Linear Algebra, Rational approximation and Orthogonal polynomials

Adhemar Bultheel and Marc Van Barel

Both authors are at

K.U.Leuven
Department of Computer Science,
Celestijnenlaan 200A
B-3001 Heverlee
Belgium

email addresses:

`Adhemar.Bultheel@cs.kuleuven.ac.be`
`Marc.VanBarel@cs.kuleuven.ac.be`

Correspondence to:
Adhemar Bultheel

Preface

It is very interesting to see how the same principles and techniques are developed in different fields of mathematics, quite independently from each other. It is only after a certain maturity is reached that a result is recognized as a variation on an old theme. This is not always a complete waste of effort because each approach has its own merits and usually adds something new to the existing theory. It is virtually impossible and there is no urgent need to stop this natural development of things. If tomorrow a new application area is emerging with its own terminology and ways of thinking, it is not impossible that an existing method is rediscovered and only later recognized as such. It will be the twist in the formulation, the slightly different objectives, an extra constraint etc. that will revive the interest in an old subject. The availability of new technological or mathematical tools will force researchers to rethink theories that have long been considered as complete and dead as far as research is concerned.

In this text we give a good illustration of such an evolution. For the underlying principle we have chosen the algorithm of Euclid. It is probably the oldest known nontrivial algorithm, which can be found in the most elementary algebra textbooks. It is certainly since the introduction of digital computers and the associated algorithmic way of thinking that it received a new impetus by its many applications which resulted in an increasing interest, but its simplicity in dealing with situations that are, at least in certain problem formulations not completely trivial, explains its success. Already in ancient times, long before modern computers became an essential part of the scene, the algorithm has been used in many applications.

In its original form, it deals with a geometrical problem. At least, Euclid himself describes it in his 7th book of the *Elements* as a way of constructing the largest possible unit rule which measures the length of two given rules as an integer times this unit rule. Nowadays we recognize the Euclidean algorithm as a method to compute the greatest common divisor of two integers or of two polynomials. This may seem a trivial step, yet, it links geometry, algebra and number theory. Of course, the distinction between different mathematical disciplines is purely artificial and is invented by mathematicians. Luckily, the self-regulating mathematical system maintains such links between the different types of mathematicians, and prevents them from drifting too far apart. This trivial observation would not justify this text if there weren't many more applications of this computational method.

It was recognized with the invention of continued fractions that the algorithm does not only compute the final result: the greatest common divisor, but all the intermediate results also appear as numerators or denominators in a (finite) continued fraction expansion of a rational number. The next, quite natural step, is to apply the algorithm to test whether a number is rational or not and to see what happens when the number is not rational. The algorithm will go on indefinitely, and we get an infinite continued fraction expansion. The study of these expansions became a useful and powerful tool in number theory.

The same technique can be applied to (infinite formal) power series, rather than just polynomials, and again the algorithm will end after a finite number of steps if we started from a representation of a rational fraction or it will give an infinite (formal) expansion which might or might not converge in a certain region e.g., of the complex plane. This hooks up the Euclidean algorithm with (rational) approximation theory. From an algebraic point of view, the Euclidean domains (and several of its generalizations) became a study object in their own right. From the approximation side, the kind of rational approximants that you get are known as Padé approximants. Although this theory has celebrated its hundredth anniversary, the recognition of the Euclidean algorithm as an elegant way of constructing approximants in a nonnormal table came only a couple of decades ago.

Traditionally, the computation of a Padé approximant was done via the solution of a linear system of equations. The matrix of the system has a special structure because it contains only $O(n)$ different elements while the entries on an antidiagonal are all the same. It is a Hankel matrix. The linear algebra community became interested in solving this special type of systems because their solution can be obtained with sequential computations in $O(n^2)$ operations, rather than the $O(n^3)$ for general systems. The Hankel matrices, and the related Toeplitz matrices, show up in numerous applications and have been studied for a long time and that explains why it is important to have efficient solution methods for such systems which fully exploit their special structure. Not only the solution of such systems, but also other, related linear algebra problems can benefit from the known theory of the Euclidean algorithm.

The Lanczos algorithm, which is recently rediscovered as a fashionable research area is intimately related to the Euclidean algorithm. Again the algorithm of Euclid can serve as an introduction to the fast growing literature on fast algorithms for matrices with special structure, of which Toeplitz and Hankel matrices are only the most elementary examples.

Connected, both to Toeplitz and Hankel matrices and to Padé approximation, are orthogonal polynomials and moment problems. Given the moments for some inner product, the problem is to find the orthogonal polynomials and eventually the measure itself. For a measure with support on the real line, the moment matrix is typically a Hankel matrix, for Szegő's theory where the support is the unit circle of the complex plane, the moment matrix is typically Toeplitz. To have a genuine inner product, the moment matrices should be positive definite and strongly nonsingular, that is, all its principal leading minors are nonsingular. In Padé approximation, this has been formally generalized to orthogonality with respect to some linear functional and the denominators of the approximants are (related to) the orthogonal polynomials. However, in this situation, the moment matrix is Hankel, but in general neither positive definite nor strongly nonsingular and then the Euclidean algorithm comes again to the rescue because it helps to jump over the singular blocks in a nonnormal Padé table. A similar situation occurs in the problem of Laurent-Padé approximation which is related to a Toeplitz moment matrix, and also here the matrix is neither positive definite nor strongly nonsingular. The analogs of the Euclidean algorithm which can handle these singular situations are generalizations of the Schur and Szegő recursions in classical moment theory.

A final cornerstone of this text that we want to mention here is linear systems theory. This is an example of an engineering application where many, sometimes deep, mathematical results come to life. Here, both mathematicians and engineers are active to the benefit of both. The discipline is relatively young. It was only since the nineteenthirties that systems theory became a mathematical research area. In the current context we should mention the minimal partial realization problem for linear systems. It is equivalent with a Padé approximation problem at infinity. The minimality of the realization is however important from a practical point of view. The different formulation and the extra minimality condition makes it interesting because classical Padé approximation doesn't give all the answers and a new concept of minimal Padé approximation is the natural equivalent in the theory of Padé approximation. A careful examination of the Euclidean algorithm will reveal that it is actually a variant of the Berlekamp-Massey algorithm. The latter was originally developed as a method for handling error-correcting codes by shift register synthesis. It became known to the engineering community as also solving the minimal partial realization problem. Another aspect that makes systems theory interesting in this aspect is that in this area it is quite natural to consider systems with n inputs and m outputs. When m and n are equal to 1, we get the scalar theory, but with m and n larger than 1, the moments, which are called Markov parameters in this context, are matrices and the denominators of the realizations are square polynomial matrices and the numerators rectangular polynomial matrices. So, many, but not all, of the previously mentioned aspects are generalized to the matrix or block matrix case for multi-input multi-output systems. In the related areas of linear algebra, orthogonal polynomials and Padé approximation, these block cases are underdeveloped to almost nonexisting at all. The translation of results from multi-input multi-output systems theory to the fields mentioned above was one of the main incentives for putting the present text together.

In this book we shall only consider the scalar theory and connections that we have sketched above.

Excellent textbooks exist on each of the mentioned areas. In most of them the Euclidean algorithm is implicitly or explicitly mentioned, but the intimate interplay between the different fields is only partially covered. It is certainly not our intention to replace any of these existing books, but we want in the first place put their interconnection at the first plan and in this way we hope to fill an empty space.

We make the text as selfcontained as possible but it is impossible to repeat the whole theory. If you are familiar with some of the fields discussed it will certainly help in understanding our message. For the theory of continued fractions and their application in Padé approximation as well as in number theory, you can consult the book by Perron *Die Lehre von den Kettenbrücken* [202] and Wall's *Analytic theory of continued fractions* [237, 238] which are classics, but Jones and Thron's *Continued fractions, analytic theory and applications* [161] can be considered as a modern classic in this domain. The most recent book on the subject is *Continued fractions with applications* [183] by Lorentzen and Waadeland who include also many applications, including orthogonal polynomials and signal processing. For Padé approximations you may consult e.g. Baker's classic: *Essentials of Padé Approximation* [5], but a more recent substitute is the new edition of the book *Padé Approximants* by Baker and Graves-Morris [6, 7]. See also [69]. The connection between Padé approximation and formal orthogonal polynomials is explicitly discussed by Brezinski in *Padé-type approximation* *and general orthogonal polynomials* [24]. And much more on formal orthogonal polynomials is given in A. Draux's *Polynômes orthogonaux formels – applications* [87]. For the linear algebra aspects, the book by Heinig and Rost *Algebraic methods for Toeplitz-like matrices and operators* [144] is a cornerstone that resumes many results. For the theory of linear systems there is a vast literature, but Kailath's book *Linear systems* [165] is a very good introduction to many of the aspects we shall discuss.

First we guide the reader from the simplest formulation of the Euclidean algorithm to a more abstract formulation in a Euclidean domain. We give some of its variants and some of its most straightforward applications.

In a second chapter we discuss some aspects and applications in linear algebra, mainly including the factorization of Hankel matrices.

In the third chapter, we give an introduction to the Lanczos algorithm for unsymmetric matrices and some of its variants.

The fourth chapter on orthogonal polynomials translates the previous results to orthogonal polynomials with respect to a general biorthogonal form with an arbitrary moment matrix. The Hankel matrices that were studied in previous chapters are a very special case. We also give some results about Toeplitz matrices which form another important subclass. As a preparation for the matrix case, we give most formulations in a noncommutative field which forces us to use left/right terminology. This is not really essential, but it forces us to be careful in writing the products and inverses so that the results reflect already (at least partially) the block case.

Chapter 5 treats Padé approximations. Perhaps the most important results of this chapter are the formulations of minimal Padé problems and the method to solve them.

The next chapter gives a short introduction to linear systems and illustrates how the previous results can be used in this context. It includes a survey of recent developments in stability tests of Routh-Hurwitz and Schur-Cohn type.

Finally, Chapters 7 and 8 give some less elaborated perspectives of further applications which are closely related to what has been presented in the foregoing chapters.

Chapter 7 gives a general framework for solving very general rational interpolation problems of which (scalar) Padé approximants are a special case. It also introduces the look-ahead strategy for solving such problems and which is most important in numerical computations. It is left to the imagination of the reader to translate the look-ahead ideas to all the other interpretations one can give to these algorithms in terms of rational approximation, of orthogonal polynomials, iterative methods for large matrices, solution of structured systems etc.

The last chapter introduces the application of the Euclidean algorithm in the set of Laurent polynomials to the factorization of a polyphase matrix into a product of elementary continued fraction-like matrices. These polyphase matrices occur is the formulation of wavelet transforms and the factorization is interpreted as primal and dual lifting steps which allow for an efficient computation of wavelet transform and its inverse.

While we were preparing this manuscript, we became aware of the Ph. D. thesis by Marlis Hochbruck [151] which treats similar subjects. As for the iterative solution of systems, Claude Brezinski is preparing another volume in this series [28] which is completely devoted to this subject.

It is clear that in this monograph all the topics of the project ROLLS are present. That is Rational approximation, Orthogonal functions, Linear algebra, Linear systems, and Signal processing. The remarkable observation is that the Euclidean algorithm, in one form or another, is a "greatest common divisor" of all these topics.

Acknowledgement: This research was supported by the Belgian National Fund for Scientific Research (NFWO) project LANCZOS under contract #2.0042.93 and the Human Capital and Mobility project ROLLS under contract CHRX-CT93-0416.

Contents