Functional Analysis

James Emery

4/13/2010

Contents

0.1 Green's functions in Ordinary Differential Equations

See chapter 5 in **A First Course in Partial Differentials,** Hans W. Weinberger, 1965, Blaisdell. The influence function.

0.2 Integral Equations

0.2.1 Fredholm Equations

0.2.2 Volterra Equations

0.3 Fourier Series

0.4 Banach Spaces

0.5 Hilbert Spaces

0.6 Quantum Mechanics

0.7 Topological Vector Spaces

0.8 The Sturm-Liouville Boundary Value Problem

Joseph Liouville (March 24, 1809 September 8, 1882) was a French mathematician.

A differential equation defined on the interval having the form of
and the boundary conditions

is called as Sturm-Liouville boundary value problem or Sturm-Liouville system, where , ; the weighting function are given functions; , , , are given constants; and the eigenvalue is an unspecified parameter.

Orthogonality and General Fourier Series

The non-trivial (non-zero) solutions , , of the Sturm-Liouville boundary value problem only exist at certain , . is called eigenvalue and is the eigenfunction.

0.9 Introduction

Linear algebra is the study of finite dimensional vector spaces and linear transformations. A vector space is a quadruple $(V, F, +, *)$. V is a set of vectors, F is a field of scalars, + is the operation of vector addition, and * is the operation of scalar multiplication. We usually do not write the multiplication operator, that is, we write $\alpha * v$ as αv. Let $\alpha, \beta \in F$ and $u, v, w \in V$. The following axioms are satisfied:

1. $u + v = v + u$
2. $u + (v + w) = (u + v) + w$
3. There is a zero element so that $0 \in V$ so that $u + 0 = u$.
4. For each $u \in V$, there is an inverse element -u so that $u + (-u) = 0$
5. $\alpha(u + v) = \alpha u + \alpha v$.
6. $(\alpha + \beta)u = \alpha u + \beta u$.
7. $(\alpha\beta)u = \alpha(\beta u)$.
8. $1u = u$

A finite set of vectors $v_1, v_2, ...v_n$ is linearly independent if

$$\alpha_1 v_1 + \alpha_2 v_2 + ... + \alpha_n v_2 = 0$$

implies that each α_i is zero. Otherwise the set is called linearly dependent. A subset S of V is a subspace of V if the sum of any two elements in S is in S and the scalar product of any element in F with any element in S is in S.

That is S is closed under addition and scalar multiplication. The subspace spanned by vectors

$$v_1, v_2, ...v_n$$

is

$$S = \{\alpha_1 v_1 + \alpha_2 v_2 + ... + \alpha_n v_n : \alpha_i \in F\}$$

Theorem The nonzero vectors $v_1, v_2, ...v_n$ are linearly dependent if and only if some one of them is a linear combination of the preceeding ones.
Proof. Suppose v_k can be written as a linear combination of $v_1, ...v_{k-1}$. Then we have a linear combination of $v_1, ...v_k$ set equal to zero with $\alpha_k = -1$, so that these vectors are linearly dependent. Conversely, suppose $v_1, ...v_n$ are dependent. Then we can find a set of α_i so that

$$\alpha_1 v_1 + \alpha_2 v_2 + ... + \alpha_n v_n = 0,$$

and at least one of them is not zero. Let k be the largest index for which α_i is not zero, then dividing by α_k we find that v_k is a linear combination of the preceeding vectors.
Corollary. Any finite set of vectors contains a linear independent subset that spans the same space.
Theorem. Let vectors $v_1, v_2, ..., v_n$ span V. Suppose the vectors $u_1, u_2, ..., u_k$ are linearly independent. Then $n \geq k$.
Proof. The set $u_1, v_1, v_2, ..., v_n$ is linearly dependent and spans V, so some v_j is dependent on its predecessors. Then the set $u_1, v_1, v_2, ..., v_{j-1}, v_{j+1}, ..., v_n$ spans V, and is dependent. We may continue this, adding a u_i while removing a v_j and still having a set that spans V and is dependent. This can be continued until the u_i are exhausted, otherwise the v_j would be exhausted first and some subset of the $u_1, u_2, ..., u_k$ would then be dependent, which is not possible. Therefore there are more v_j than u_i, which forces

$$n \geq k.$$

Definition A *basis* of a vector space V is a set of linearly indepencent vectors that spans V. A vector space is finite dimensional if it has a finite basis.
Theorem Suppose a vector space has a finite basis

$$A = \{v_1, v_2, ..., v_n\}.$$

Then any other basis also has n elements.

Proof Let $B = \{u_1, u_2, u_3, ...\}$, which is possibly infinite, be a second basis of V. By the previous theorem $n \geq k$ for any subset

$$u_1, u_2, ..., u_k$$

of B. It follows that

$$B = \{u_1, u_2, u_3, ..., u_m\},$$

for some m, and that $n \geq m$. Reversing the role of A and B, we apply the previous theorem again to get $m \geq n$, which proves the corollary. We conclude that the dimension of a finite dimensional vector space can be well defined as the number of elements in any basis. Any vector $v \in V$ can be represented as a linear combination of the basis elements:

$$v = \alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3 + ... + \alpha v_n v_n.$$

This representation is unique, because if we subtract two different representations we would get a representation of the zero vector as a linear combination of the basis vectors with at least one nonzero coefficent, which contradicts that the vectors are linearly independent.

The scalar coefficients are called the coordinates of v, and form a element in the cartesian n-product of the scalar field F. These n-tuples of scalars themselves form an n dimensional vector space and are isomorphic to the original vector space. Let U and V be two vector spaces. A function

$$T : U \to V$$

is called a linear transformation if

1. For $u_1, u_2 \in U, T(u_1 + u_2) = T(u_1) + T(u_2)$.

2. For $\alpha \in F, u \in U, T(\alpha u) = \alpha T(u)$.

A linear transformation from U to itself, is called a linear operator. Associated with every linear transformation is a matrix. Let

$$\{u_1, u_2, u_3, ..., u_n\}$$

be a bases of U and let

$$\{v_1, v_2, ..., v_m\}$$

a bases of V. For each u_j, $T(u_j)$ is in V, so that it may be written as a linear combination of the basis vectors, we have

$$T(u_j) = \sum_{i=1}^{m} a_{ij} v_i.$$

Now let $u \in U$ and let its coordinates be $x_1, ..., x_n$. Vector u is represented by the coordinate vector

$$\begin{bmatrix} x_1 \\ x_2 \\ ... \\ x_n \end{bmatrix}$$

We have

$$\begin{aligned} T(u) = T(\sum_{j=1}^{n} x_j u_j) &= \sum_{j=1}^{n} x_j T(u_j) \\ &= \sum_{j=1}^{n} x_j \sum_{i=1}^{m} a_{ij} v_i \\ &= \sum_{i=1}^{m} (\sum_{j=1}^{n} a_{ij} x_j) v_i \\ &= \sum_{i=1}^{m} y_i v_i, \end{aligned}$$

where the $y_1, y_2, ..., y_m$ are the components of the vector $T(u)$ in vector space V. We have shown that the coordinate vector x of u is mapped to the coordinate vector y of $T(u)$ by matrix multiplication,

$$\begin{bmatrix} y_1 \\ y_2 \\ ... \\ y_n \end{bmatrix} = \begin{bmatrix} a_{11} & ...... & a_{1n} \\ a_{21} & ...... & a_{2n} \\ ... & ...... & ... \\ a_{m1} & ...... & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ ... \\ x_n \end{bmatrix}$$

Now suppose we have two linear transformations

$$T : U \to V,$$

and

$$S : V \to W.$$

The composite transformation is

$$ST : U \to W,$$

defined by $ST(u) = S(T(u))$.
Theorem. If A is the matrix of linear transformation S, and B is the matrix of linear transformation T, then the matrix multiplication product AB is the matrix of ST.

0.10 Determinants

A determinant is a multilinear functional defined on a square matrix. A linear functional is a mapping from a vector space to a field of scalars. A multilinear functional is a function defined on a cartesian product of the vector space. The column vectors of a matrix may be considered to be vectors of the vector space V. The set of n column vectors constitute a point in the cartesian product. The functional is linear in the sense that

$$f(v_1, v_2, ..., v_k + v'_k, ..., v_n) = f(v_1, v_2, ..., v_k, ..., v_n) + f(v_1, v_2, ..., v'_k, ..., v_n),$$

and

$$f(v_1, v_2, ..., \alpha v_k, ..., v_n) = \alpha f(v_1, v_2, ..., v_k, ..., v_n).$$

A multilinear functional is alternating if interchanging a pair of variables changes the sign of the function.
Definition. A determinant $D(A)$ of an n dimensional square matrix A is the unique alternating multilinear functional defined on the n column vectors of A, which takes value 1 on the identity matrix.
Properties.

1. If two column vectors of a matrix are identical, then the determinant is zero. This follows because interchanging the columns changes the sign of the determinant, but the new matrix has not changed, so the value of the determinant is the same. The determinant must be zero.

2. Adding a multiple of one column to a second does not change the value of the determinant. This is clear from

$$D(v_1, ..., v_i, ..., v_j + \alpha v_i, ..., v_n)$$

$$= D(v_1, ..., v_i, ..., v_j, ..., v_n) + \alpha D(v_1, ..., v_i, ..., v_i, ..., v_n)$$

$$= D(v_1, ..., v_i, ..., v_j, ..., v_n) + 0.$$

Example To compute the determinant of

$$\begin{bmatrix} 2 & 2 \\ 3 & 4 \end{bmatrix}$$

subtract the first column from the second, then three times the second from the first, getting

$$\begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$$

So $D(A) = 2D(I) = 2$, where I is the identity matrix. Once we have a matrix in diagonal form, we see from its definition as a multilinear functional, that the determinant is equal to the product of each multiplier of each column times the determinant of the identity. That is, the value equals the product of the diagonal elements.

Example *Cramers's Rule* Suppose we have a system of n equations in n unknowns $x_1, x_2, ..., x_n$ written in the form

$$x_1 v_1 + x_2 v_2 + ..., x_n v_n = v.$$

We have

$$D(v, v_2, v_3, ..., v_n) = D(x_1 v_1 + x_2 v_2 + ... + x_n v_n, v_2, ..., v_n)$$

$$= x_1 D(v_1, v_2, v_3, ..., v_n).$$

So that unknown x_1 is given by

$$x_1 = \frac{D(v, v_2, v_3, ..., v_n)}{D(v_1, v_2, v_3, ..., v_n)}.$$

There is clearly a similar expression for each of $x_2, ...x_n$.

To compute a determinant we can perform permutations on the columns and add scalar multiples of columns to other columns, to put the matrix into diagonal form. Once in diagonal form, (or triangular form) the determinant equals the product of the diagonal elements.

There is an alternate definition of the determinant involving permutations. Let a be a n by n matrix. Consider the sum

$$\sum_{\sigma} s(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} ... a_{n\sigma(n)}$$

where σ is a permutation of the integers $1, 2, 3, 4, ..., n$ and $s(\sigma)$ is the sign of the permutation. The sign of the identitiy permutation is one, and interchanging a pair of elements, a transposition, changes the sigh of the permuntation. Notice that this is a multilinear functional, and is alternating, and further equals one on the identity matrix. Therefore it is the unique such functional, and so is equal to the determinant.

Properties.

1. $D(A^T) = D(T)$.
2. $D(AB) = D(A)D(B)$.
3. Expansion by minors about row i. Let A_{ij} be the matrix obtained from A by deleting row i and column j. Then

$$D(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{i,j} D(A_{ij}).$$

0.11 Permutations

See the paper **Group Theory** by James Emery (Document groups.tex).

0.12 Kernal and Range

Let U be a n dimensional vector space. Let T be a linear transformation from U to a vector space V. Let $K(T)$ be the kernel of T and $R(T)$ the range of T. Then

$$dim(K(T)) + dim(R(T)) = n$$

Let $u_1, u_2, ..., u_p$ be a basis of the kermal of T. It can be extended to be a full basis of U. Suppose V is also n dimensional. Let A be the matrix of T with respect to this basis. Then the first column of A must be all zeroes, so that $D(A)$ is zero. Let U have a second basis and let S be the linear transformation mapping the first bases to the second. Let B be the matrix of S. Then B has an inverse B^{-1} and

$$D(B)D(B^{-1}) = D(BB^{-1}) = D(I) = 1.$$

Then $D(B)$ is not zero. The matrix of T with respect to the second basis is $D(AB) = D(A)D(B)$, so that in general the determinant of the matrix of T with respect to bases of U and V is zero if and only if the kernal of T is not zero, and this is true if and only if T has an inverse.

0.13 Inner Product Spaces

Let V be a vector space with an inner product (u, v) . A basis can be used to construct an orthonormal basis (Gram-Schmidt orthogonalization). An inner product space is also a normed linear space with the norm

$$\|u\| = (u, u)^{1/2}.$$

The Cauchy-Schwartz inequality is

$$(u, v)^2 \leq \|u\|^2 \|v\|^2$$

The triangle inequality is

$$\|u + v\| \leq \|u\|^2 + \|v\|^2.$$

0.14 Gram-Schmidt orthogonalization

Given a vector v and a vector u, the projection of v onto u has length

$$\|v\| \cos(\theta),$$

where θ is the angle between v and u. This can be written as

$$\frac{(v, u)}{\|u\|}$$

The projection of v in the direction of u, or the component of v in the direction u is

$$\frac{(v,u)}{\|u\|}\frac{u}{\|u\|} = \frac{(v,u)}{\|u\|^2}u.$$

So suppose we have an orthogonal set $W = w_1, w_2, ...w_k$ and a vector v, not in W. Let V be the space spanned by W and v. We can find a replacement for v, w_{k+1}, so that $w_1, w_2, ...w_k, w_{k+1}$ is an orthogonal basis for V. We subtract from v its projections in the direction of each of the $w_1, w_2, ...w_k$. So

$$w_{k+1} = v - \frac{(v,w_1)}{\|w_1\|^2}w_1 - \frac{(v,w_2)}{\|w_2\|^2}w_2 - \frac{(v,w_k)}{\|w_k\|^2}w_k$$

Using this process, given a basis of a space, we can compute from this basis an orthogonal basis. See **Numerical Methods,** Dahlquist and Bjorck, 1974, Prentice-Hall, for the modified Gram-Schmidt process, which is an equivalent process with better numerical stability.

0.15 Quadratic Forms

A symmetric matrix defines a quadratic form. Every quadratic form can be diagonalized.

0.16 Canonical Forms

Jordon Normal form, Rational Normal form, Triangular form, LU decomposition.

0.17 Upper Triangular Form

For every matrix M, there is a a matrix T so that

$$T^{-1}MT$$

is upper triangular. One may make an eigenvector of M the first column of T, to get a partitioned matrix where the first column has zeroes below the first element of the first column. Then one may apply induction to the remaining lower dimensional submatrix. See Richard Bellman **The Stability Theory of Differential Equations** Dover, for a simple proof.

0.18 Isometries, Rotations, Orthogonal Matrices

An isometry is a transformation that preserves length.

$$||T(v)|| = ||v||.$$

Let T be a linear transformation that is an isometry. Let T have a matrix representation M in some orthogonal basis. Then we shall show that M is an orthogonal matrix. This means that the norm of any column is one, and that any two columns are orthogonal.

We have

$$Mv \cdot Mv = v \cdot v$$

That is, if v is a column vector, then

$$(Mv)^T(Mv) = v^T M^T M v = v^T v.$$

Let v_i be the column vector whose ith element is one, and all other elements are zero. Then letting $P = M^T M$, we see that $P_{ii} = 1$. Letting $v = v_i + v_j$, for i not j, we find that $P_{ij} = -P_{ji}$. But P is symmetric. So $P_{ij} = 0$. Thus P is the identity matrix. So we have shown that the inverse of an orthogonal matrix is its transpose. This also implies that its row vectors are orthonormal. In two space or three space, a rotation matrix, representing a rotation about a fixed axis, is clearly an isometry. Hence it is orthogonal. Because

$$\det(M) = \det(M^T),$$

the determinant must be 1, or -1. Clearly any real eigenvalue of an isometry must be 1 or -1. This follows from the value of the determinant. An orthogonal matrix whose determinant is 1, called proper. A proper orthogonal transformation represents a rotation. A proper orthogonal matrix defines three Euler angles, which explicitly shows that the matrix represents a rotation. See Nobel, **Applied Linear Algebra.** Also for the determination of the rotation axis of the matrix, which is an eigenvector, see the discussion in the paper on rotation matrices contained in an issue of the IEEE Journal on Computational Geometry and Graphics. Also see the computer codes in the Emery Fortran and C++ libraries.

0.19 Rotation Matrices

See the paper **Rotations** by James Emery (Document rotation.tex).

0.20 Exponentials of Matrices and Operators in a Banach Algebra

See See the paper **Rotations** by James Emery (Document rotation.tex).

0.21 Eigenvalues

0.22 Unitary Transformations

0.23 Transpose, Trace, Self-Adjoint Operators

The trace of a matrix is the sum of the diagonal elements. The trace has the following property. If A and B are square matrices then

$$trace(AB) = trace(BA).$$

This follows because

$$\begin{aligned} trace(AB) &= \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}b_{ji} \\ &= \sum_{j=1}^{n}\sum_{i=1}^{n} a_{ij}b_{ji} \\ &= \sum_{j=1}^{n}\sum_{i=1}^{n} b_{ji}a_{ij} \end{aligned}$$

$$= trace(BA).$$

If

$$M' = PMP^{-1},$$

then

$$trace(M') = trace(PMP^{-1}) = trace(P^{-1}PM) = trace(M).$$

0.24 The Spectral Theorem

0.25 Tensors

A tensor is a multilinear functional.

0.26 Projection Operators

0.27 Linear Algebra Bibliography