Contents

1 INFORMATION ON THE PRACTICAL SESSIONS

1.1 Structure of Course

The following introductory course on Computational Linear Algebra (Computer Practical Sessions) comprises 10% of the AMTH246 unit on Mathematical Methods.

There are **8 compulsory** computer practicals (Pracs 1-8).
The material in these practicals is **examinable**.

Following these there are **3 optional** computer practicals, using **DERIVE**, which you may wish to complete. The material in these 3 practicals is **not examinable**. The practical work is assessed through two assignments, which can be found at the end of this booklet.

1.2 Software

For the practicals in this course, you need to purchase *The Student Edition* of **MATLAB** [1] (book and disks). There are also a number of freeware Matlab clones available. **SciLab** is used widely in our on campus computing labs, it is probably the best of these freeware clones and is available from

`http://www-rocq.inria.fr/scilab/`

The freeware clones do not have the copious documentation and theoretical background that accompanies the Matlab software – you need to be reasonably self-sufficient.

- It is each student's responsibility to get the required software and install it in time to begin the course.
- If for some reason you cannot get **MATLAB** (or clone) installed in time to begin the practicals, please contact Dr Tim Dalby.
- (Optional) **DERIVE**[2] is available through the URL

  `http://education.ti.com/us/product/software/derive/down/download.html`

[1] **MATLAB** is a trademark of **The MathWorks Inc.**
[2] **DERIVE** is a trademark of Texas Instruments Inc.

1.3 Lecturer

Lecturer: Dr Tim Dalby (02) 6773 2279

Email messages may be sent to: timd@turing.une.edu.au

Departmental Fax: (02) 6773 3312

1.4 References

"Elementary Linear Algebra" (7th edition) Howard Anton
(John Wiley & Sons, Inc)

2 COMPUTATIONAL LINEAR ALGEBRA AN INTRODUCTION

In previous units (e.g. MATH111), you have been introduced to some linear algebra, largely through the use of matrices. Although algebra and calculus may appear at first to be separate branches of mathematics, in many situations, they are simply different ways of viewing the same mathematical problem. The set of continuous functions of the real numbers can be viewed as elements of a vector space, and thereby analysed algebraically.

The main purpose of this course is to elaborate upon this theme, and to make use of **MATLAB** to perform some calculations based upon it.

We demonstrate the relationship between continuous functions and their differentiation on the one hand, and the representation of a function as a vector, with differentiation represented as a change of basis (matrix multiplication) on the other.

It goes without saying that computers are essential tools in modern applied mathematics. The practical exercises introduce you to **MATLAB** and **DERIVE** (optional), both of which are extremely useful in assisting with calculations and which have excellent plotting facilities for displaying results visually.

One of the most convenient aspects of **MATLAB** is that all variables are assumed to be multidimensional, saving us the work of having to specify that a variable is a vector or matrix in each situation. As long as the dimensions match, **MATLAB** will perform operations on matrices without us having to write a lot of code along the lines of "`for i = 1,n do ...`".

Another convenient aspect of **MATLAB** is its attached packages, more details of which can be found through **MATLAB**'s WWW adress (see manual).

For more advanced use, **MATLAB** can be interfaced with other software. Libraries of programs exist so that we don't have to "re-invent the wheel" every time we need to program a mathematical process.

DERIVE is very useful as a calculator. It has been designed to handle mathematically oriented problems, leaving all answers in exact form e.g. π, e unless a decimal approximation is specifically requested. It handles vector calculus quite well, and is user-friendly.

In section 3, we give some theoretical background into vector space theory. We include some examples that can be calculated without the use of a computer in this section.

Section 4 consists of the computer practicals required for this unit.

Section 5 contains some optional practicals using **DERIVE**. Some of the calculational work of Vector Calculus can be done using software such as **DERIVE**. You may find that this package is helpful in checking your calculus assignment answers. However please note that: **We require you to demonstrate a thorough understanding of the calculus component and, as such, computer output will not be accepted as solutions to the calculus asssignments. You must show all working for your calculations.**

Sections 6 and 7 consist of two assignments to be posted by the due dates shown. These assignments contribute 5% **each** to your total mark for the semester.

3 MATHEMATICAL BACKGROUND

The aim of this section is to provide a background to your work in the practical sessions. These notes contain the formal mathematical definitions and theorems on which the practicals are based – this material is not examinable, but you will require some understanding of it to get through your practical sessions.

Don't be discouraged if the abstract presentation of this background material is unfamiliar to you. Gloss through the section relevant to the practical. Attempt the practical and then reread the background material. In this way, you will see how the abstract ideas work in concrete examples.

3.1 Vector Spaces

The idea of an abstract vector space is a natural generalisation of the intuitive ideas of vectors in the plane and 3–space. However when freed from its naive setting of "arrows in space" the vector space becomes one of the most widely used concepts in mathematics. It is at the foundation of such diverse areas of mathematics as differential geometry, numerical analysis and functional analysis.

In general a vector space is defined "over a field F" (the scalars and their associated rules for addition and multiplication.) In these notes the field F will always denote the field of real numbers or the field of complex numbers.

Definition 3.1.1 *A* vector space *over a field F is a system $\{V, F, +, \ \}$ such that*

(a) $\{F, +, \ \}$ is a field with "addition" denoted by $+$ and "multiplication" denoted by juxtaposition (i.e. for a, b in F, ab denotes a multiplied by b.) The identity elements are denoted 1 and 0 (for us F is $\mathbb{R}$ or $\mathbb{C}$ so these symbols all have their usual meaning.)

(b) V is a set whose elements are called vectors which we will denote by boldface letters such as $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}$ and $\boldsymbol{x}$. There are two operations defined on V and F: vector addition $(+)$ and multiplication by a scalar denoted by juxtaposition i.e. for a in F and $\boldsymbol{u}$ in V, $a\boldsymbol{u}$) is their scalar product.

These operations must satisfy the following ten axioms

1. *For any $\boldsymbol{u}$ and $\boldsymbol{v}$ in V, $\boldsymbol{u} + \boldsymbol{v}$ is an element of V (closure under vector addition).*
2. *For any $\boldsymbol{u}, \boldsymbol{v}$ and $\boldsymbol{w}$ in V, $\boldsymbol{u} + (\boldsymbol{v} + \boldsymbol{w}) = (\boldsymbol{u} + \boldsymbol{v}) + \boldsymbol{w}$ (associativity of vector addition)*

3. *For any* $\boldsymbol{u}$ *and* $\boldsymbol{v}$ *in* V, $\boldsymbol{u}+\boldsymbol{v}=\boldsymbol{v}+\boldsymbol{u}$ *(vector addition is commutative)*
4. *There is a unique vector* $\mathbf{0}$ *in* V, *called the zero vector, such that* $\boldsymbol{u}+\mathbf{0}=\boldsymbol{u}$ *for all* $\boldsymbol{u}$ *in* V *(additive identity)*
5. *For each* $\boldsymbol{u}$ *in* V *there is a unique vector,* $-\boldsymbol{u}$, *in* V *such that* $\boldsymbol{u}+-\boldsymbol{u}=0$ *(additive inverse)*
6. *For any* $\boldsymbol{u}$ *in* V *and any* k *in* F $k\boldsymbol{u}$ *is in* V *(closure under scalar multiplication)*
7. *For any* $\boldsymbol{u}$ *in* V *and any* h,k *in* F, $(h+k)\boldsymbol{u}=h\boldsymbol{u}+k\boldsymbol{u}$ — *distributive laws*
8. *For any* $\boldsymbol{u}$ *and* $\boldsymbol{v}$ *in* V *and any* k *in* F, $k(\boldsymbol{u}+\boldsymbol{v})=k\boldsymbol{u}+k\boldsymbol{v}$ — *distributive laws*
9. *For any* h,k *in* F *and any* $\boldsymbol{u}$ *in* V, $h(k\boldsymbol{u})=(hk)\boldsymbol{u}$
10. *For any* $\boldsymbol{u}$ *in* V, $1\boldsymbol{u}=\boldsymbol{u}$, *where 1 is the multiplicative identity in* F.

Theorem 3.1.1 *For all* $\boldsymbol{u}$ *in* V *and* k *in* F

(a) $0\boldsymbol{u}=\mathbf{0}$

(b) $(-1)\boldsymbol{u}=-\boldsymbol{u}$

(c) $k\mathbf{0}=\mathbf{0}$

Examples of Vector Spaces

1. For fixed integer n the set of all n-tuples of numbers from F (i.e. real numbers or complex numbers) is a vector space over F with the following definitions

$$(x_1,x_2,\cdots,x_n)+(y_1,y_2,\cdots,y_n)=(x_1+y_1,x_2+y_2,\cdots,x_n+y_n)$$

and $k(x_1,x_2,\cdots,x_n)=(kx_1,kx_2,\cdots,kx_n)$, for $k,x_1,x_2,\cdots,x_n$ in F.

2. For fixed integer n let $P_n(F)$ denote the set of all polynomials in x with coefficients in F of degree not exceeding n (and including the zero polynomial). Then $P_n(F)$ is a vector space over F with the usual rules for addition of polynomials ("collecting like terms") and multiplication of polynomials by scalars from F (i.e. $\mathbb{R}$ or $\mathbb{C}$). In fact, we can simply represent a polynomial of degree n by an $(n+1)-$tuple of its coefficients:

$$a_0+a_1x+a_2x^2+\cdots+a_nx^n \text{ is represented as}$$
$$(a_0,a_1,a_2,\cdots,a_n)$$

The vector space $P_n(F)$ is then *isomorphic* to that of example 1, above.

3. Let V be the set of all real valued differentiable functions on $[0, 1]$, with operations defined by
$$(f+g)(x) = f(x) + g(x)$$
$$(kf)(x) = kf(x), k \text{ in } \mathbb{R}.$$

4. $\mathbb{R}^{m,n}$ the set of all real m by n matrices is a vector space over $\mathbb{R}$.

5. $\mathbb{R}^n$ is a real vector space (this is our motivating example).

Definition 3.1.2 *Let $\{V, f, +, \}$ be a vector space. A subset of vectors S, of V, is said to form a* subspace *$\{S, F, +, \}$ if and only if $\{S, F, +, \}$ is a vector space.*

Notice that most of the axioms of a vector space are inherited by S, in fact only the two closure axioms need to be verified for the system $\{S, F, + \}$.

Theorem 3.1.2 *A nonempty subset of vectors S forms a subspace $\{S, F, +, \}$ of the vector space $\{V, F, +, \}$ if and only if S is closed under the two operations of vector sum and multiplication of vectors by scalars (axioms 1 and 6, above).*

Examples of Subspaces:

1. Let S be the set of all n tuples (of example 1, above) with $x_1 = 0$.

2. Let m and n be positive integers with $m \leq n$, then $P_m(F)$ is a subspace of $P_n(F)$.

3. Let S be the set of all constant functions on $[0, 1]$. S is a subspace of example 3, above.

A general means of generating subspaces is that of forming linear combinations of vectors from a nonempty subset.

Definition 3.1.3 *Let S be a nonempty subset of a vector space V (we will abuse notation from now on and simply refer to the vector space V rather than $\{V, F, +\}$). The* span *of S, denoted $< S >$, is the set of all linear combinations of vectors in S, i.e. the collection of all finite sums of the form $k_1\boldsymbol{u}_1 + k_2\boldsymbol{u}_2 + \cdots + k_m\mathbf{u}_m$ with the k_i in F and the $\boldsymbol{u}_i$ in S, $m = 1, 2, \cdots$.*

As anticipated, above, we have

Theorem 3.1.3 *If S is a nonempty subset of vectors from a vector space V, then $< S >$ is a vector subspace of V.*

3.2 Linear Independence and Bases

Definition 3.2.1 *Let V be a vector space over F and let $S = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_m\}$ be a finite subset of vectors from V. S is said to be* linearly independent *if and only if every equation of the form $k_1\boldsymbol{u}_1 + k_2\boldsymbol{u}_2 + \cdots + k_m\boldsymbol{u}_m = \mathbf{0}$, k_i in F, implies $k_1 = k_2 = \cdots = k_m = 0$.*

An infinite set T is said to be linearly independent if every finite subset of T is linearly independent. The empty set is linearly independent. Any set which is not independent is linearly dependent. A simple consequence of the definition is:

Theorem 3.2.1 *Any subset of an independent set is independent and any set containing a dependent subset is dependent.*

Another straightforward consequence of the definition is:

Theorem 3.2.2 *Let S be a finite set consisiting of m nonzero vectors of a vector space. Then S is linearly dependent if and only if there exists a set of $m-1$ vectors in S which span S.*

Notice that any set of vectors will contain a maximal linearly independent subset, i.e. if S is a set of vectors S contains a subset A of linearly independent vectors such that the set $\{A, \boldsymbol{u}\}$ is dependent for all $\boldsymbol{u}$ in S.

Definition 3.2.2 *A maximal linearly independent subset of a vector space V is called a* basis *of V.*

Notice that if B is a basis of V then $V = < B >$ and every vector of V can be written as a linear combination of elements of B. If $B = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_n\}$ then for any $\boldsymbol{v}$ in V there exists scalars $k_1, k_2, \cdots, k_n$ such that $\boldsymbol{v} = k_1\boldsymbol{u}_1 + k_2\boldsymbol{u}_2 + \cdots k_n\boldsymbol{u}_n$. It is easy to show that the scalars k_i are unique for the given basis. This means that a vector uniquely determines an n-tuple of scalars $(k_1, k_2, \cdots, k_n)$ which can be regarded as the coordinates of $\boldsymbol{v}$ relative to the basis B. We can also show that the number of elements in a basis is unique, this prompts the following definition.

Definition 3.2.3 *The* dimension *of a vector space V is the number of elements in any basis, it is denoted by* $\dim(V)$.

We end this subsection with three important theorems.

Theorem 3.2.3 *Any linearly independent set of vectors in an n-dimensional vector space V can be extended to a basis.*

Proof: Let $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots \boldsymbol{u}_m\}$ be an independent set and let $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n\}$ be a basis. Let $S =< \boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_m >$, then if $\mathbf{v}_i$ is in S for all $i = 1, 2, \cdots, n$ we have $S = V$. Otherwise, $\mathbf{v}_j$ is not in S for some j and $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \mathbf{u}_m, \boldsymbol{v}_j\}$ is independent. We repeat the argument until the enlarged set spans V.

Theorem 3.2.4 *Let $S = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \mathbf{u}_n\}$ be a subset of vectors of an n-dimensional vector space V. Then*

(a) S is a basis for V if and only if S is linearly independent.

(b) S is a basis for V if and only if $V =< S >$.

Proof: Exercise

Theorem 3.2.5 *Let $B = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \mathbf{u}_n\}$ and $C = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n\}$ be bases for an n-dimensional vector space V. Then there is a unique non-singular n by n matrix $A = (a_{ij})$, with elements a_{ij} in F, such that*

$$\boldsymbol{u}_i = \sum_{j=1}^{n} a_{ij}\boldsymbol{v}_j, \quad i = 1, 2, \cdots, n.$$

Proof: $\boldsymbol{u}_i$ is in V and C is a basis of V so $\mathbf{u}_i$ is a linear combination of the elements of C. Denote this as

$$\boldsymbol{u}_i = \sum_{j=1}^{n} a_{ij}\boldsymbol{v}_j$$

.

The matrix is non singular because its row vectors are linearly independent (and none of the row vectors can contain all zeros.) This matrix clearly has an inverse as the $\boldsymbol{v}_i$ can be written as linear combinations of the $\boldsymbol{u}_j$.

Finally, we note that if $\boldsymbol{w}$ is in V it can be represented as an n-tuple in many ways, depending on the choice of basis. If $(k_1, k_2, \cdots, k_n)$ is the representation with respect to the basis B and $(h_1, h_2, \cdots, h_n)$ is the representation with respect to the basis C then, with the notation of the last theorem,

$$(h_1, h_2, \cdots, h_n) = (k_1, k_2, \cdots, k_n)A$$

Example: Consider $P_2(\mathbb{R})$, the vector space of ploynomials of degree 2. A suitable basis is

$$\begin{aligned}\boldsymbol{u}_1 &= 1 \ \ (= x^0)\\ \boldsymbol{u}_2 &= x\\ \boldsymbol{u}_3 &= x^2\end{aligned}$$

Any $\boldsymbol{v}$ in $P_2(\mathbb{R})$ can be written as $\boldsymbol{v} = k_0\mathbf{u}_1 + k_1\boldsymbol{u}_2 + k_2\boldsymbol{u}_3 = k_0 + k_1x + k_2x^2$, with the k's in $\mathbb{R}$. Another basis is

$$\begin{aligned}\boldsymbol{v}_1 &= 1 = \boldsymbol{u}_1\\ \boldsymbol{v}_2 &= (x-1) = \boldsymbol{u}_2 - \boldsymbol{u}_1\\ \boldsymbol{v}_3 &= (x-1)^2 = \boldsymbol{u}_3 - 2\boldsymbol{u}_2 + \boldsymbol{u}_1, \text{ or}\\ \boldsymbol{u}_1 &= \boldsymbol{v}_1\\ \boldsymbol{u}_2 &= \boldsymbol{v}_2 + \boldsymbol{v}_1\\ \boldsymbol{u}_3 &= \boldsymbol{v}_3 + 2\boldsymbol{v}_2 + \boldsymbol{v}_1\end{aligned}$$

Consequently, the matrix A giving the change of basis from C to B (theorem 3.2-5, above) is

$$A = \begin{pmatrix} 1 & 0 & 0\\ 1 & 1 & 0\\ 1 & 2 & 1\end{pmatrix}$$

So if (a, b, c) is the representation of an element $\mathbf{p}$ of $P_2(\mathbb{R})$ with respect to B, i.e. $\mathbf{p} = a\boldsymbol{u}_1 + b\boldsymbol{u}_2 + c\boldsymbol{u}_3 = a + bx + cx^2$, then its representation with respect to C, (d, e, f), is given by

$$(d, e, f) = (a, b, c)A, \quad \text{i.e.}$$

$$\begin{aligned}d &= a + b + c\\ e &= b + 2c\\ f &= c\end{aligned}$$

We verify this

$$\begin{aligned}d + e(x-1) + f(x-1)^2 &= (d - e + f) + (e - 2f)x + fx^2\\ &= (a + b + c - b - 2c + c) + (b + 2c - 2c)x + cx^2\\ &= a + bx + cx^2.\end{aligned}$$

Exercises:

Consider the vector space $P_3(\mathbb{R})$, with standard basis

$$\begin{aligned}\boldsymbol{u}_0 &= 1\\ \boldsymbol{u}_1 &= x\\ \boldsymbol{u}_2 &= x^2\\ \boldsymbol{u}_3 &= x^3\end{aligned}$$

1. Another basis for $P_3(\mathbb{R})$

$$\begin{aligned} \boldsymbol{v}_0 &= 1 \\ \boldsymbol{v}_1 &= 1+x \\ \boldsymbol{v}_2 &= 1+x+x^2 \\ \boldsymbol{v}_3 &= 1+x+x^2+x^3 \end{aligned}$$

   (a) Verify that these elements do indeed form a basis for $P_3(\mathbb{R})$.

   (b) What is the matrix, A, giving the change of basis $\{\boldsymbol{u}_j\}_{j=0}^3 \longrightarrow \{\boldsymbol{v}_j\}_{j=0}^3$?

2. Yet another basis for $P_3(\mathbb{R})$ is

$$\begin{aligned} \boldsymbol{w}_0 &= 1 \\ \boldsymbol{w}_1 &= 1-x \\ \boldsymbol{w}_2 &= 1-x+x^2 \\ \boldsymbol{w}_3 &= 1-x+x^2-x^3 \end{aligned}$$

   (a) Find B the matrix for the change of basis $\{\boldsymbol{v}_j\}_{j=0}^3 \longrightarrow \{\boldsymbol{w}_j\}_{j=0}^3$

   (b) Find C the matrix for $\{\boldsymbol{u}_j\}_{j=0}^3 \longrightarrow \{\boldsymbol{w}_j\}_{i=0}^3$

   (c) What is the relationship between A, B and C?

3.3 Linear Transformations of Vector Spaces

Definition 3.3.1 *A* linear transformation*, T, from a vector space V to a vector space W (both over the field F) is a mapping of V into W such that for all $\boldsymbol{u}$ and $\boldsymbol{v}$ in V and all a and b in F,*

$$T(a\boldsymbol{u} + b\boldsymbol{v}) = aT(\boldsymbol{u}) + bT(\boldsymbol{v})$$

Definition 3.3.2 *The* image *of V under T is the subset of vectors in W given by*

$$Im_T(V) = \{T(\boldsymbol{v}) : \boldsymbol{v} \in V\}$$

Definition 3.3.3 *The* kernel *of T is the set of vectors in V given by*

$$Ker_T(V) = \{\boldsymbol{v} \in V : T(\boldsymbol{v}) = \boldsymbol{0}\}$$

We give the following two important theorems without proof.

Theorem 3.3.1 *Let $T : V \to W$ be a linear transformation of vector spaces V and W. Then $Ker_T(V)$ is a vector subspace of V and $Im_T(V)$ is a vector subspace of W.*

The fact that the kernel and image are vector spaces lets us make the following definitions.

Definition 3.3.4 *The* nullity *of a linear transformation T is the dimension of $Ker_T(V)$, denoted $n(T)$.*

Definition 3.3.5 *The* rank *of a linear transformation T is the dimension of $Im_T(V)$, denoted $r(T)$.*

Theorem 3.3.2 *Let $\{\boldsymbol{u}_1, \cdots, \boldsymbol{u}_{n(T)}\}$ be a basis for $Ker_T(V)$. Extend this to a basis $\{\boldsymbol{u}_1, \cdots, \boldsymbol{u}_{n(T)}, \boldsymbol{u}_{n(T)+1}, \cdots, \boldsymbol{u}_n\}$ for V (where $n = \dim V$). Then $\{T(\boldsymbol{u}_{n(T)+1}), \cdots, T(\boldsymbol{u}_n)\}$, is a basis for $ImT(V)$ and $n(T) + r(T) = n$.*

Linear transformations can be given a matrix representation. We will be interested in transformations from V to itself, i.e. $T : V \to V$. Let $\{\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n\}$ be a basis for V, a vector $\boldsymbol{u}$ in V can be written as

$$\boldsymbol{u} = \sum_{i=1}^{n} x_i \boldsymbol{u}_i$$

so that $T(\boldsymbol{u}) = \sum_{i=1}^{n} x_i T(\boldsymbol{u}_i)$. The $T(\boldsymbol{u}_i)$ will be (linearly independent) linear combinations of the $\boldsymbol{u}_i$ so we write $T(\boldsymbol{u}_i) = \sum_{j=1}^{n} t_{ij}\boldsymbol{u}_j$, for some scalars t_{ij}. Then, $\boldsymbol{u} \to T(\boldsymbol{u})$ is represented by $(x_1, \cdots, x_n) \to (x_1, \cdots, x_n)(t_{ij})$ where (t_{ij}) is the matrix with elements t_{ij}.

Examples of Linear Transformations.

1. Rotations of points in $\boldsymbol{R}^n$ about a fixed axis.
2. Matrices acting on row or column vectors.
3. Linear differential operators acting on various function spaces.

We now define the important concepts of eigenvectors and eigenvalues associated with a linear transformation.

Definition 3.3.6 *A nonzero vector* $\boldsymbol{v}$ *satisfying*

$$T(\boldsymbol{v}) = \lambda \boldsymbol{v}$$

is called an eigenvector *of* T *and* λ *is called an* eigenvalue *of* T *associated with the eigenvector* $\boldsymbol{v}$.

In our matrix representation, above, we have

$$(y_1, \cdots, y_n)(t_{ij}) = \lambda(y_1, \cdots, y_n), \text{ where } \boldsymbol{v} = \sum_{i=1}^{n} y_i \boldsymbol{u}_i$$

so that λ must satisfy the characteristic equation $\det((t_{ij}) - \lambda I) = 0$, I the n by n unit matrix. You are probably more used to the eigenvector equation in transposed form

$$(t_{ij})^t \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \lambda \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}.$$

3.4 Quadratic Forms

A quadratic form on $\mathbb{R}^n$ (or $\mathbb{C}^n$) is a homogeneous quadratic function on $\mathbb{R}^n$. The general quadratic form on $\mathbb{R}^n$ may be presented as

$$\begin{aligned} Q(x) &= \sum_{i,j=1}^{n} A_{ij} x_j x_j, \quad \text{where the } A_{ij}\text{'s are constants and } x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n. \\ &= xAx^t, \quad \text{where } A = (A_{ij}) \text{ and } {}^t \text{ denotes the transpose – a column vector here.} \end{aligned}$$

Note that A may be taken as a *symmetric matrix*.

Definition 3.4.1 *A quadratic form is said to be* non-degenerate *if* $Q(x) \neq 0$ *for all non-zero* x.

Definition 3.4.2 *A quadratic form is said to be* positive definite *if* $Q(x) > 0$ *for all non-zero* x.

Definition 3.4.3 *A quadratic form is said to be* negative definite *if* $Q(x) < 0$ *for all non-zero* x.

(The terms *positive semi-definite* and *negative semi-definite* are used when the strict inequalities are relaxed to include the possibility of equality).

A positive definite quadratic form defines an inner product on R^n:

$$< x, y > = xAy^t$$

Quadratic forms arise in many guises in mathematics, physics, economics and so on; conic sections, moments of inertia and special relativity to give but a few examples.

It can be shown that any symmetric matrix can be diagonalized i.e. that there exists a non-singular matrix S such that $D = S^{-1}AS$ is a diagonal matrix. In fact $D = diag(\lambda_1, \lambda_2, \ldots, \lambda_n)$ where $\lambda_1, \lambda_2, \ldots, \lambda_n$ are the eigenvalues of A. The matrix S is not quite unique. This arbitrariness is removed by demanding that $S^{-1} = S^t$. The reason for this condition is as follows: $D = D^t = S^t A (S^{-1})^t = S^{-1}AS$. The last equality is then identically true if we impose the condition $S^{-1} = S^t$. A matrix satisfying this last condition is called an *orthogonal* matrix. (The matrix S is the matrix whose columns are the eigenvectors of A.)

Returning to our quadratic form we have

$$\begin{aligned} Q(x) &= xAx^t \\ &= x(SDS^{-1})x^t, \quad \text{as} \quad D = S^{-1}AS \quad \text{we have} \quad A = SDS^{-1}. \\ &= x(SDS^t)x^t, \quad \text{as} \quad S^t = S^{-1}. \\ &= (xS)D(xS)^t \end{aligned}$$

So by "rotating" our axes $x \longrightarrow X = xS$ by the orthogonal transformation S the quadratic form $Q(x)$ has been diagonalized in terms of the new variables X. The new axes are determined by the rotation S, e.g. where the old x_1 axis is given by the vector $(1, 0, 0, \ldots, 0)$ the new x_1 axis is $(1, 0, 0, \ldots, 0)S$.

As $D = diag(\lambda_1, \lambda_2, \ldots, \lambda_n)$, where the λ_i are the eigenvalues of A, we have

$$Q(x) = \lambda_1(X_1)^2 + \lambda_2(X_2)^2 + \ldots + \lambda_n(X_n)^2, \quad \text{where} \quad X = xS.$$

The new axes are called the *principal axes* of the quadratic form.

Exercises.

1. Let $Q = 4x^2 + 8xy + y^2$
   (a) Find the orthogonal transformation which diagonalizes this form.
   (This is the matrix S whose columns are the eigenvectors of$A$.)
   (b) Find the principal axes.
   (These are simply the rows of S, since $X = xS$.)
   (c) Find the diagonalized form of Q.
   (i.e. Write Q in terms of the new variables X and Y - there should be no "cross-terms")
2. Rotate the conic $5x^2 - 4xy + 2y^2 = 30$ to its principal axes and so identify it.
3. Rotate to principal axes the quadric surface $x^2 + 6xy - 2y^2 - 2yz + z^2 = 24$.

4 PRACTICAL SESSIONS (COMPULSORY)

4.1 Practical 1 – Vector Bases Using MATLAB

A common vector space in many applications is the vector space $P_n(\mathbb{R})$ of polynomials of degree at most n. Since any polynomial $p(x)$ of degree at most n can be written as linear combination

$$p(x) = a_0 \cdot 1 + a_1 \cdot x + a_2 \cdot x^2 + \cdots + a_n \cdot x^n$$

of the set of polynomials

$$\begin{aligned} \boldsymbol{u}_0(x) &= 1 \\ \boldsymbol{u}_1(x) &= x \\ \boldsymbol{u}_2(x) &= x^2 \\ \boldsymbol{u}_3(x) &= x^3 \\ &\vdots \\ \boldsymbol{u}_n(x) &= x^n \end{aligned}$$

it follows that the set of polynomials $\boldsymbol{u}_i(x), i = 0, \cdots, n$ is a basis for the vector space $\mathcal{P}_n$ and that the vector space has dimension $n+1$. The basis $\boldsymbol{u}_i(x) = x^i, i = 0, \cdots, n$ will be referred to as the *standard basis* for the vector space $P_n(\mathbb{R})$.

There are other bases for $P_n(\mathbb{R})$ which occur in applications. One of these is the set of *Chebyshev polynomials* $T_n(x)$ defined by the recurrence relation:

$$\begin{aligned} T_0(x) &= 1 \\ T_1(x) &= x \\ T_n(x) &= 2xT_{n-1}(x) - T_{n-2}(x). \end{aligned}$$

The next few terms of the sequence are

$$\begin{aligned} T_2(x) &= 2x^2 - 1 \\ T_3(x) &= 4x^3 - 3x \\ T_4(x) &= 8x^4 - 8x^2 + 1 \end{aligned}$$

We will work within the vector space $P_4(\mathbb{R})$. The Chebyshev polynomials $T_0(x), T_1(x), \cdots, T_4(x)$ can be expressed in terms of the standard basis $\boldsymbol{u}_i(x), i = 0, \cdots, 4$:

$$\begin{aligned} T_0(x) &= \boldsymbol{u}_0(x) \\ T_1(x) &= \boldsymbol{u}_1(x) \\ T_2(x) &= 2\boldsymbol{u}_2(x) - \boldsymbol{u}_0(x) \\ T_3(x) &= 4\boldsymbol{u}_3(x) - 3\boldsymbol{u}_1(x) \\ T_4(x) &= 8\boldsymbol{u}_4(x) - 8\boldsymbol{u}_2(x) + \boldsymbol{u}_0(x). \end{aligned}$$

The matrix A giving the change of basis from the standard basis to the Chebyshev basis is

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

This matrix can be entered in **MATLAB** as

```
A = [1 0 0 0 0; 0 1 0 0 0; -1 0 2 0 0; 0 -3 0 4 0; 1 0 -8 0 8]
```

Or as

```
A = [1  0  0 0 0
     0  1  0 0 0
    -1  0  2 0 0
     0 -3  0 4 0
     1  0 -8 0 8]
```

The inverse of the matrix A will allow us to express the standard basis in terms of the Chebyshev basis. The **MATLAB** (or **scilab**) command

```
B = inv(A)
```

will give us this inverse. The rows of the matrix B are the elements of the standard basis expressed as linear combinations of the Chebyshev polynomials. For example, the fourth row gives us

$$\boldsymbol{u}_3 = \frac{3}{4}T_1 + \frac{1}{4}T_3.$$

The matrices A and B can be used to convert vectors (= polynomials) back and forth between the standard and Chebyshev bases.

Example

Express the polynomial

$$p(x) = x^4 - 3x + 5$$

as a linear combination of Chebyshev polynomials.

The **MATLAB** calculation is as follows:

```
p = [5 -3 0 0 1]
p1 = p*B
```

giving

$$p(x) = \frac{1}{8}T_4(x) + \frac{1}{2}T_2(x) - 3T_1(x) + 5\frac{3}{8}T_0(x).$$

Problem

Another important family of polynomials are the Hermite polynomials defined by

$$\begin{aligned} H_0(x) &= 1 \\ H_1(x) &= 2x \\ H_n(x) &= 2xH_{n-1}(x) - 2(n-1)H_{n-2}(x), \quad n = 2, 3... \end{aligned}$$

(a) Write down $H_2(x)$, $H_3(x)$ and $H_4(x)$.

(b) Write down the matrix expressing the first five Hermite polynomials as linear combinations of the standard basis for $P_n(\mathbb{R})$.

The remainder of the problem should be done using **MATLAB**.

(c) Find the matrix for transforming the standard basis to Hermite polynomials.

(d) Using the matrices for changing between the standard and Chebyshev bases, calculate the matrices for changing between Hermite and Chebyshev bases.

(e) Express

$$2.37H_3(x) - 1.21H_1(x) + 0.13H_0(x)$$

in terms of Chebyshev polynomials and

$$-3.78T_4(x) + 5.02T_3(x)$$

in terms of Hermite polynomials.

4.2 Practical 2 – Inner Products

Definition: If $\boldsymbol{u}$ and $\boldsymbol{v}$ are vectors in a vector space V then $(\boldsymbol{u}, \boldsymbol{v})$ is called an *inner product* (also called a scalar or dot product) if $(.,.)$ is a mapping $V \times V \to F$ with the following properties

1. $(\boldsymbol{u}, \boldsymbol{v}) = \overline{(\boldsymbol{v}, \boldsymbol{u})}$, where the overbar denotes complex conjugate.
2. $(k_1\boldsymbol{u} + k_2\boldsymbol{v}, \boldsymbol{w}) = k_1(\boldsymbol{u}, \boldsymbol{w}) + k_2(\boldsymbol{v}, \boldsymbol{w})$, for any $k_1, k_2 \in F$ and $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in V$.
3. $(\boldsymbol{u}, \boldsymbol{u}) > 0$ if $\boldsymbol{u} \neq 0$, $(\boldsymbol{u}, \boldsymbol{u}) = 0$ if and only if $\boldsymbol{u} = \boldsymbol{0}$.

   A vector space with an inner product is called an *inner product space.*

Examples of Inner Product Spaces

1. In $\mathbb{R}^n$ with $\boldsymbol{u} = (x_1, \cdots, x_n)$ and $\boldsymbol{v} = (y_1, \cdots, y_n)$ the Euclidean Inner Product is
$$(\boldsymbol{u}, \boldsymbol{v}) = \sum_{i=1}^{n} x_i y_i.$$
   (This is just the usual dot product when we are dealing with $\mathbb{R}^3$.)
2. Important examples (on which this practical will be based) are obtained from the vector spaces of polynomials $P_n(F)$. We consider $P_n(\mathbb{R})$ on the interval $(-1, 1)$. Then, if $w = w(x)$ is an integrable function on $(-1, 1)$ we define an inner product by
$$(\boldsymbol{u}, \boldsymbol{v}) = \int_{-1}^{1} u(x)v(x)w(x)dx, \text{ where on the right } \boldsymbol{u} \text{ and } \boldsymbol{v}$$
   are considered to be functions of x.
3. We could use the inner product of example 2 on a larger space than $P_n(\mathbb{R})$, for example the vector space of integrable functions on $(-1, 1)$. There are many other interesting inner products one can define on such function spaces, for example the vector space of C^1 functions on $(-1, 1)$ can be given the inner product
$$(\boldsymbol{u}, \boldsymbol{v}) = \int_{-1}^{1} [u(x)v(x) + u'(x)v'(x)]dx$$

Definition: A vector space V is called a *normed vector space* if for each $\boldsymbol{u}$ in V there is a function $\|.\|$ with the properties

1. $\|\boldsymbol{u}\| \geq 0$ if $\boldsymbol{u} \neq 0$ and $\|\boldsymbol{u}\| = 0$ implies $\boldsymbol{u} = \boldsymbol{0}$
2. $\|k\boldsymbol{u}\| = |k|\|\boldsymbol{u}\|$

3. $\|\boldsymbol{u}+\boldsymbol{v}\| \leq \|\boldsymbol{u}\| + \|\boldsymbol{v}\|$ (triangle inequality.)

An inner product space is always a normed vector space we define the induced norm by $\|\boldsymbol{u}\| = (\boldsymbol{u}, \boldsymbol{u})^{1/2}$ (in $\mathbb{R}^3$ this will just give us the usual length or magnitude of a vector).

In this weeks practical we will be concerned with producing sets of polynomials which form an orthogonal set of basis vectors for $P_4(\mathbb{R})$. Such sets are extremely important in applications. These sets are produced using a general procedure which may be applied to any inner product space.

The **Gram-Schmidt Orthogonalization Procedure** is an inductive technique to generate a mutually orthogonal set from any linearly independent set of vectors. Given the linearly independent set of vectors $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$, we set $\boldsymbol{v}_1 = \boldsymbol{u}_1$. If we take $\boldsymbol{v}_2 = \boldsymbol{u}_2 - k\boldsymbol{v}_1$, the requirement that $(\boldsymbol{v}_1, \boldsymbol{v}_2) = 0$ determines k as

$$k = (\boldsymbol{u}_2, \boldsymbol{v}_1)/\|\boldsymbol{v}_1\|^2$$

so

$$\boldsymbol{v}_2 = \boldsymbol{u}_2 - \{(\boldsymbol{u}_2, \boldsymbol{v}_1)/\|\boldsymbol{v}_1\|^2\}\boldsymbol{v}_1.$$

In other words, we get $\boldsymbol{v}_2$ by subtracting from $\boldsymbol{u}_2$ the projection of $\boldsymbol{u}_2$ onto $\boldsymbol{v}_1$. Proceeding inductively, we find

$$\boldsymbol{v}_n = \boldsymbol{u}_n - \sum_{j=1}^{n-1} \{(\boldsymbol{u}_n, \boldsymbol{v}_j)/\|\boldsymbol{v}_j\|^2\}\boldsymbol{v}_j$$

from which it is clear that $(\boldsymbol{v}_i, \boldsymbol{v}_j) = 0$ for all $i, j = 1, 2, \cdots, n$ with $i \neq j$.
(That is, the new set of vectors are mutually orthogonal.)

Let us apply the Gram-Schmidt procedure to $P_4(\mathbb{R})$ with the inner product

$$(\boldsymbol{u}, \boldsymbol{v}) = \int_{-1}^{1} u(x)v(x)dx \tag{1}$$

If we write $\boldsymbol{u} = x_0\boldsymbol{u}_0 + \cdots + x_4\boldsymbol{u}_4$ and $\boldsymbol{v} = y_0\boldsymbol{u}_0 + \cdots + y_4\boldsymbol{u}_4$ (where $\boldsymbol{u}_0, \boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3, \boldsymbol{u}_4$ is the standard basis) then

$$(\boldsymbol{u}, \boldsymbol{v}) = (x_0, \cdots, x_4)A(y_0, \cdots, y_4)^t \quad \text{(where } t \text{ is the transpose)}$$

where A is the 5 by 5 matrix:

$$\begin{bmatrix} \int_{-1}^{1} 1dx & \int_{-1}^{1} xdx & \int_{-1}^{1} x^2dx & \int_{-1}^{1} x^3dx & \int_{-1}^{1} x^4dx \\ \int_{-1}^{1} xdx & \int_{-1}^{1} x^2dx & \int_{-1}^{1} x^3dx & \int_{-1}^{1} x^4dx & \int_{-1}^{1} x^5dx \\ \int_{-1}^{1} x^2dx & \int_{-1}^{1} x^3dx & \int_{-1}^{1} x^4dx & \int_{-1}^{1} x^5dx & \int_{-1}^{1} x^6dx \\ \int_{-1}^{1} x^3dx & \int_{-1}^{1} x^4dx & \int_{-1}^{1} x^5dx & \int_{-1}^{1} x^6dx & \int_{-1}^{1} x^7dx \\ \int_{-1}^{1} x^4dx & \int_{-1}^{1} x^5dx & \int_{-1}^{1} x^6dx & \int_{-1}^{1} x^7dx & \int_{-1}^{1} x^8dx \end{bmatrix}$$

which equals

$$\begin{bmatrix} 2 & 0 & 2/3 & 0 & 2/5 \\ 0 & 2/3 & 0 & 2/5 & 0 \\ 2/3 & 0 & 2/5 & 0 & 2/7 \\ 0 & 2/5 & 0 & 2/7 & 0 \\ 2/5 & 0 & 2/7 & 0 & 2/9 \end{bmatrix}.$$

(Make sure that you can derive this from the definition of the inner product (1).)

You can now use **MATLAB** to calculate the orthogonal basis given by the Gram-Schmidt procedure.

$$\begin{aligned}
\boldsymbol{v}_0 &= \boldsymbol{u}_0 \\
\boldsymbol{v}_1 &= \boldsymbol{u}_1 - \{(\boldsymbol{u}_1, \boldsymbol{v}_0)/(\boldsymbol{v}_0, \boldsymbol{v}_0)\}\boldsymbol{v}_0 \\
&= \boldsymbol{u}_1 = x, \text{ check this!} \\
\boldsymbol{v}_2 &= \boldsymbol{u}_2 - \{(\boldsymbol{u}_2, \boldsymbol{v}_0)/(\boldsymbol{v}_0, \boldsymbol{v}_0)\}\boldsymbol{v}_0 - \{(\boldsymbol{u}_2, \boldsymbol{v}_1)/(\boldsymbol{v}_1, \boldsymbol{v}_1)\}\boldsymbol{v}_1 \\
&= x^2 - 1/3, \text{ check this!} \\
\boldsymbol{v}_3 &= \boldsymbol{u}_3 - \{(\boldsymbol{u}_3, \boldsymbol{v}_0)/(\boldsymbol{v}_0, \boldsymbol{v}_0)\}\boldsymbol{v}_0 - \{(\boldsymbol{u}_3, \boldsymbol{v}_1)/(\boldsymbol{v}_1, \boldsymbol{v}_1)\}\boldsymbol{v}_1 \\
&\quad -\{(\boldsymbol{u}_3, \boldsymbol{v}_2)/(\boldsymbol{v}_2, \boldsymbol{v}_2)\}\boldsymbol{v}_2 \\
&= x^3 - 3x/5, \text{ check this!} \\
\boldsymbol{v}_4 &= \ ?
\end{aligned}$$

To do the calculation first enter $\boldsymbol{u}_0$ to $\boldsymbol{u}_4$ as row vectors and enter the matrix A into **MATLAB** i.e.

```
u0 = (1,0,0,0,0)
u1 = (0,1,0,0,0)
u2 =
u3 =
u4 =
```

Then we set

```
v0 =  u0
v1 =  u1 - ((u1*A*v0')/(v0*A*v0'))*v0
v2 =
v3 =
v4 =
```

[Note: $'$ is the **MATLAB** transpose command.]

Now write down the associated polynomials. These polynomials are called Legendre polynomials, they occur frequently in solutions to problems in mathematical physics.

Exercise: Obtain the orthogonal polynomials associated with the inner product

$$(\boldsymbol{u}, \boldsymbol{v}) = \frac{8}{\pi} \int_{-1}^{1} u(x)v(x)w(x)dx, \text{ where } w(x) = 1/(1-x^2)^{1/2}.$$

If you have **DERIVE** you can use it to calculate the integrals required for the matrix A.

- Do you recognize the polynomials? They are essentially the Chebyshev polynomials of Prac 1.

4.3 Practical 3 – Gram–Schmidt Orthogonalization

In this weeks practical we will again be working with the vector space $P_4(\mathbb{R})$. Firstly we will apply the Gram-Schmidt process to $P_4(\mathbb{R})$ with the inner product given as an exercise at the end of Practical 2.

$$(\boldsymbol{u}, \boldsymbol{v}) = \frac{8}{\pi}\int_{-1}^{1}\left(u(x)v(x)/\sqrt{(1-x^2)}\right)dx$$

As usual we represent a polynomial $u(x)$ in $P_4(\mathbb{R})$ by a row vector $(x_0, x_1, \cdots, x_4)$, representing $v(x)$ by $(y_0, y_1, \cdots, y_4)$ we have

$$(\boldsymbol{u}, \boldsymbol{v}) = (x_0, \cdots, x_4)A(y_0, \cdots, y_4)^t$$

where A is the matrix

$$A = [A_{ij}], \quad \text{with } A_{ij} = \frac{8}{\pi}\int_{-1}^{1}\left(x^{i+j}/\sqrt{(1-x^2)}\right)dx \ \text{ and } i, j = 0, 1, 2, 3, 4.$$

These integrals can be done using **DERIVE**. The resulting matrix is

$$A = \begin{pmatrix} 8 & 0 & 4 & 0 & 3 \\ 0 & 4 & 0 & 3 & 0 \\ 4 & 0 & 3 & 0 & 5/2 \\ 0 & 3 & 0 & 5/2 & 0 \\ 3 & 0 & 5/2 & 0 & 35/16 \end{pmatrix}$$

We will now utilise the **MATLAB** function file, `gram.m`, which has been placed in the **MATLAB** directory in the mathematics labs at U.N.E. (Externals may type the following file into **MATLAB** within a session, or edit a file called gram.m and keep it in the same directory in which you are using **MATLAB**.) This file contains the following:

```
function z=gram(x,y)
[m,n]=size(x);
z=zeros(x);
h=zeros(x);
z(1,:)=y(1,:);
for i=1:m-1
f=z*x*z';
d=y*x*z'
 for j=1:i, d(j,i)=0;
 end
h(:,i)=d(:,i)/f(i,i);
z=y-h*z;
end
```

For **scilab** create a text file called gram containing

```
function z=gram(x,y)
[m,n]=size(x);
z=zeros(x);
h=zeros(x);
z(1,:)=y(1,:);
for i=1:m-1
f=z*x*z';
d=y*x*z'
 for j=1:i, d(j,i)=0;
 end
h(:,i)=d(:,i)/f(i,i);
z=y-h*z;
end
endfunction
```

and then use getf('gram') to load the program. Now enter the matrix A.
Next you must input the matrix U, whose *rows* are the original linearly independent set of vectors. We will use the standard basis $\boldsymbol{u}_0, \cdots, \boldsymbol{u}_4$ as this set. This means that the matrix U is just the 5 by 5 identity matrix. Enter U (which can be done as U=eye(5,5) in **scilab**).

We can now invoke the function file, `gram`

```
>gram(A,U)
```

The result is then displayed on the screen. This matrix has row vectors which are the vectors $\boldsymbol{v}_0, \cdots, \boldsymbol{v}_4$.

- Write them out as polynomials. They are essentially the Chebyshev polynomials of Prac.1.
- Repeat the exercise with inner product $(\boldsymbol{u}, \boldsymbol{v}) = \displaystyle\int_0^\infty e^{-x} u(x)v(x)dx$.

4.4 Practical 4 – Linear Transformations of Vector Spaces

See section 3.3 for theoretical background into Linear Transformations of Vector Spaces.

In this practical, we will consider the effects of the linear *differential* operator,

$$L = (1 - x^2)\frac{d^2}{dx^2} - 2x\frac{d}{dx}$$

on $P_4(\mathbb{R})$. We have

$$\begin{aligned} L(\boldsymbol{u}_0) &= 0 \\ L(\boldsymbol{u}_1) &= -2x \\ L(\boldsymbol{u}_r) &= -r(r+1)\boldsymbol{u}_r + r(r-1)\boldsymbol{u}_{r-2}, \quad \text{for } r \geq 2. \end{aligned}$$

If we write

$$\boldsymbol{u}(x) = \sum_{i=0}^{n} x_i \boldsymbol{u}_i$$

then the linear transformation $\boldsymbol{u}(x) \to L(\boldsymbol{u}(x))$ can be represented as

$$(x_0, \cdots, x_4) \to (x_0, \cdots, x_4)(t_{ij})$$

where the matrix (t_{ij}) is

$$t = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 & 0 \\ 2 & 0 & -6 & 0 & 0 \\ 0 & 6 & 0 & -12 & 0 \\ 0 & 0 & 12 & 0 & -20 \end{pmatrix}$$

MATLAB has predefined procedures which will calculate the eigenvalues and eigenvectors of a matrix.

The command,

```
>eig(A)
```

returns the eigenvalues of the matrix A, i.e. the numbers λ such that $Ax = \lambda x$, where the column vector x is the eigenvector corresponding to λ. The command `eig(A)` returns the eigenvalues as a column vector.

The eigenvalues and eigenvectors can be obtained with the double assignment statement,

```
>[x,d] = eig(A)
```

This yields the eigenvalues as the diagonal elements of d and the normalized eigenvectors as the columns of x. In **scilab** we use the command

```
-->[x,d] = spec(A)
```

For our problem we will require, `eig(t')`.

- Why the transpose of t?
- Write out the eigenvectors in polynomial form.
- Do you recognise them?
- Can you write a formula for the eigenvalue in terms of the associated polynomial (i.e in terms of n)?
- What differential equation do the eigenvector polynomials satisfy?

4.5 Practical 5 – Applications

PROBLEM 1

In discussing fluid flow past an infinite cylinder we will be looking for solutions to Laplace's equation

$$(*) \qquad \nabla^2 v = 0$$

in cylindrical coordinates (ρ, ϕ, z), with $v = v(\rho, \phi)$. In fact there is a particular separable solution $v = R(\rho)T(\phi)$ where $T(\phi) = \cos k\phi$.

We now look for more general separable solutions.

Substituting $v = R(\rho)T(\phi)$ into (*) gives (in cylindrical polar coordinates), after some rearrangement,

$$\frac{\rho^2}{R(\rho)}[\frac{1}{\rho}R'(\rho) + R''(\rho)] = \frac{-T''(\phi)}{T(\phi)}$$

The L.H.S. is a function of ρ only and the R.H.S. is a function of ϕ only. Hence each must be constant, λ, say. We have

$$T'' + \lambda T = 0 \quad \text{and} \quad \rho^2 R'' + \rho R' = \lambda R$$

The first equation leads to trig or exponential solutions for T, depending on the sign of λ (e.g. $T = const. \cos\phi$ for $\lambda = 1$). Our task then is to solve the eigenvalue problem

$$\rho^2 R'' + \rho R' = \lambda R$$

over certain function spaces (polynomials, in our case). We will obtain two types of solution:

1. Polynomial solutions in $x = \rho$ and
2. Polynomial solutions in $x = \frac{1}{\rho}$.

Your first step in, each case, is to determine the matrix representing the linear transformation $v \longrightarrow Lv$, where

$$L = x^2 \frac{d^2}{dx^2} + x\frac{d}{dx} \quad \text{and} \quad v = \sum_{i=0}^{4} a_i x^i \quad \text{is an element}$$

of $P_4(R)$, i.e. a polynomial of order 4 in x (with $x = \rho$ in case 1 and $x = \frac{1}{\rho}$ in case 2).

Once you have obtained this matrix T, say, then $v \longrightarrow Lv$ is represented by

$$(a_0, \ldots, a_4) \longrightarrow (a_0, \ldots, a_4)T$$

Now enter T into **MATLAB**. Next use the procedure eig, to find the eigenvectors and eigenvalues.

- Rewrite the eigenvectors in polynomial form.
- What is the eigenvector (in each case) for $\lambda = 1$?
- Does this conform with our known solution?
- Can you guess the general eigenvector and eigenvalue?
- What form would you expect a general solution $v = v(\rho, \phi)$ to take?

PROBLEM 2

Solve the eigenvalue problem

$$Ly = \lambda y$$

$L \equiv (x^2 - 3x + 1)\frac{d^2}{dx^2} + x\frac{d}{dx}$ over the order 4 polynomials in x, i.e. $P_4(R)$.

4.6 Practical 6 – Applications

1. Solve the eigenvalue problem

$$x(1-x)\frac{d^2y}{dx^2} + (2-x)\frac{dy}{dx} = \lambda y$$

over the polynomials in x of order 4.

[The above differential equation is a particular case of the hypergeometric equation, many of the special functions of mathematical physics are solutions to equations of this type.]

2. Solve the eigenvalue problem

$$x^3(1-x)\frac{d^3y}{dx^3} + x^2(2-x)\frac{d^2y}{dx^2} + x\frac{dy}{dx} = \lambda y$$

over the polynomials in $1/x$ of order 4.

4.7 Practical 7 – Applications

All of the eigenvalue problems we have dealt with so far have been involved linear operators which act differentiably on our vector space. Another important class of linear operators in the sciences are the integral operators (particularly in transport problems e.g. heat transfer in stellar physics and chemistry). A particularly important class of integral operators can be written in the form

$$u \to I(u) = \int_{y=a}^{b} K(x,y)u(y)dy,$$

where u is an element of a suitable vector space of functions and the function of two variables $K(x,y)$ is called the kernel of the integral operator. Here a and b are two parameters which may depend on x.

If u is a polynomial in some variable then we can solve eigenvalue problems involving integral operators by converting them into problems in matrices and row vectors (just as we did for linear differential operators).

Solve the eigenvlaue problem

$$I(u) = \lambda u, \text{ where } I(u) = \int_{y=a}^{b} K(x,y)u(y)dy$$

in the following cases

(a) $K(x,y) = x^5/y^6,\ a = x\ b = \infty$ and u a polynomial in x of at most degree 4.

[Hint: you have first to determine the effect of I on u, if $u = u_0 + u_1x + u_2x^2 + u_3x^3 + u_4x^4$, where the u_i are constants, then $I(u) = u_0I(1) + u_1I(x) + u_2I(x^2) + u_3I(x^3) + u_4I(x^4)$. So you must evaluate the integrals

$$I(y^n) = \int_x^\infty (x^5/y^6)y^n dy = \int_x^\infty x^5 dy/y^{6-n}$$

for $n = 0, 1, 2, 3, 4$.]

(b) $K(x,y) = x^5e^{x(1-y)},\ a = 1,\ b = \infty,\ x > 0$ and u a polynomial in x of at most degree 4.

[Hint: to determine the effect of the integral operator on u you have to evaluate integrals of the type

$$\int_{y=1}^\infty x^5e^{x(1-y)}y^n dy,\ n = 0, 1, 2, 3, 4,$$

which can involve several integrations by parts. If you have **DERIVE** you can use it to integrate:

```
int(x^5*exp(x*(1-y))*y^n,y,1,inf)
```

for each of the cases $n = 0, 1, 2, 3, 4$.]

4.8 Practical 8 – Quadratic Forms

We use the notation:

- $Q(x) = xAx^t, x$ *a* (row) vector in $\mathbb{R}^n$
- A a symmetric n by n matrix.
- $D = S^{-1}AS$, the diagonal matrix with the eigenvalues of A down it's diagonal.
- $X = (X_1, X_2, \ldots, X_n) = xS = (x_1, x_2, \ldots, x_n)S, X$ is the row vector of new (rotated) coordinates.
- $\lambda_1, \lambda_2, \ldots, \lambda_n$, the eigenvalues of A.
- If e_i is the row vector giving the x_i axis, i.e. e_i has a 1 in ith column with zeroes elsewhere, then e_iS is the ith principal axis. In fact e_iS is just the ith row of S, i.e. the principal axes are just the row vectors of S.

The calculation of the orthogonal matrix S is made easier by noting that $SD = AS$. So that, if the column vectors of S are $c_1, c_2, \ldots, c_n$ (each c_i a column vector), we have

$$\begin{aligned} S &= (c_1, c_2, \ldots, c_n), \quad \text{and} \\ SD &= (\lambda_1 c_1, \lambda_2 c_2, \ldots, \lambda_n c_n) \\ &= AS \\ &= (Ac_1, Ac_2, \ldots, Ac_n) \quad \text{or} \\ Ac_i &= \lambda_i c_i, i = 1, 2, \ldots, n. \end{aligned}$$

So c_i must be an eigenvector for the eigenvalue λ_i. The orthogonality condition $S^tS = I$ becomes

$$\begin{aligned} c_i^t c_i &= 1, \quad \text{for} \quad i = 1, 2, \ldots, n. \quad \text{Thus} \\ S &= (c_1, c_2, \ldots, c_n) \end{aligned}$$

where the column vectors c_i are eigenvectors of A normalized so that $c_i^t c_i = 1$.

MATLAB has the predefined procedure eig which will calculate the eigenvalues and eigenvectors of A. In fact, **MATLAB** gives us the normalized eigenvectors, so that the declaration $[s, d] = eig(A)$ actually gives us $S = (c_1, c_2, \cdots, c_n)$.

EXERCISES

For each of the following quadratic forms find

(a) The orthogonal transformation diagonalizing the form.

(b) The principal axes [recall that these are the rows of S].

(c) The diagonalized form of Q,

1. $4x^2 + 16y^2 + 9z^2 + 8xy + 6xz + 12yz$.
2. $x^2 + 25y^2 + z^2 + 16u^2 + 4w^2 + 10yz - 40yu$.
3. $(x_1)^2 + 9(x_2)^2 - 16(x_3)^2 + (x_4)^2 - 12x_2x_3 + 20x_3x_4$.

5 PRACTICAL SESSIONS (OPTIONAL)

5.1 Practical 9 – An Introduction to DERIVE

Arithmetic

Most of the **DERIVE** menu option commands are not applicable until you have entered some expressions on the screen. Use the Author command to create a new expression. It displays the prompt Author expression and waits for you to type in an expression. You may use the **DERIVE** line editor to change or correct any part of your expression, when you are satisfied with your typed expression you press Enter. Try entering the decimal 32.789. If the expression is syntactically correct it is displayed in the window and you are returned to the command menu. If the expression is not syntactically correct the program emits an error beep, you must then correct your author line. Try entering and then editing, 32.7.89.

The four basic arithmetical operations addition, subtraction, multiplication and division are denoted by +,-,* and /, respectively; you can use parenthese () to control the order of operations. Try entering

$$9 + 2 * 7$$

Now try

$$(9 + 2) * 7$$

Don't forget to press the Enter key after you finish typing each author line.

Exponentiation is denoted by `^`. The operator `^` has a higher precedence than most other operations, so you will often need parentheses to get the correct order of operations in expressions involving `^`. For example, the square root of four must be entered as `4^(1/2)`, rather than `4^1/2`, try it.

In all the expressions you entered above you will have noticed that **DERIVE** did not automatically simplify your typed result. To simplify an expression you issue the Simplify command. Enter the expression

```
5*(7-3+2*3)^4/(2+3^2)
```

and simplify the result. **DERIVE** does exact arithmetic (no approximations, unless asked for) with the result of a simplication being displayed in rational form, where possible. You can get an approximate answer by using the approx command. The various commands operate on the expression number specified or on the highlighted expression, highlighting is controlled by the arrow keys (whole expressions or different parts of an expression may be highlighted) – try simplifying some of your earlier expressions.

If you have no interest in the unsimplified expression you can enter and simplify in one step by using Ctrl/Enter (hold down the Ctrl key and press Enter).

If you wish to use a previous expression (or part of it) in a new author line you simply use the function keys F3 and F4 : F3 inserts the highlighted expression into the current author line, F4 inserts the highlighted expression into the current author line with parentheses. Try authoring the expression 2*(7-3+2*3) using one of your previous lines and an F-key. An alternative way to refer to previous lines is to type the number line preceded by the `#` character. Try entering `#2*#3+#4-#5`.

Algebra

DERIVE can also handle algebraic expressions, all the above arithmetic operations can be applied to expressions involving one or more unknowns. Try entering and simplifying the following

```
3*x+6^2*x-6*x
x^2*(x^3-6*y)-2*y*x^2
(x^2+2*x*y+y^2)/(x+y)
(x^3+3*x^2*y+3*x*y^2+y^3)/(x+y)
```

You can substitute numerical values for an unknown using the Manage command followed by the Substitute subcommand (remember you can simply type the capitalized letter of a command to issue that command); try setting x=1 and y=2 into your last expression.

DERIVE can also solve algebraic expressions for a specified unknown (use the solve command) to solve the equation $x^2 - xy - 6y^2 = 0$ for x you enter the left side only and then use the solve command.

Functions

DERIVE has many predefined functions (you can also define your own using the Declare command) a short list follows

Exponential Functions; `#e` is the base of the natural logarithms. If you enter `#e` it displays as $\hat{e}$.
exp(z) is the exponential function (in z). Upon simplification it is replaced by e^z. Typing `#e^z` gives the same result.

Logarithmic Functions; `ln(z)` is the natural logarithm function (in z – any other variable can be used).
If z is complex the imaginary part of `ln(z)` is between $-\pi$ and π.
`log(x,w)` is the logarithm of x to base w.

Trig Functions; All trig functions use radian angle measure. The various trig functions recognised by **DERIVE** are

```
sin(x),cos(x),tan(x),csc(x),sec(x),cot(x)
```

and **asin(x)**...inverse sin, **acos(x)**...inverse cos etc.

CALCULUS DERIVE can do symbolic limits, differentiation, Taylor polynomial approximation, integration, series summation and extended products.

The Calculus commands so not automatically simplify the expressions concerned, this can be done using the Simplify command or directly using Ctrl/Enter. We will look at the differentiation and integration packages:

Differentiation – enter an expression and differentiate it using the Calculus command with the differentiate subcommand. Try it on (x+1)/(2*x+3). Alternatively, the derivative of an expression, u, with respect to an independent variable x, can be authored by entering the expression dif(u,x). The n-th order derivative is entered by typing dif(u,x,n). You can also take derivatives of expressions involving an unknown function – provided you have first used the Declare Function command. Try some examples, don't forget you can get the result in one step by using Ctrl/Enter. Note that **DERIVE** is actually calculating partial derivatives.

Integration - Use the Calculus command with the Integrate subcommand to find an antiderivative or definite integral of an expression, the command prompts you to enter the name of the integration variable, the lower limit and the upper limit; if the last two fields are left blank the antiderivative is returned. Alternatively, the antiderivative of an expression, u, with respect to x can be entered by typing int(u,x); the definite integral with lower limit a and upper limit b is entered by typing int(u,x,a,b). Try it!

Graphics DERIVE can produce 2-D and 3-D graphs, this is done using the Plot command on a highlighted expression. If the expression contains one unknown we get a 2-D plot, if the expression contains two unknowns we get a 3-D plot (being the graphs of equations such as y=f(x) and z=f(x,y)). Try graphing the following

```
y=x*sin(x^2)
z=cos((x^2+y^2)/4)/(x^2+y^2+pi))
```

Note that you only enter the right hand sides. There are many menu options available for controlling the appearance of the graph, eg in the 3-D plot you may control the number of grids using the Grid subcommand (use the Tab key for moving between fields in this and other commands).

5.2 Practical 10 – Vector Calculus Using DERIVE

Vectors and Matrices A vector is entered by authoring an expression of the form $[x_1, x_2, x_3, x_4, \cdots, x_n]$ where $x_1, \cdots, x_n$ are the (ordered) elements of the vector. For example,

```
[x+y,x^2,z].  Note the "square brackets".
```

is a three dimensional vector.

Alternatively, you may use the Declare command with the vectoR subcommand. After requesting the dimension of the vector these commands prompt you to enter the elements (individually) of the vector. If you use Ctrl/Enter, rather than Enter, after entering the last element the elements of the resulting vector are simplified. Try these various commands with the vector

```
[x*(x-y),3*y^2-y^2+a,(z^2+2*z+1)/(z+1)].
```

A matrix is a vector of vectors each having the same dimension; each inner vector is a row of the matrix. For example,

```
[[1,2,3,],[4,5,6]]
```

displays as the 2 by 3 matrix with first row 1 2 3 and second row 4 5 6. Try it!

Alternatively, use the Declare command followed by the Matrix subcommand. Again, using Ctrl/Enter will automatically simplify the elements of the matrix. The vector function may be used to generate a vector or matrix from an expression evaluated at a sequence of points. Authoring vector (u, k, n) produces (after simplification) a vector with elements generated by simplifying u with the variable k stepping from 1 through n in steps of size 1. For example,

```
vector(2*x,x,7)
```

simplifies to `[2,4,6,8,10,12,14]`. Authoring `vector(u,k,m,n)` produces a vector of `n-m+1` elements simplifying `u` with variable `k` stepping from `m` through `n` in steps of size 1. Try `vector(y!,y,2,6)`. Authoring `vector(u,k,m,n,s)` produces a vector of `(m-k+1)/s` elements, rounded down, generated by simplifying `u` with variable `k` stepping from `m` through `n` in steps of size `s`. Try `vector(sin(q),q,0,pi/2,0.2)`.

A nested call on vector can be used to generate a matrix. Try

```
vector(vector(j+k,k,1,4),j,1,4)
```

The identity matrix is created using the identity matrix function, for example

```
identity_matrix (4)
```

produces the 4 by 4 identity matrix.

Individual elements (or rows) of a vector or matrix may be extracted using the highlight with the F3 and F4 keys. Alternatively, use the element function, `element(v,n)` or `element(M,j,k)`.

The following operations on vectors and matrices are defined

+ addition of vectors or matrices

- subtraction of vectors or matrices

* multiplication of a vector or matrix by a scalar

/ division of a vector or matrix by a scalar

. dot product of two vectors or matrix product of two matrices (or vector and matrix).

`cross` the cross product of two vectors; eg, `cross([1,2,3],[4,5,6])` produces `[-3,6,-3]`.

` (back-accent) produces the transpose of a matrix, eg, ``[[1,2,3],[4,5,6]]` `` gives `[[1,4],[2,5],[3,6]]`.

`det` produces the determinant of a square matrix; eg, `det[[1,2],[3,4]]`.

`trace` usage: `trace[[1,2],[3,4]]`.

^ raises square matrices to a specified power, usage: `[[1,2],[2,3]]^n`. If `n=-1`, we get the inverse, if it exists.

`row_reduce` usage: `row_reduce (A,B)`, reduces the augmented matrix A|B to row echelon form (eg, in solving the linear system `A.X=B`).

`charpoly(A,z)` gives the characteristic polynomial of `A` in terms of `z`.

`eigenvalues(A,x)` gives the eigenvalues of `A` in terms of `x`.

`Vector Differential Operators`. The standard operators are defined as follows: `grad(u)` produces the (three dimensional) gradient of a specified scalar `u`. To produce the grad of an unspecified function you must declare the function, this is most quickly done as

```
f(x,y,z):=
```

Now try authoring `grad(f(x,y,z))`.

`div` produces the divergence of a vector. Usage: $div([u_1, u_2, u_3])$, where u_1, u_2, u_3 are the components of the vector in terms of `x,y,z`

`laplacian` produces ∇^2 of a given scalar function of `x,y,z`. Usage: $laplacian(f)$.

`curl` produces the curl of a given vector. Usage: $curl([u_1, u_2, u_3])$.

Examples

1. Compute the gradient of the following functions
   (a) xyz
   (b) $x^2 - y^2 - z^2 \sec(z/x)$
2. Compute div and curl of the following
   (a) $\tan(x/y)\mathbf{i} + \sec(y/z)\mathbf{j} + \tan(z/x)\mathbf{k}$
   (b) $x^3\mathbf{i} + y^3\mathbf{j} + z^3\mathbf{k}$
3. Verify that the following function is harmonic

$$a + bln(x^2 + y^2),\ \text{where } a \text{ and } b \text{ are constants}$$

5.3 Practical 11 – Vector Calculus Using DERIVE

In Practical 10 we met the **DERIVE** version of the vector differential operators grad, div, curl and laplacian. **DERIVE** can also calculate scalar (and vector) potentials, calculate with vector differential operators in orthogonal curvelinear coordinates (e.g. spherical polar coordinates) and calculate Jacobians.

The potential function is used to compute the scalar potential of a vector field. **DERIVE** computes the potential using line integrals (in much the same way as you do) however **DERIVE** does not check to see if the original vector field is conservative; you can, of course, get **DERIVE** to check if the vector field is conservative using the curl function.

Examples:

1. Try

   $$\text{potential}[x/(x^2+y^2+z^2)^{1/2},\ y/(x^2+y^2+z^2)^{1/2},\ z/(x^2+y^2+z^2)^{1/2}]$$

   (The answer is $r + c$, c some constant)

2. **WARNING: DERIVE will attempt to find a potential function even if the given vector is not conservative.**
   Try

   $$\text{potential}\ [y^2, 2xz, x^3]$$

   Now check the result by taking grad of this potential and comparing it with the original vector field $y^2\mathbf{i} + 2xz\mathbf{j} + x^3\mathbf{k}$.
   The two vectors are not the same.
   $y^2\mathbf{i} + 2xz\mathbf{j} + x^3\mathbf{k}$ cannot be conservative.
   Check this by taking its curl.

3. Try Question 4 of Assignment 1 (of the vector calculus assignments): Verify that each of the following vector fields is conservative and then find, for each, a scalar potential.

(a) $\mathbf{F} = \mathbf{i} - z\mathbf{j} - y\mathbf{k}$

(b) $\mathbf{F} = (3x^2yz - 3y)\mathbf{i} + (x^3z - 3x)\mathbf{j} + (x^3y + 2z)\mathbf{k}$

Conservative vector fields present a very nice geometric picture which can be displayed using the 2D and 3D plotting facilities of **DERIVE**.

A normal to an equipotential surface (i.e. a level surface of the potential) is given by **F**, as **F**=gradf (where f is the potential), and gradf is always normal to the surfaces f=constant. The field lines for **F** have **F** as a tangent vector so the field lines are curves orthogonal to the equipotential surfaces.

Try **DERIVE**'s plot facility to plot some of the equipotential surfaces, using the potential functions you obtained in (a) and (b).
Remember you only enter the "right hand side" of the equation to be plotted – **DERIVE** plots $z = H(x, y)$, but you enter only $H(x, y)$.

To plot several potential surfaces, choose several constants. We then need solve the equation $f = constant$ for z. In **DERIVE** we enter `f=constant` (where f is the potential), and use the soLve option to solve for z.

6 Assignment 1

Show all working. Include MATLAB transcripts where relevant.

Question 1 For each of the following inner products defined on $P_4(\mathbb{R})$ find

1. The matrix A representing the inner product.
2. A basis $\boldsymbol{v}_0$, $\boldsymbol{v}_1$, $\boldsymbol{v}_2$, $\boldsymbol{v}_3$, $\boldsymbol{v}_4$ othonormal with respect to the inner product.
3. The matrix T giving the change of basis from the standard basis 1, x, x^2, x^3, x^4 to the basis $\{\boldsymbol{v}_i\}_{i=0}^4$.

(a)

$$(f, g) = \int_{-1}^{1} x^4 f(x)g(x)\, dx.$$

(b)

$$(f, g) = \int_{0}^{\infty} e^{-x} f(x)g(x)\, dx.$$

Question 2 For the following differential operators L acting on $P_4(\mathbb{R})$ find

1. The matrix T representing the linear transformation.
2. The eigenvalues and eigenvectors (in terms of $P_4(\mathbb{R})$) for L.

(a)

$$L(y) = (1 - x^2)\frac{d^2y}{dx^2} + (2 - x)\frac{dy}{dx}.$$

(b)

$$L(y) = (1 - x^2 + x^3)\frac{d^3y}{dx^3} + x^2\frac{d^2y}{dx^2} - \frac{dy}{dx}.$$

7 Assignment 2

Show all working. Include MATLAB transcripts where relevant.

Question 1 For the following linear operators I acting on $P_4(\mathbb{R})$ find

1. The matrix representing the linear transformation.
2. The eigenvalues and eigenvectors (in terms of $P_4(\mathbb{R})$) for I.

(a)

$$I(u) = \int_{y=x}^{\infty} \frac{x^7}{y^8} u(y)\, dy.$$

(b)

$$I(u) = \int_{y=x}^{\infty} e^{x-y} \left[u(y) + y \frac{du}{dy}(y) \right] dy.$$

Question 2 For the following quadratic forms Q find

1. The matrix A representing the quadratic form.
2. The orthogonal transformation diagonalising Q.
3. The principal axes. The diagonalised form of Q.

(a)

$$Q = 2xz + 2yu + 2xu.$$

(b)

$$Q = 2x^2 + y^2 + 3z^2 + 2xz - 2xu + 4zu.$$