. A one-semester course on perturbation theory of Markov chains and Markov decision processes intended for graduate students. This course would cover the material of Chapters 6 and 7 and could be given as a continuation of any of the above listed courses, or it could be made self-contained if it began with Sections 2.1–2.2, possibly at the cost of omitting some of the later sections of Chapters 6 and 7. This course would require some knowledge of basic probability theory and Markov chains.

4. A one-semester course on perturbation theory in infinite dimensional spaces intended for graduate students. This course would cover the material of Chapters 8 and 9.

1.5 ▪ Future Directions

As with most branches of mathematics there is always more to be done. The interested researcher will clearly recognize that there are many opportunities for continuing the various lines of investigation outlined in this book. Below, we mention only a small sample of these.

1. There are many natural extensions to the multiparameter case.

2. Applications of infinite dimensional general results reported in Chapter 8 should be developed in a number of areas, including optimal control, signal processing, and stochastic processes.

3. Efficient numerical implementations for many of the techniques described here are yet to be devised. Much can be done in the way of numerical computation for many of the problems discussed here using standard mathematical packages such as those available within *Mathematica* and MATLAB. Nevertheless, there is much room for development of problem-specific programs that may or may not call on various standard subroutines from existing packages.

Chapter 2

Inversion of Analytically Perturbed Matrices

2.1 ▪ Introduction and Preliminaries

This chapter and the following one are devoted to a perturbation analysis of the algebraic finite dimensional linear system

$$A(z)x(z) = b(z), \tag{2.1}$$

where the matrix $A(z)$ depends analytically on the parameter z. Namely, $A(z)$ can be expanded as a power series

$$A(z) = A_0 + zA_1 + z^2 A_2 + \dots$$

with some nonzero radius of convergence. Mostly in the exposition of the present chapter, z is a complex number and A_i is a matrix with complex elements. If we want to restrict our consideration to the real numbers, we shall use ε instead of z.

In this chapter we study the linear system (2.1) with a square coefficient matrix $A(z)$. (Systems with rectangular matrices $A(z)$ will be studied in Chapter 3.) In particular, we are interested in the case of *singular perturbations* when $A(0)$ is not invertible but $A(z)$ has an inverse for $z \neq 0$, but sufficiently small. We investigate the properties of the matrix-valued function $A^{-1}(z)$. For example, we provide several methods for expanding $A^{-1}(z)$ as a Laurent series:

$$A^{-1}(z) = \frac{1}{z^s} B_{-s} + \dots + \frac{1}{z} B_{-1} + B_0 + zB_1 + \dots. \tag{2.2}$$

The first method is based on the use of augmented block-Toeplitz type matrices and the Moore–Penrose generalized inverse. The second and third methods are based on reduction techniques which allow us to work with spaces of lower dimension. Then, we give specific methods for cases of a linear perturbation $A(z) = A_0 + zA_1$ and a polynomial perturbation $A(z) = A_0 + \dots + z^p A_p$.

It is easier to explain and to understand the techniques of perturbation theory in terms of matrix inversion. However, we are cognizant that numerical analysts would most likely consider algorithms in the context of the solution of a linear system rather than a matrix inversion. The matrix A^{-1} is simply the solution to the linear equation $AX = I$. In that sense calculation of A^{-1} is equivalent to solving the linear system.

Since the methods of this chapter are essentially based on the application of generalized inverse matrices, we briefly review the main definitions and facts from the theory of generalized inverses. The interested reader can find a more detailed discussion in references provided in the bibliographic notes.

There are several types of generalized inverses. The *Moore–Penrose generalized inverse* (or *Moore–Penrose pseudoinverse*) is by far the most commonly used generalized inverse. It can be defined in either geometric or algebraic terms. First we give a "geometric" definition.

Let $A \in \mathbb{C}^{m\times n}$ be the matrix of a linear transformation from $\mathbb{C}^n$ to $\mathbb{C}^m$. And let $N(A) \subseteq \mathbb{C}^n$ and $R(A) \subseteq \mathbb{C}^m$ denote the null space and the range space of this transformation, respectively. The space $\mathbb{C}^n$ can be represented as the direct sum $N(A)\oplus N(A)^{\perp}$ and the space $\mathbb{C}^m$ can be represented as the direct sum $R(A)\oplus R(A)^{\perp}$.

Definition 2.1. *The Moore–Penrose generalized inverse of the linear transformation* $A : \mathbb{C}^n \to \mathbb{C}^m$ *is a linear transformation* $A^{\dagger} : \mathbb{C}^m \to \mathbb{C}^n$ *defined in the following way. Let* $y \in \mathbb{C}^m$, *and write* $y = y_R + y_R^{\perp}$, *where* $y_R \in R(A)$ *and* $y_R^{\perp} \in R(A)^{\perp}$. *Choose* $x \in \mathbb{C}^n$ *such that* $Ax = y_R$, *and write* $x = x_N + x_N^{\perp}$, *where* $x_N \in N(A)$ *and* $x_N^{\perp} \in N(A)^{\perp}$. *Then* $A^{\dagger}y = x_N^{\perp}$.

Of course, the generalized inverse matrix is just the matrix representation of the corresponding generalized inverse transformation. Next we give an equivalent algebraic definition.

Definition 2.2. *If* $A \in \mathbb{C}^{m\times n}$, *then the Moore–Penrose generalized inverse (or pseudoinverse) is the matrix* $A^{\dagger} \in \mathbb{C}^{m\times n}$ *uniquely defined by the equations*

$$AA^{\dagger}A = A, \tag{2.3}$$

$$A^{\dagger}AA^{\dagger} = A^{\dagger}, \tag{2.4}$$

$$(AA^{\dagger})^* = AA^{\dagger}, \tag{2.5}$$

$$(A^{\dagger}A)^* = A^{\dagger}A, \tag{2.6}$$

where $(\ )^*$ *denotes a conjugate transpose matrix.*

There are several methods for the computation of Moore–Penrose generalized inverses. The best known and, perhaps, the most computationally stable method is based on the singular value decomposition (SVD). Let $r = r(A)$ be the rank of $A \in \mathbb{C}^{m\times n}$. And let $D = \operatorname{diag}\{\sigma_1,\ldots,\sigma_r\}$ be an invertible diagonal matrix, whose diagonal elements are the positive square roots of the nonzero eigenvalues of A^*A repeated according to multiplicity and arranged in descending order. The numbers $\sigma_1,\ldots,\sigma_r$ are usually referred to as the *singular values* of A. Define also two unitary matrices $U \in \mathbb{C}^{m\times m}$ and $V \in \mathbb{C}^{n\times n}$ as follows: u_k, the kth column of matrix U, is a normalized eigenvector of A^*A corresponding to the eigenvalue σ_k^2 and $v_k = Au_k/\sigma_k$. Then, the SVD is given by

$$A = V\begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix}U^*.$$

Now, the generalized inverse $A^{\dagger} \in \mathbb{C}^{n\times m}$ can be written in the form

$$A^{\dagger} = U\begin{bmatrix} D^{-1} & 0 \\ 0 & 0 \end{bmatrix}V^*. \tag{2.7}$$

It is easy to check that the above expression for $A^{\dagger}$ indeed satisfies all four equations (2.3)–(2.6); see Problem 2.1.

The following well-known properties of the Moore–Penrose generalized inverse will be used in what follows:

$$(A^*)^\dagger = (A^\dagger)^*, \tag{2.8}$$
$$A^* = A^*AA^\dagger = A^\dagger AA^*, \tag{2.9}$$
$$(A^*A)^\dagger = A^\dagger A^{*\dagger}, \tag{2.10}$$
$$A^\dagger = (A^*A)^\dagger A^* = A^*(AA^*)^\dagger. \tag{2.11}$$

One can immediately conclude from Definition 2.1 that the generalized inverse is an equation solver. We have the following formal result.

Lemma 2.1. *Consider the assumed feasible system of linear equations*

$$Ax = b, \tag{2.12}$$

where $A \in \mathbb{C}^{m\times n}$ *and* $b \in \mathbb{C}^m$. *Then* x *is a solution of this system if and only if*

$$x = A^\dagger b + v,$$

where $v \in \mathbb{C}^{n\times 1}$ *belongs to the null space of* A, *that is*, $Av = 0$.

The next lemma provides a simple condition for the feasibility of linear systems (see Problem 2.2).

Lemma 2.2. *The system of linear equations* (2.12) *is feasible if and only if* $w^*b = 0$ *for all vectors* $w \in \mathbb{C}^{m\times 1}$ *that span the null space of the conjugate transpose matrix* A^*, *that is*, $A^*w = 0$.

An important particular case of the Moore–Penrose generalized inverse is the so-called *group inverse*, defined as follows.

Definition 2.3. *Suppose that* A *is a square matrix. The group inverse* A^g, *if it exists, is characterized as the unique matrix satisfying the following three equations:*

$$AA^gA = A, \tag{2.13}$$
$$A^gAA^g = A^g, \tag{2.14}$$
$$AA^g = A^gA. \tag{2.15}$$

Existence of the group inverse of $A \in \mathbb{C}^{n\times n}$ is equivalent to the existence of a decomposition of the space $\mathbb{C}^n$ into a *direct* sum of the null space and the range of A (see Problem 2.3).

We now show that computing the Moore–Penrose generalized inverse reduces to computing the group inverse of a square symmetric matrix. This result seems to be new or, at least, not widely reported.

Lemma 2.3. *The Moore–Penrose generalized inverse of* A *can be calculated by the formulae*

$$A^\dagger = (A^*A)^gA^* = A^*(AA^*)^g. \tag{2.16}$$

Proof: By (2.11), to prove the above formulae, we need only verify that the Moore–Penrose generalized inverse $(A^*A)^\dagger$ is also the group inverse of A^*A. Thus, we need to verify that $(A^*A)^\dagger$ satisfies (2.13)–(2.15). It is obvious that (2.13) and (2.14) hold, since by definition, the generalized inverse satisfies (2.3) and (2.4). The last identity (2.15) is obtained via

$$\begin{aligned}(A^*A)(A^*A)^\dagger &= A^*AA^\dagger A^{*\dagger} = A^*AA^\dagger A^{\dagger *} = A^*A^{\dagger *}\\ &= (A^\dagger A)^* = A^\dagger A = (A^*A)^\dagger (A^*A),\end{aligned}$$

using (2.10), (2.8), (2.9), (2.6), and (2.11), respectively. Thus, the matrix $(A^*A)^\dagger$ satisfies its analogue of (2.15), and, therefore, $(A^*A)^\dagger = (A^*A)^g$, which immediately yields (2.16). □

Now let us discuss another type of generalized inverse, the so-called *Drazin inverse*. The Drazin inverse can be defined and calculated in the following way: If $A \in \mathbb{C}^{n\times n}$, then it can be represented by the decomposition

$$A = W\begin{bmatrix} S & 0 \\ 0 & N \end{bmatrix} W^{-1}, \tag{2.17}$$

where S is invertible and N is nilpotent. Then, the Drazin inverse is defined by

$$A^\# = W\begin{bmatrix} S^{-1} & 0 \\ 0 & 0 \end{bmatrix} W^{-1}.$$

Note that the Drazin inverse is not an equation solver. However, based on algebraic properties, Drazin inverses have more in common with usual matrix inverses than Moore–Penrose generalized inverses do. In spectral theory of linear operators the Drazin inverse is also known as *reduced resolvent*.

The group inverse is also a particular case of the Drazin inverse. Namely, whenever for a matrix A the group inverse exists, A can be decomposed into (2.17) with $N = 0$. In fact, the group inverse represents the case when the Moore–Penrose generalized inverse and the Drazin inverse coincide.

2.2 ▪ Inversion of Analytically Perturbed Matrices: Algebraic Approach

2.2.1 ▪ Laurent series and fundamental equations

Let $\{A_k\}_{k=0,1,\dots} \subseteq \mathbb{C}^{n\times n}$ be a sequence of matrices that defines an analytic matrix-valued function

$$A(z) = A_0 + zA_1 + z^2A_2 + \cdots . \tag{2.18}$$

The above series is assumed to converge in some nonempty neighborhood of $z = 0$. In such a case we say that $A(z)$ is an *analytic perturbation* of the matrix $A_0 = A(0)$. Assume the inverse matrix $A^{-1}(z)$ exists in some (possibly punctured) disc centred at $z = 0$. We are primarily interested in the case when A_0 is singular. The next theorem shows that $A^{-1}(z)$ can be expanded as a Laurent series.

Theorem 2.4. *Let $A(z)$ be an analytic matrix-valued function of z in some nonempty neighborhood of $z = 0$ and such that $A^{-1}(z)$ exists in some (possibly punctured) disc centered at*

$z=0$. Then, $A^{-1}(z)$ possesses a Laurent series expansion

$$A^{-1}(z)=\frac{1}{z^s}(X_0+zX_1+\cdots), \tag{2.19}$$

where $X_0 \neq 0$ and s is a natural number, known as the order of the pole at $z=0$.

Proof: Using the Cramer formula, we can write

$$A^{-1}(z)=\frac{\operatorname{adj}A(z)}{\det A(z)}. \tag{2.20}$$

Since the determinant $\det A(z)$ and the elements of the adjugate matrix $\operatorname{adj}A(z)$ are polynomials in $a_{ij}(z)$, $i,j=1,\ldots,n$, they are analytic functions of z. The division of two analytic functions yields a meromorphic function. Since n is finite, the order of the pole s in the matrix Laurent series (2.19) is finite as well. □

We would like to note that the above proof is essentially based on the finiteness of the dimension of the underlying space. The case of infinite dimensional spaces will be treated in Chapter 8.

Example 2.1. *Let us consider the following example of the analytically perturbed matrix:*

$$A(z)=\begin{bmatrix} 1-z & 1+z \\ 1-2z & 1-z \end{bmatrix}.$$

According to the formula (2.20), the inverse is given by

$$A^{-1}(z)=\frac{1}{-z(1-3z)}\begin{bmatrix} 1-z & -1-z \\ -1+2z & 1-z \end{bmatrix}.$$

Next, to obtain the Laurent series (2.19), we just expand $(\det A(z))^{-1}=1/(-z(1-3z))$ as a scalar power series, multiply it by $\operatorname{adj}A(z)$, and collect coefficients with the same power of z. In this case, we have

$$A^{-1}(z)=\left(-\frac{1}{z}-3-9z^2-\ldots\right)\begin{bmatrix} 1-z & -1-z \\ -1+2z & 1-z \end{bmatrix}$$
$$=\frac{1}{z}\begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}+\begin{bmatrix} -2 & 4 \\ 1 & -2 \end{bmatrix}+z\begin{bmatrix} -6 & 12 \\ 3 & -6 \end{bmatrix}+\ldots.$$

Of course, the direct application of the Cramer formula (2.20) as in the above example is very inefficient as a method of deriving the Laurent series (2.19). Thus, the main purpose of this section is to provide efficient computational procedures for calculating the Laurent series coefficients X_k, $k\geq 0$.

In fact, we present three methods for computing the coefficients of the Laurent series (2.19) for the inverse of the analytically perturbed matrix (2.18). The first method is based on a direct application of the Moore–Penrose generalized inverse matrix. The other two methods are based on a so-called *reduction technique*. All three methods depend essentially on equating coefficients of powers of z.

By substituting the series (2.18) and (2.19) into the identity $A(z)A^{-1}(z)=I$ and collecting coefficients of the same powers of z, one obtains the following system, which we