Quadratic Form Theory and Differential Equations

John Gregory
DEPARTMENT OF MATHEMATICS
SOUTHERN ILLINOIS UNIVERSITY AT CARBONDALE
CARBONDALE, ILLINOIS

1980

ACADEMIC PRESS
A Subsidiary of Harcourt Brace Jovanovich, Publishers
New York London Toronto Sydney San Francisco

ACADEMIC PRESS, INC.
111 Fifth Avenue, New York, New York 10003

United Kingdom Edition published by
ACADEMIC PRESS, INC. (LONDON) LTD.
24/28 Oval Road, London NW1 7DX

Library of Congress Cataloging in Publication Data

Gregory, John, Date
Quadratic form theory and differential equations.

(Mathematics in science and engineering)
Bibliography: p.
Includes index.
1. Forms, Quadratic. 2. Differential equations,
Partial. 3. Calculus of variations. I. Title.
II. Series.
QA243.G73 512.9'44 80 520
ISBN 0-12-301450-6

PRINTED IN THE UNITED STATES OF AMERICA

80 81 82 83 9 8 7 6 5 4 3 2 1

Preface

Historically, quadratic form theory has been treated as a rich but misunderstood uncle. It appears briefly, almost as an afterthought, when needed to solve a variety of problems. A partial list of such problems includes the Hessian matrix in n-dimensional calculus; the second variational (Jacobi or accessory) problem in the calculus of variations and optimal control theory; Rayleigh–Ritz methods for finding eigenvalues of real symmetric matrices; the Aronszajn–Weinstein methods for solving problems of vibrating rods, membranes, and plates; oscillation, conjugate point, and Sturm comparison criteria in differential equations; Sturm–Liouville boundary value problems; spline approximation ideas for numerical approximations; Gershgorin-type ideas (and the Euler–Lagrange equations) for banded symmetric matrices; Schrödinger equations; and limit-point–limit-circle ideas of singular differential equations in mathematical physics.

A major purpose of this book is to develop a unified theory of quadratic forms to enable us to handle the mathematical and applied problems described above in a more meaningful way. Our development is on four levels and should appeal to a variety of users of mathematics. For the theoretically inclined, we present a new formal theory of approximations of quadratic forms/linear operators on Hilbert spaces. These ideas allow us to handle a wide range of problems. They also allow us to solve these problems in a qualitative and quantitative manner more easily than with more conventional methods. Our second level of development is qualitative in nature. Using this theory, we can derive very general qualitative comparison results such as generalized Sturm separation theorems of differential equations and generalized Rayleigh–Ritz methods of eigenvalues. Our theory is also quantitative in nature. We shall derive in level three an approximation theory that can be applied in level four to give numerical algorithms that are easy to implement and give good numerical results.

Our development will provide several bonuses for the reader. A major advantage is that our numerical theory and algorithms are designed to be used with high-speed computers. The computer programs are small and easy to implement. They trade detailed analysis by and sophistication on the part of the user for large numbers of computer computations that can be performed in fractions of milliseconds. Another advantage is that our four levels can be understood and used (virtually) independently of each other. Thus our numerical algorithms can be understood and implemented by users with little mathematical sophistication. For example, for eigenvalue problems, we need no understanding of projection operators, Hilbert spaces, convergence, Green's functions, or resolvent operators. We need only the idea of the Euler–Lagrange equation, an idea that we can obtain a discrete solution as a result of level one, a one-step–three-term difference equation, and an interval-halving procedure.

As with any mathematical theory, we shall leave the reader with several research problems still unanswered. In the area of discrete mathematics, we present for splines and for real symmetric banded or block diagonal symmetric matrices a use that may stimulate further research. For those problems in optimal control theory, we expect our methods, which give qualitative results, to give quantitative results similar to those obtained for the calculus-of-variations case. For the area of limit-point–limit-circle differential equations and singular differential equations (Bessel, Legendre, Laguerre), we expect our ideas to carry over to this very important area of mathematical physics. For the area of differential equations, we hope that our ideas on integral-differential equations can lead to new ideas for oscillation theory for non-self-adjoint problems.

Our concept of quadratic form theory began with the landmark *Pacific Journal of Mathematics* paper by Professor Magnus Hestenes in 1951. For many years, he was convinced that there should be a unified method for problems dealing with a quadratic form $J(x)$ on a Hilbert space $\mathscr{A}$. A major part of his work depends upon two nonnegative integer-valued functions s and n, which correspond to the number of negative and zero eigenvalues of $J(x)$. In subsequent years, Hestenes and his students showed how this theory could be applied to solve a multitude of applied problems.

In 1970 the author developed, in a Ph.D. thesis under Professor Hestenes at the University of California, Los Angeles, an approximating theory of quadratic forms $J(x;\sigma)$ defined on Hilbert spaces $\mathscr{A}(\sigma)$, where σ is a parameter in a metric space. In this and subsequent work, this approximation theory has been used to solve the types of problems listed above. A major part of our work involves the development and interpretation of inequalities concerning $s(\sigma)$ and $n(\sigma)$ as σ approaches a fixed member σ_0 of the matrix space Σ.

In Chapter 1 we take a look backward at more classical methods and ideas of quadratic forms. It may initially be read briefly for flavor and interest since this material is not completely necessary for subsequent chapters. We begin this chapter with finite-dimensional quadratic forms. Many of these ideas will be new to even the sophisticated reader and will appear in an infinite-dimensional context in later parts of the text. The topics include the duality between quadratic forms and symmetric matrices, stationary conditions of f: $\mathbb{R}^n \rightarrow \mathbb{R}^1$, Rayleigh–Ritz methods, and eigenvalues as comparison parameters. Section 1.2 contains a brief introduction to the calculus of variations and in particular the second variation. Of interest is that the Euler–Lagrange necessary conditions are differential equations. In Section 1.3 we cover a general theory of integration by parts and multiplier rules. Section 1.4 explores briefly the relationship between quadratic forms and differential equations. Many examples are included, covering the simpler second-order problems to the more difficult $2n$th-order control theory or partial differential equations.

Chapter 2 may also be initially read for flavor by all but the theoretical mathematician since it contains our theoretical machinery and results. Section 2.1 contains the basic Hilbert space material, which was given by Hestenes and which forms the basis of our approximation theory. The majority of the material in Section 2.2 is more general than needed for the remainder of this book. Section 2.3 is our fundamental theoretical section yielding nonnegative integer inequalities. Briefly, if $s(\sigma)$ and $n(\sigma)$ correspond to the number of negative and zero eigenvalues of a quadratic form or symmetric matrix, then for σ "close to" σ_0 we obtain $s(\sigma_0) \leq s(\sigma) \leq s(\sigma) + n(\sigma) \leq s(\sigma_0) + n(\sigma_0)$. This innocent-looking inequality is used extensively throughout this book.

Chapter 3 is a complete discussion of the second-order problem, and the reader is strongly advised to begin here. We have made a serious attempt to make our ideas in this chapter conceptually clear and descriptive so as to be readily understood. In a real sense, Chapter 3 is a book unto itself. The nontheoretical parts may be understood by senior-level students in mathematics and the physical sciences. Once grasped, the remainder of the book can at least be read for the flavor of more general examples. Formal proofs have been postponed until the last section of this chapter. We begin Chapter 3 with a discussion of the duality of focal-point theory of quadratic forms and the oscillation theory of differential equations. Section 3.2 contains approximation ideas and shows how to build numerical solutions for differential equations. Sections 3.3 and 3.4 contain general theories for eigenvalue problems. The unified setting yields numerical–eigenvalue–focal-point theories and results, as well as efficient and accurate computer algorithms.

Chapter 4 contains the most general ordinary-differential-system–quadratic-form problem, namely, the self-adjoint $2n$th-order integral-differential case begun by Hestenes and Lopez. The exposition is primarily theoretical, but in Section 4.4 we do give numerical ideas of higher-order spline approximations and banded symmetric matrices. Section 4.1 contains the work of Lopez relating quadratic forms and differential equations. Section 4.2 contains our approximation theory. Section 4.3 presents a general comparison theory and results that are applicable to a variety of problems.

Chapter 5 contains the elliptic partial differential equation theory begun by Hestenes and Dennemeyer; this theory is contained in Section 5.1. The numerical construction of conjugate (or focal) surfaces for Laplacian-type partial differential equations, including eigenvalue results, is given in Section 5.2. In Section 5.3 we give a separation-of-variables theory for quadratic forms and new ideas for block tridiagonal matrices.

Chapter 6 contains a general theory of quadratic control problems begun by Hestenes and Mikami. In particular, in Section 6.1 we generalize the concepts of oscillation, focal, and conjugate point to focal intervals and show how to count and approximate them. The concept of abnormality is the key idea here, which distinguishes conjugate-point (calculus-of-variations) problems and focal-interval (optimal control theory) problems. In Section 6.2 we apply these ideas to solutions of differential equations. In Section 6.3 we give two nontrivial examples to illustrate abnormality. Finally, in Section 6.4 we apply our approximation ideas a second time to obtain an approximation theory of focal intervals.

It should be evident that we have been influenced by many distinguished scholars whose works cover several centuries. We should like particularly to acknowledge the work and guidance of Professor Magnus Hestenes in the beginning of this effort. Quadratic form theory is only one of at least four major mathematical areas that bear his stamp. To paraphrase one of our most illustrious forefathers, "If we have seen further than others, it is because we have stood on the shoulders of giants." We should like to acknowledge Lewis Williams and Ralph Wilkerson for their support in the generation of computer algorithms that appear in this text. We acknowledge Joseph Beckenbach for his fine illustrations, Sharon Champion for her expert typing and patience in reading handwritten pages, and the author's charming wife, Virginia, for her editorial corrections. Finally, the author would like to thank Professor Richard Bellman for inviting him to write this book at an early stage of its development, thus providing the encouragement to complete the task.

Contents

Preface ix

Chapter 0 A Few Introductory Remarks 1

Chapter 1 Introduction to Quadratic Forms and Differential Equations

1.0 Introduction 4
1.1 The Finite-Dimensional Case 5
1.2 The Calculus of Variations 25
1.3 Fundamental Lemmas (Integration by Parts) 31
1.4 Quadratic Forms and Differential Equations 38

Chapter 2 Abstract Theory

2.0 Introduction 58
2.1 Hilbert Space Theory 59
2.2 Further Ideas of Hestenes 62
2.3 Approximation Theory of Quadratic Forms 73

Chapter 3 The Second-Order Problem

3.0 Introduction 82
3.1 The Focal-Point Problem 83
3.2 The Numerical Problem 88
3.3 The Eigenvalue Problem 103
3.4 The Numerical Eigenvalue Problems 114
3.5 Proofs of Results 135

Chapter 4 The 2nth-Order Problem

4.0 Introduction 140
4.1 The Signature Theory of Lopez 143
4.2 Approximation Theory 156
4.3 Comparison Results 160
4.4 Higher-Order Numerical Problems and Splines 166

Chapter 5 Elliptic Partial Differential Equations

5.0 Introduction 174
5.1 Summary 175
5.2 The Numerical Problem 188
5.3 Separation of Variables 197

Chapter 6 The Quadratic Control Problem

6.0 Introduction 201
6.1 Focal-Interval Theory of Quadratic Forms 202
6.2 Focal Arcs of Differential Equations 208
6.3 Two Examples 215
6.4 An Approximation Theory of Focal Intervals 221

Postscript The Numerical Problem Revisited

1 The $x(t)x'(t)$ Term 225
2 Cheap Boundary-Value Methods 226
3 Systems 227
4 Nonlinear Problems 229

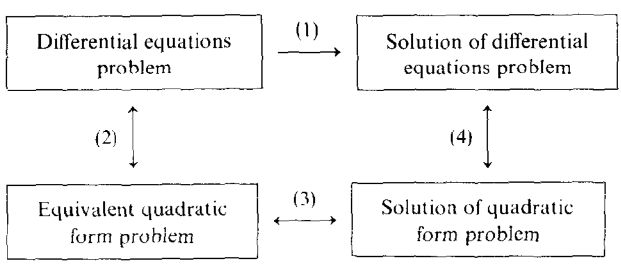
References 231

Index 235

Chapter 0 A Few Introductory Remarks

On May 14, 1979, I had just arrived at the Technical University in Wroclaw, Poland. My luggage and hence my notes had not yet arrived due to the inefficiency of the American (not the Polish) airlines. There was a scheduled Monday morning seminar, and I was asked if I should like to speak, essentially on the spur of the moment. I replied, as one must in those circumstances, "Of course." It seems reasonable that the summary of such a presentation before a charming though general audience, having some language difficulties, might form an introduction to this book.

From the point of view of this book, the following diagram is fundamental:



Usually, people working on differential equations proceed on path (1). By this statement, we mean that they have their own methods to solve their problems. Thus, a numerical problem might call for divided difference methods, while oscillation theory might call for Sturm theory type arguments. Our approach will be to convert the differential equation into the equivalent quadratic form, path (2); solve this quadratic form problem, path (3); then convert back into the solution, path (4). These methods seem to require more steps. However, the steps are often easier to accomplish and are more enlightening. We get better results, methods, and ease of applicability. In addition, we have more flexibility and more ability to generalize to more difficult problems with less additional effort.

Three example problem areas come to mind, and we shall quickly describe them in the next few paragraphs, deferring a more thorough explanation until Chapter 3. We ask the reader to skim the next few paragraphs for the cream and not be concerned about details. Equally important, we ask the reader to note that these examples can be easily combined by our ideas, a process not easily performed on path (1). We shall illustrate a numerical oscillation eigenvalue theory of differential equations at the end of the next few paragraphs.

Let $L(x)$ be a linear self-adjoint, differential operator, and $Q(x)$ be the associated quadratic form, such as our most elementary infinite example

$$L(x) = x''(t) + x(t) = 0, \tag{1}$$

$$Q(x) = \int_0^b (x'^2 - x^2)\, dt, \tag{2a}$$

and

$$Q(x, y) = \int_0^b [x'(t)y'(t) - x(t)y(t)]\, dt. \tag{2b}$$

For (1) we wish to study conjugate or oscillation points relative to $t = 0$; that is, point λ such that there is a nontrivial solution of (1), denoted $x_0(t)$, such that $x_0(0) = x_0(\lambda) = 0$. (1) is the Euler–Lagrange equation of (2). It is obtained by integration by parts or a divergence theorem. Let $\mathscr{B}(\lambda)$ denote the collection of smooth functions such that $x(t)$ is in $\mathscr{B}(\lambda)$ implies $x(0) = 0$ and $x(t) \equiv 0$ on $[\lambda, b]$. We shall see that $\mathscr{B}(\lambda)$ is a subspace of a Hilbert space. For (2), we wish to determine the signature $s(\lambda)$, that is, the dimension of $\mathscr{C}$ where $\mathscr{C}$ is a maximal subspace of $\mathscr{B}(\lambda)$ with respect to the property that $x \neq 0$ in $\mathscr{C}$ implies $Q(x) < 0$. That is, $s(\lambda)$ is the dimension of a negative space of $\mathscr{B}(\lambda)$. Let $n(\lambda) = \dim\{x \text{ in } \mathscr{B}(\lambda) \mid Q(x, y) = 0 \text{ for } y \text{ in } \mathscr{B}(\lambda)\}$. These two nonnegative indices correspond, respectively, to the number of negative and zero eigenvalues of a real symmetric matrix.

Instead of finding the zeros of (1) subject to $y(0) = 0$, path (1), we convert $L(x)$ to $Q(x)$, path (2), solve the signature $s(\lambda)$ for each $0 \le \lambda \le b$, path (3), and finally use the result that for λ_0 in $[0, b]$,

$$s(\lambda_0) = \sum_{\lambda < \lambda_0} n(\lambda). \tag{3}$$

Thus, $s(\lambda_0)$ counts the number of oscillation points before $t = \lambda_0$.

Similarly, the eigenvalue differential equation $L(x; \xi) = x'' + \xi x = 0$, $x(0) = x(\pi)$ is converted to a quadratic form

$$J(x; \xi) = J(x) - \xi K(x) = \int_0^\pi x'^2\, dt - \xi \int_0^\pi x^2\, dt.$$

This time, $s(\xi)$ is the signature of $J(x;\xi)$ on a smooth space of functions defined on $[0, \pi]$ vanishing at the end points. We solve this problem, path (3); then establish the equivalence between an eigenvalue ξ_0 and the discontinuity in $s(\xi_0)$, path (4).

Similarly, for numerical problems, we convert (1) to a quadratic form (2a), in path (2), numerically approximate (2a) by a finite-dimensional quadratic form, path (3), and then show that this approximation leads to a numerical solution that converges to the desired result in a very strong, derivative norm sense, path (4). As we have remarked before, our methods allow us to combine these three problems in a relatively simple manner to obtain a numerical oscillation theory of eigenvalues.

Chapter 1 Introduction to Quadratic Forms and Differential Equations

1.0 Introduction

The purpose of this chapter is to present to the reader much of the beauty and many of the fundamental ideas of quadratic forms. This chapter is an introduction to the remainder of this book. It may be read (and reread) for interest and examples, or it may be skipped entirely by those who are only interested in specific problems.

Section 1.1 treats the finite-dimensional case or equivalently a real symmetric matrix. Since most readers may be familiar with the usual ideas, we have included several topics that illustrate important ideas which are not commonly known nor understood. We believe even the expert will find these topics of interest. The topics are: (a) the duality between finite-dimensional quadratic forms and matrices; (b) optimal or stationary conditions of $f:\mathbb{R}^n \to \mathbb{R}^1$ and in particular second-order conditions (the Hessian); (c) the (finite) Rayleigh–Ritz method for obtaining eigenvalues; and (d) eigenvalues as comparison parameters.

Section 1.2 contains a brief introduction to the calculus of variations. Of special interest is the second variation functional or the stationary value conditions when the original functional is quadratic. The major idea is that the necessary condition for quadratic functionals leads to a self-adjoint differential system of equations. Some interesting examples are given.

In Section 1.3 we explore the fundamental tool in our work, i.e., integration by parts. We show that these ideas can be put on a sound mathematical basis. Of special note is the use of multiplier rules.

In Section 1.4 we explore briefly the relationship between quadratic forms and differential equations. In particular, two indices of quadratic forms are introduced and their relationship with solutions of differential equations with boundary value problems are given for many interesting problem settings. This section also contains many examples that the reader should find helpful. We note that these indices correspond to the number of negative and zero eigenvalues of a real symmetric (possibly infinite) matrix.

Our initial idea was to include a section on the Aronszajn–Weinstein theory of eigenvalues for compact operators since classically these ideas provide one of the most beautiful uses of Hilbert space theory. However, with the use of a computer we have developed numerical algorithms (Chapters 3 and 5) that surpass the computational algorithms of those classical methods in speed, accuracy, and feasibility. The interested reader may consult Gould [12] for the best explanation of these methods.

1.1 The Finite-Dimensional Case

In this section we treat four topics. Our criteria of which topics to include and of the degree of each topic were based on a personal judgment, based upon interesting ideas and what we feel is needed to understand quadratic form theory and the remainder of this book. Whenever possible we shall avoid technical details, results, and settings and use an expository style. The first topic deals with the duality between real finite-dimensional quadratic forms on $\mathbb{R}^n$ and real symmetric matrices. The second topic deals with optimal or stationary conditions of a function $f:\mathbb{R}^n \to \mathbb{R}^1$ and second-order necessary conditions involving the symmetric matrix A (the Hessian) with elements $a_{ij} = \partial^2 f/\partial x_i \partial x_j$ evaluated at a stationary point. The third topic is the (finite) Rayleigh–Ritz method for obtaining the eigenvalues of a real symmetric matrix. Our fourth and last topic is the concept of eigenvalues as companion parameters between a real symmetric matrix A and the identity, or more generally another real symmetric matrix B. We have also added some ideas on Lagrange multipliers for yet another view of eigenvalue theory and extremal problems. In fact, as we shall indicate in subsequent sections and chapters, this is often the correct, more practical view of eigenvalues.

We begin the first topic by assuming that $\mathscr{H}$ is a finite-dimensional, real inner product space and $Q(x)$ is a quadratic form defined on $\mathscr{H}$. The remainder of the book will be concerned with extending these concepts, along with the "meaning" of nonpositive eigenvalues, to infinite-dimensional quadratic forms $Q(x)$ and Hilbert spaces $\mathscr{H}$. Our model of $\mathscr{H}$ in dimension n is usually $\mathbb{R}^n$ and of $Q(x)$ is $x^{\mathrm{T}}Ax = (Ax, x)$, where A is an $n \times n$ real symmetric matrix, x an n vector, and x^{T} the transpose of x. For completeness, we include some topics involving background material in the next few paragraphs.

We assume that the reader is familiar with the definition of $(\mathscr{H}, \mathbb{R}^1, +, \cdot)$ as a real vector space, $\mathscr{S}$ a subspace of $\mathscr{H}$, linear combinations, linear independence and linear dependence, span, and basis. If $\mathscr{H}$ is a vector space, an *inner product on* $\mathscr{H}$ is a function $(\cdot,\cdot)\colon \mathscr{H} \times \mathscr{H} \to \mathbb{R}^1$ such that if x, y, z in $\mathscr{H}$ and c in $\mathbb{R}^1$, then $(x + y, z) = (x,z) + (y,z)$, $(cx, y) = c(x, y)$, $(x, y) = (y, x)$, and $(x,x) \geq 0$ with equality if and only if $x = 0$ in $\mathscr{H}$. The usual example is $(x, y) = y^{\mathrm{T}}x = a_j b_j$, where repeated indices are summed, $x = (a_1, \ldots, a_n)^{\mathrm{T}}$, and $y = (b_1, \ldots, b_n)^{\mathrm{T}}$ in $\mathbb{R}^n$. The *norm* of x is a function $\|\cdot\|\colon \mathscr{H} \to \mathbb{R}^1$ such that $\|x\|$ is the positive square root of (x, x).

The following ideas are found in most standard texts, for example, Hoffman and Kunze [32].

Theorem 1 *If $\mathscr{H}$ is an inner product space, then for any x, y in $\mathscr{H}$ and c in $\mathbb{R}^1$ we have*

(i) $\|cx\| = |c|\,\|x\|$,
(ii) $\|x\| \geq 0$, $\|x\| = 0$ if and only if $x = 0$,
(iii) $|(x, y)| \leq \|x\|\,\|y\|$,
(iv) $\|x + y\| \leq \|x\| + \|y\|$.

Statement (iii) is the well-known Cauchy–Schwartz inequality, and (iv) is the triangular inequality. We remark that both of these inequalities hold in the more general case of a real symmetric matrix A associated with a quadratic form $Q(x) = (Ax, x)$ if $Q(x)$ is nonnegative, i.e., $x \neq 0$ implies $Q(x) \geq 0$. The inner product is the special case with $A = I$. We shall make these concepts clearer below, but for now let $Q(x, y) = (Ax, y) = (x, Ay) = Q(y, x)$ be the bilinear form. Conditions (iii) and (iv) become, respectively,

(iii)′ $|Q(x, y)| \leq \sqrt{Q(x)}\sqrt{Q(y)}$,
(iv)′ $\sqrt{Q(x + y)} \leq \sqrt{Q(x)} + \sqrt{Q(y)}$.

Condition (iii)′ follows since for λ real,

$$\begin{aligned} 0 \leq Q(x + \lambda y) &= Q(x + \lambda y, x + \lambda y) \\ &= Q(x, x) + Q(\lambda y, x) + Q(x, \lambda y) + Q(\lambda y, \lambda y) \\ &= Q(x) + 2\lambda Q(x, y) + \lambda^2 Q(y). \end{aligned}$$

If $x = 0$ or $y = 0$, we have equality in (iii)′. The fact that the quadratic function of λ has no roots or one double root implies the discriminant "$B^2 - 4AC$" of the quadratic formula is nonpositive; otherwise we would obtain two real values of λ, and hence $f(\lambda) = \lambda^2 Q(y) + \lambda[2Q(x, y)] + Q(x)$ is negative for some $\lambda = \lambda_0$. Thus $B^2 - 4AC = 4Q^2(x, y) - 4Q(x)Q(y) \leq 0$, and hence $Q^2(x, y) \leq Q(x)Q(y)$. If $Q(x) > 0$, equality holds if and only if $x + \lambda y = 0$.

For (iv)′, $Q(x+y) = Q(x) + 2Q(x,y) + Q(y) \leq Q(x) + 2|Q(x,y)| + Q(y) \leq Q(x) + 2\sqrt{Q(x)}\sqrt{Q(y)} + Q(y) = (\sqrt{Q(x)} + \sqrt{Q(y)})^2$. Since $Q(x+y) \geq 0$ we may take square roots of both sides to obtain (iv)′. By "dropping the Q" we obtain the usual proofs of (iii) and (iv) in Theorem 1.

The vector x is *orthogonal* to y if $(x, y) = 0$. The vector x is *orthogonal* to $\mathscr{S}$ (a subset of $\mathscr{H}$) if $(x, y) = 0$ for all y in $\mathscr{S}$. $\mathscr{S}$ is an *orthogonal* set if $(x, y) = 0$ for all $x \neq y$ *in* $\mathscr{S}$. $\mathscr{S}$ is an *orthonormal set* if $\mathscr{S}$ is an orthogonal set and $\|x\| = 1$ for all x in $\mathscr{S}$. The Gram–Schmidt orthogonalization process provides that if $\{x_1, x_2, \ldots, x_n\}$ are n linearly independent vectors, there exists an orthonormal set of vectors $\{y_1, y_2, \ldots, y_n\}$ such that $\text{span}\{x_1, \ldots, x_k) = \text{span}\{y, \ldots, y_k\}$, where $1 \leq k \leq n$. The vectors $\{y_k\}$ are defined inductively by $y_1 = x_1/\|x_1\|$ and $y_{m+1} = z_{m+1}/\|z_{m+1}\|$, where (assuming $y_2, \ldots, y_m$ have been found)

$$z_{m+1} = x_{m+1} - \sum_{k=1}^{m} (x_{m+1}, y_k) y_k. \tag{1}$$

In fact, z_m is the solution to the projection or best approximation problem illustrated by Fig. 1.

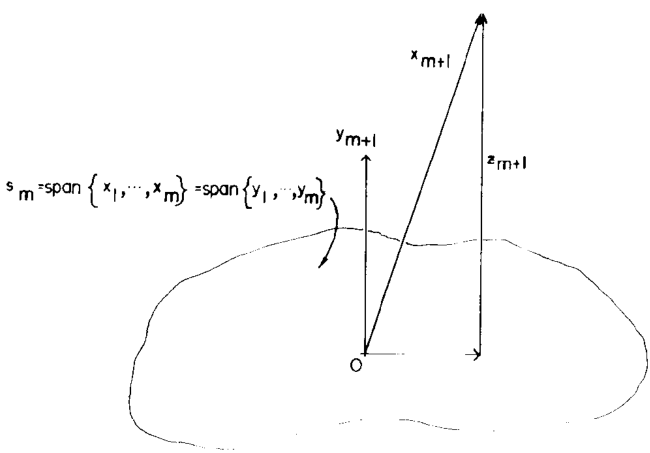


Fig. 1

If $\mathscr{H}$ is a vector space over $\mathbb{R}^1$, then $L: \mathscr{H} \to \mathscr{H}$ is a linear operator if x, y in $\mathscr{H}$ and c in $\mathbb{R}^1$ imply $L(cx + y) = cL(x) + L(y)$. It is well known (but bothersome to state precisely) that there is an *isomorphism* between the set of linear operators $L(\mathscr{H})$ and the set $\mathscr{M}_{n \times n}$ of $n \times n$ matrices, where $\mathscr{H}$ is an n-dimensional vector space.

However, before we move on, let us illustrate the above definitions and concepts by assuming $\mathscr{H} = \{x(t) = a_0 + a_1 t + a_2 t^2 + a_3 t^3\}$, where a_k in

$\mathbb{R}^1$ ($k = 0, 1, 2, 3$) with $L = D$, the derivative operator. Choosing the standard basis $\{1, t, t^2, t^3\}$ of $\mathscr{H}$ with coordinates $\bar{a} = (a_0, a_1, a_2, a_3)^{\mathrm{T}}$ in $\mathbb{R}^4$, we note that

$$\begin{aligned} D(1) &= 0 = 01 + 0t + 0t^2 + 0t^3, \\ D(t) &= 1 = 11 + 0t + 0t^2 + 0t^3, \\ D(t^2) &= 2t = 01 + 2t + 0t^2 + 0t^3, \\ D(t^3) &= 3t^2 = 01 + 0t + 3t^2 + 0t^3. \end{aligned}$$

Thus D, is identified with the 4×4 matrix

$$M = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

while $D(-3x + 4x^2) = -3 + 8x$ since

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ -3 \\ 4 \\ 0 \end{pmatrix} = \begin{pmatrix} -3 \\ 8 \\ 0 \\ 0 \end{pmatrix}$$

Note that $D(t^k)$ determine the components of the columns of M.

Similarly, this four-dimensional space becomes an inner product space if we define either

$$(x, y)_1 = a_0 b_0 + a_1 b_1 + a_2 b_2 + a_3 b_3$$

or

$$(x, y)_2 = \int_{-1}^{1} p(t) x(t) y(t)\, dt,$$

where

$$\begin{aligned} x(t) &= a_0 + a_1 t + a_2 t^2 + a_3 t^3 \\ y(t) &= b_0 + b_1 t + b_2 t^2 + b_3 t^3, \end{aligned}$$

and $p(t) > 0$ and integrable. Note that if $x(t) = t$ and $y(t) = t^3$, then $x(t)$ is orthogonal to $y(t)$ using $(\ ,\ )_1$, but $x(t)$ is not orthogonal to $y(t)$ using $(\ ,\ )_2$ since in this case with $p(t) = 1$ for example,

$$(x, y)_2 = \int_{-1}^{1} (t)(t^3)\, dt = \tfrac{1}{5} t^5 \Big|_{-1}^{1} = \tfrac{2}{5} \neq 0.$$

Similarly $\|x\|_1 = \sqrt{1 \cdot 1} = 1$, *while* $\|x\|_2 = \sqrt{\tfrac{2}{3}}$ since

$$\|x\|_2^2 = (x, x)_2 = \int_{-1}^{1} t^2 dt = \tfrac{1}{3} t^3 \Big|_{-1}^{1} = \tfrac{2}{3}.$$