Advanced Linear Algebra

ThinkBS: Basic Sciences in Engineering Education

Kadir Has University, Turkey

What is Linear Algebra about? I

These lecture notes are based on the textbook by K. Hoffman and Kunze, Ed 2,(1971). We assume a basic knowledge of matrix algebra, matrix manipulations, matrix inversions, eigenvalue and eigenvector computations

Outline

- We introduce the notion of "vector space" as a set of objects with certain operations.
- We give the set of polynomials and functions with standard addition operation as an example.
- We describe linear independence, bases and coordinates in this general setting.
- We describe linear operators, the matrix of a linear operator and we discuss how the differentiation of polynomials is represented as the multiplication of a matrix with a vector.

What is Linear Algebra about? II

- We introduce the notion of invertibility of a matrix and we discuss the solution of linear systems of equations in terms of matrix algebra;
- We define the determinant of a matrix and introduce the characteristic polynomial of a matrix and show that every matrix annihilates its own characteristic polynomial.
- We define the minimal polynomial of a matrix as the monic polynomial of lowest degree that is annihilated by the given matrix.
- We define similar matrices as the ones that represent the same linear transformation in terms of different bases. We discuss how the characteristic and minimal polynomials are used in the determination of similarity classes.

What is Linear Algebra about? III

- We discuss the "diagonalizability" of linear operators. We define functions of matrices, in particular we discuss how one can define the exponential of a matrix.
- We introduce "inner product spaces" and generalize the notion of orthogonality in abstract linear spaces.
- We discuss complete orthonormal sets and their applications in Fourier series.

Vector spaces I

The definition of vector spaces relies on a number of notions from abstract algebra. We briefly review these notions.

- A "group" G is a set with a single operation on it; this operation has to satisfy certain rules.
- A "ring" R is a set with two operations defined on it. The first operation is usually denoted by $+$ and $(R, +)$ is a group. The second operation is usually denoted bu $\cdot$; elements of R need not have multiplicative inverses.. These two operations satisfy certain compatibility relations.
- A "field" F is a set with two operations $(F, +, \cdot)$, F is a group with respect to the addition, 0 is the additive identity, and nonzero elements form a group with respect to multiplication. These operations satisfy certain compatibility relations. The real and complex numbers are typical examples of fields. In these lectures we will work with these fields and denote them by F.

Vector spaces II

- A vector space V over a field F is a set consisting of objects called "'"vector" with an addition operation defined on it. The set V is a group with respect to the addition operation. This means, the addition of vectors is commutative, associative, it has a zero element and each vector has an additive inverse.

$$+ : V \to V$$

- The fact that $(F, +)$ is a group tells us readily that for example a half space cannot be a vector space, because inverses (negadives) of vectors are not included.

Vector spaces III

- The elements of the field are called "scalars" as opposed to vectors. The structure of a vector space includes the "multiplication by scalars"

$$F \times V \to V$$

  this operation gives a vector. (this should not be confused with scalar multiplication that takes two vectors and gives a scalar). The operations of multiplication by scalars and addition are subject to certain compatibility conditions.
- In short, a vector space is a set of objects that we know how to add and how to multiply by scalars, subject to certain rules. Typical example are pointed line segments in R^n, on column vectors with n elements.

Reference of Abstract Algebra is: Fraleigh, J. B. (2003). A first course in abstract algebra. 7th Edition. Pearson Education.

Groups

A Group is a set G with a binary operation (usually shown as $*$) $* : G \times G \to G$ such that satisfies the followings axioms:

1. For every element $g_1, g_2, g_3 \in G$,
   $(g_1 * g_2) * g_3 = g_1 * (g_2 * g_3) \qquad$ (associativity of $*$)

A Group is a set G with a binary operation (usually shown as $*$) $* : G \times G \to G$ such that satisfies the followings axioms:

1. For every element $g_1, g_2, g_3 \in G$,
   $(g_1 * g_2) * g_3 = g_1 * (g_2 * g_3)$ (associativity of $*$)
2. There is an element e in G such that for all $g \in G$,
   $g * e = e * g = g$ (identity element for $*$)

Groups

A Group is a set G with a binary operation (usually shown as $*$) $* : G \times G \to G$ such that satisfies the followings axioms:

1. For every element $g_1, g_2, g_3 \in G$,
   $(g_1 * g_2) * g_3 = g_1 * (g_2 * g_3)$ (associativity of $*$)
2. There is an element e in G such that for all $g \in G$,
   $g * e = e * g = g$ (identity element for $*$)
3. For each element $g \in G$, there exists an element $g' \in G$ such that $g * g' = g' * g = e$ (inverse of g with respect to $*$)

Examples of vector spaces I

Row vectors, column vectors, matrices, with componentwise addition and multiplication by scalars are typical examples of vector spaces. The set of functions also form vector spaces. Here we should define what we meen by the addition of two functions. If we have two functions $f(x)$ and $g(x)$, in order to define their sum $f+g$, we should define its value at each x. Normally we do thșs by

$$(f+g)(x) = f(x) + g(x).$$

Here note that the $+$ sign on the left hand side is the addition operation in the vector space of functions, while the $+$ sign at the right hand side is the addition operation in the field of real numbers.

Examples of vector spaces II

As a special case of the vector space of functions we will work with the vector space of polynomials of degree at most n. These are of the form

$$p(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0.$$

Here the addition is the usual addition of polynomials, we note that the 0 polynomial is the polynomial that is identially zero, i.e, all of the a_i's are zero.

Subspaces I

A subset S of a vector space V is called a subspace, if it is itself a vector space. The nontrivial thing to check is usually to ensure that S is closed under the vector space operations. In particular the zero vector should be in S. For example lines that do not pass through the origin are not subspaces.

Linear independence; Spanning set, Basis I

A set of vectos$\{X_1, X_2, \ldots, X_k\}$ in a vector space V is called **linearly independent** if the sum

$$c_1X_1 + c_2X_2 + \cdots + c_kX_k = 0$$

implies

$$c_1 = c_2 = \cdots = c_k = 0.$$

Let us give an example: V is the vector space of polynomials of order less than or equal to 3, $X_1 = (x+1)^2$, $X_2 = (x-1)$. We form the linear combination

$$c_1(x+1)^2 + c_2(x-1) = 0$$

and expand

$$c_1(x^2+2x+1) + c_2(x-1) = c_1x^2 + (2c_1+c_2)x + (c_1-c_2) = 0$$

As zero polynomial is the polynomial all of whose coefficients are zero, we should have

$$c_1 = c_2 = 0$$

these the set is linearly independent.
Example: Show that $f_1 = sin^2 x$, $f_2 = cos^2 x$ and $f_3 = 1$ are not linearly independent.
A set of vectors $\{X_1, \dots, X_k\}$ is called a "spanning set for V", if for any X in V, we can find scalars c_i such that

$$X = c_1 X_1 + c_2 X_2 + \cdots + c_k X_k.$$

Example

$$S = \{x^3, x^2, x, x+1, x-1\}$$

is a spanning set for the space of polynomials of degree less than equal to 3. We should find c_i's such that

$$ax^3 + bx^2 + cx + d = c_1(x^3) + c_2(x^2) + c_3(x) + c_4(x-1) + c_5(x+1)$$

Collection terms we get

$$ax^3+bx^2+cx+d = c_1(x^3)+c_2(x^2)+(c_3+c_4+c_5)(x)+(-c_4+c_5)(1)$$

Equating coefficients we get

$$c_1 = a, \quad c_2 = b, \quad c_3 + c_4 + c_5 = c, \quad -c_4 + c_5 = d.$$

Not that the solution is not unique. But the definition is not asking for a uniquely defined solution for the c_i's.
If a spanning set is linearly independent, then these coefficients are uniquely defined. In the case, the spanning set is called a "basis".

A “basis” for a vector space V is a linearly independent set that spans V.
It can be shown that in the context of finite dimensional vector spaces all bases t-have the same number of elements, and this number is called the dimension of the vector space.
When we work with vectors we are used to represent them by sets of numbers. This representation tacitly assumes that we are using a “standard basis” and these numbers are the coefficients of given vector with respect to this standard basis. For example when we see the vector

$$X = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

we assume that it denotes the point in R^3 whose x component is 1, y component is 2 and z component is 3. On the other hand if we were using the space of polynomials of degree less than or equal to 2, with a basis $\{x^2, x, 1\}$ we should understand it as the polynomial

$$p(x) = x^2 + 2x + 3.$$

We might consider another basis for these polynomials, say $\{(x-1)^2, (x-1), 1\}$. Then

$$p(x) = 1(x-1)^2 + 2(x-1) + 3$$

would denote a completely different polynomialç
We see that, a set of numbers, as coordinates makes sense only when we specify the basis.
Recall that a basis consists of a set of vectors. Let's denote these as $B = \{e_i\}$. Since they form a spanning set, any v can be written as

$$v = v_1 e_1 + v_2 e_2 + \cdots + v_n e_n.$$

Linear independence; Spanning set, Basis VI

the v_i's are called the coefficients of the vector v with respect to the basis B.

If we choose another basis $\tilde{B} = \{\tilde{e}_i\}$, then the components will be denoted by $\tilde{v}_i$ and we should have

$$v = v_1 e_1 + v_2 e_2 + \cdots + v_n e_n = \tilde{v}_1 \tilde{e}_1 + \tilde{v}_2 \tilde{e}_2 + \cdots + \tilde{v}_n \tilde{e}_n.$$

We will show how to express the new components $\tilde{v}_i$ in terms of the new basis elements and the old components

Linear independence; Spanning set, Basis VII

$$v = v_1 e_1 + v_2 e_2 + \cdots + v_n e_n = \tilde{v}_1 \tilde{e}_1 + \tilde{v}_2 \tilde{e}_2 + \cdots + \tilde{v}_n \tilde{e}_n.$$

We know the expression of the new basis vectors with respect to the old basis:

Thus we have a change of basis matrix P whose columns are the components of the new basis elements with respect to the old basis. Symbolically we denote this as $\tilde{e}^t = e^t P$. Let X and $\tilde{X}$ the components of the same vector v with respect to the old and to the new basis:

$$v = e^t X = \tilde{e}^t \tilde{X} = e^t P \tilde{X}$$

It follows that $X = P\tilde{X}$ or $\tilde{X} = P^{-1}X$. Note that P is invertible since the columns are linearly independent.

Example: Rotation in R^2 by an angle θ. $X = P\tilde{X}$ or $\tilde{X} = P^{-1}X$.

Example: Change of basis for the vector space of polynomials

It will be useful to think of a linear operator as an operation on vectors that satisfied the linearity rules. Let T denore a linear operator from a vector space V to a vector space W.

$$T: \quad V \to W.$$

, T should be well defined, i.e, it has to be defined on all elements of V and its value $T(X)$ should be an element of W. The linearity rules are

$$T(X + kY) = T(X) + kT(Y).$$

Since T is linear it is sufficient to define it on the basis elements.

Linear operators; Kernel, range, nullity, rank II

Examples: Rotations, reflections, projections in the plane.

Example: Differentiation of polynomials.

Linear operators; Kernel, range, nullity, rank IV

The matrix of a linear transformation is the matrix whose columns are the images of the basis elements of the domain, expressed with respect to the basis of the range space.

Examples: Matrices for rotations, reflections, projections in the plane.

Example: Matrices for the differentiation of polynomials.

Change of basis: Let $T : V \to W$ and $\tilde{e}^t = e^t P$ on V and $\tilde{f}^t = f^t Q$ on W. Let's denote the matrix of the linear transformation with respect to the old bases by A, i.e, $T(e^t) = f^t A$. With respect to the new basis

$$T(\tilde{e}^t) = \tilde{f}^t \tilde{A}.$$

$$\begin{aligned} T(\tilde{e}^t) &= T(e^t P) \\ &= T(e^t) P \\ &= f^t A P \end{aligned}$$

$$\tilde{f}^t \tilde{A} = f^t Q \tilde{A},$$

Hence

$$f^t AP = f^t Q\tilde{A}$$

It folows that

$$AP = Q\tilde{A}$$

or

$$\tilde{A} = Q^{-1}AP.$$

Expression with coordinate indices

Example: Matrix of differentiation, how the differentiation of polynomials is represented as the multiplication of a matrix with a vector.

If $W = V$ and we use the same basis in the domain and in the range

$$\tilde{A} = P^{-1}AP.$$

Those matrices that are related by the formula above are in fact matrices of the same linear transformation with respect to different bases.

These matrices are called "similar matrices".

Rank, nullity, invertibility

- If T is a map from V to W, then the null space of T is the set of vectors in V such that $TX = 0$. is a subset of V. It can be shown that it is a subspace. This subspace is called the kernel of T, $Ker(T)$. Its dimension is called the nullity of T.

Linear operators; Kernel, range, nullity, rank XII

- If T is a map from V to W, then the range space of T is the set of vectors in W such that $Y = TX$, for some X in V. It is is a subset of W. It can be shown that it is a subspace. This subspace is called the range space or image of T, $Im(T)$. Its dimension is called the rank of T.

Linear operators; Kernel, range, nullity, rank XIII

- Examples: Reflections, projections differentiations.
- The rank nullity theorem

$$dim(V) = rank(T) + nullity(T)$$

-

Linear functionals, the dual space I

- A special form of linear transformations are linear maps from V to $W = F$.
- The set of such linear transformations is a vector space. It is called the dual space of V and it is denoted by V^*.
- If the basis for V is $\{e_i\}$ then the set of linear transformations f_i defined by

$$f_i(e_j) = \delta_{ij}$$

  is called the dual basis. Here $\delta_{ij} = 1$ if $i = j$ and zero otherwise.
- The double dual V^{**} is defined as the dual of V^*. If V is a finite dimensional vector space, then

$$V^{**} = V.$$

Solution set of linear systems of equations I

A system of linear equations in matrix form is $AX = B$. If $B = 0$ then the system is called homogeneous. otherwise it is called inhomogeneous. The solvability of the system depends on the rank/nullity of A

Day 2: Determinants; Characteristic polynomial, minimal polynomial I

The determinant of a matrix can be defined as a polynomial function of the rowas or of the columns of a matrix, subject to certain rules. In particular

$$det(AB) = det(A)det(B).$$

We prefer to give a recursive definition as below

Day 2: Determinants; Characteristic polynomial, minimal polynomial II

The characteristic polynomial $k_A(\lambda)$ of a matrix A is defined as

$$det(A - \lambda I)$$

where I is the identity matrix and λ is a scalar. If A is an $n \times n$ matrix its characteristic polynomial has order n. (we may assume that $k_A(\lambda)$ is monic. **Cayley-Hamilton theorem** Every matrix satisfies its characteristic polynomial.
Furthermore

$$det(P^{-1}AP) = det(P^{-1})det(A)det(P) = det(A)$$

Thus the characteristic polynomial of similar matrices are the same. The minimal polynomial $m_A(\lambda)$ of a matrix A is the monic polynomial of least degree that is satisfied by A. **Theorem** The minimal polynomial divides the characteristic polynomial.

Day 2: Determinants; Characteristic polynomial, minimal polynomial III

How to prove the existence of the minimal polynomial? This will need the fact that polynomials in one indeterminate form a so-called principal ideal domain. An idea is a set of polynomials I with the property that whenever $p(x)$ is in I then $q(x)p(x)$ is also in I. For example, polynomials $p(x)$ such that $p(A) = 0$ form an ideal. In a principle ideal domain, every ideal has a unique generator. To prove the existence of the minimal polynomial, we first show that the set of polynomials that satisfy $p(A) = 0$ is nonempty. This is proved by the linear dependence of the powers of A. Then the generator of the ideal is the minimal polynomial. There are many proofs of the Cayley-Hamilton theorem and the relation of minimal and characteristic polynomials. We adopt the simple relationship given above. This is sufficient for a first step to the classification of linear transformations.

Day 2: Determinants; Characteristic polynomial, minimal polynomial IV

Remark Polynomials in one variable are factorizable over the complex numbers. This means both $k_A(\lambda)$ and $m_A(\lambda)$ can be written as a product of linear factors raised to certain powers.

$$k_A(\lambda) = (\lambda - \lambda_1)^{a_1}(\lambda - \lambda_2)^{a_2} \dots (\lambda - \lambda_1 k)^{a_k}, \quad a_1 + a_2 + \dots + a_k = n$$

$$m_A(\lambda) = (\lambda - \lambda_1)^{b_1}(\lambda - \lambda_2)^{b_2} \dots (\lambda - \lambda_1 k)^{b_k}, \quad b_1 \leq a_1, b_2 \leq a_2, \dots, b_k \leq$$

Similarity classes of matrices in 2 dimensions. In 2 dimensions, a characteristic polynomial is a polynomial of degree 2

$$k(x) = (x - a_1)(x - a2)$$

or

$$k(x) = (x - a_1)^2$$

That means the roots are either distinct or repeated. Here we work over the complex numbers. Thus if $k(x)$ is irreducible, then its roots are distinct. If the roots are distinct, then the minimal polynomial should be

$$m(x) = (x - a_1)(x - a_2).$$

This is actually true in any dimensions. If the characteristic polynomial has distinct roots, then the minimal and the characteristic polynomials are the same.

Remark: This is the generic case, in the sense that, if you pick a polynomial randomly, its roots will be distinct. Nevertheless those polynomials with repeated roots important because they determine similarity classes of matrices.

If the roots are repeated, then the minimal polynomial can be either

$$m(x) = (x - a_1)$$

or

$$m(x) = (x - a_1)^2$$

.

We construct matrices for each case:

Day 2: Similarity Classes IV

$n = 3$. In this case also, if the characteristic polynomial has distinct roots, then the minimal polynomial is the same, and a matrix that realize is as below. All cases are listed below:

- $k(x) = (x - a_1)(x - a_2)(x - a_3)$,
  $m(x) = (x - a_1)(x - a_2)(x - a_3)$,
- $k(x) = (x - a_1)^2(x - a_2)$, $m(x) = (x - a_1)^2(a_2)$,
- $k(x) = (x - a_1)^2(x - a_2)$, $m(x) = (x - a_1)(x - a_2)$,
- $k(x) = (x - a_1)^3$, $m(x) = (x - a_1)^3$,
- $k(x) = (x - a_1)^3$, $m(x) = (x - a_1)^2$,
- $k(x) = (x - a_1)^3$, $m(x) = (x - a_1)$.

We give matrices that realize these for each case.

Day 2: Similarity Classes V

$m(x) = (x-a)(x-b)(x-c)$, or if $m(x) = (x-a)(x-b)$ put $c = a$, or if $m(x) = (x-a)$, put $b = c = a$ If the minimal polynomial is linear the matrix is realized by a diagonal matrix. The converse is actually true:

Theorem. A matrix is diagonalizable, if and only if its minimal polynomial is a product of linear factors.

$$\begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix}$$

Day 2: Similarity Classes VI

The minimal polynomial contains a quadratic factor
$m(x) = (x-a)^2(x-b)$ or $m(x) = (x-a)^2$ (put $b = a$)

$$\begin{pmatrix} a & 0 & 0 \\ 1 & a & 0 \\ 0 & 0 & b \end{pmatrix}$$

Day 2: Similarity Classes VII

The minimal polynomial contains a cubic factor $m(x) = (x - a)^3$

$$\begin{pmatrix} a & 0 & 0 \\ 1 & a & 0 \\ 0 & 1 & a \end{pmatrix}$$

Day 2: Similarity Classes VIII

There are matrices with the same characteristic polynomial but different minimal polynomials.

Question: Are these matrices similar?

If matrices A and B have different characteristic polynomials, they cannot be similar. Because if they were similar they would have the same characteristic polynomial.

If matrices A and B have different minimal polynomials, they cannot be similar. Because if they were similar they would have the same minimal polynomial.

Can we say that if matrices that have the same characteristic and minimal polynomials are similar? NO

In 3 dimensions, the matrices above belong to distinct similarity classes (this can be checked directly).

To summarize, if the minimal polynomial has degree 3

$$(x-a)^3, \quad (x-a)^2(x-b), \quad (x-a)(x-b)(x-c)$$

if the minimal polynomial has degree 2

$$(x-a)^2, \quad (x-a)(x-b)$$

if the minimal polynomial has degree 1

$$(x-a)$$

In 4 dimensions the possibilities for the minimal polynomial are
$m(x)$ has degree 4

$$(x-a)^4, \quad (x-a)^3(x-b), \quad (x-a)^2(x-b)^2, \quad (x-a)^2(x-b)(x-c), \quad (x-$$

Day 2: Similarity Classes XI

$m(x)$ has degree 3

$$(x-a)^3, \quad (x-a)^2(x-b), \quad (x-a)(x-b)(x-c)$$

Day 2: Similarity Classes XII

$m(x)$ has degree 2

$$(x-a)^2, \quad (x-a)(x-b),$$

Day 2: Similarity Classes XIII

$m(x)$ has degree 1

$$(x - a),$$

Day 2: Similarity Classes XIV

Remark, in the case $k(x) = (x-a)^4$, $m(x) = (x-a)^2$, there are 2 possibilities, this shows that the equality of characteristic and minimal polynomials is not enough the determine similarity classes. As one goes to higher dimensions, there is room for placing blocks of different sizes and it is easy to construct such examples.

A REDUCE program to check whether all matrices with $m(x) = (x - a)^2$ are similar or not.

Day 2: Similarity Classes XVI

We found a matrix P that satisfies $AP = PB$ but P is not invertible.

Some examples in higher dimensions

Eigenvalues and eigenvectors I

We have seen how to construct examples of matrices with given characteristic and minimal polynomials.
Consider the matrix equation

$$(A - \lambda I)X = 0$$

This means X lies in the kernel of $(A - \lambda I)$, in other words we should have

$$det(A - \lambda I) = 0$$

The roots of the characteristic equation are called **eigenvalues** of A and the vectors X that satisfy $(A - \lambda_i I)X = 0$ are called eigenvectors associated with the eigenvalue λ_i.
Actually the set of eigenvectors associated to a given eigenvalue form a vector space. One should rather talk of an eigenspace associated to an eigenvalue.

Eigenvalues and eigenvectors II

eigenvectors and eigenspaces can also be defined as follows. A vector X is an eigenvector of A if there is a scalar λ such that $AX = \lambda X$.

Example: Projection operator:

Eigenvalues and eigenvectors III

Examples of similarity classes $n = 2$

Eigenvalues and eigenvectors IV

Examples of similarity classes $n = 3$

Examples of similarity classes $n = 4$

Examples of similarity classes $n = 5$

Examples of similarity classes $n = 6$

The Jordan Canonical form and the rational canonical form I

In this section we will give an outline of the proofs of the existence of canonical forms of matrices. We will need a number of basic results from abstract algebra.

- A field F is said to be algebraically closed, if every polynomial with coefficients in F has a root in F. The field of real numbers is not algebraically closed, because, for example the polynomial $x^2 + 1$ has no real roots. On the other hand the field of complex numbers is algebraically closed. If F is algebraically closed, every polynomial van be written as a product of linear factors raised to certain powers.
- The field of complex numbers is a "field extension" of real numbers; this is done by introducing the element i which is the square root of -1. It follows that, over the field of real numbers, every polynomial can be written as a product of linear or irreducible quadratic factors raised to certain powers.

The Jordan Canonical form and the rational canonical form II

- Recall that a ring R is a set on which addition and multiplication operations are defined (subject to certain rules). A subset I of R is called an ideal, if whenever p is in I, qp is also in I (one has to be careful with right/left multiplication). Polynomials in one indeterminate is a typical example for a ring. If we define I as those polynomials $p(x)$ such that $p(T) = 0$, then any multiple of p will also annihilate T, thus they will form an ideal.
- A set of ring elements is said to generate an ideal, if every member of the ideal can be written in terms of this generating set. A ring is called a "principal ideal domain", if every ideal has a unique generator. The set of polynomials in a single variable is a principal ideal domain.

The Jordan Canonical form and the rational canonical form III

- This property is used crucially in the proofs of the existence of canonical forms. An elegant example is the existence of the minimal polynomial as a uniquely defined object. The argument is as follows: If A is an $n \times n$ matrix, it belongs to an n^2 dimensional space. Therefore not all powers of A are linearly independent. They satisfy a certain linear combination relation to be equal to zero. This is an annihilating polynomial, the ideal is nonempty. It has a unique generator, and this is called the minimal plynomial.

Subspace decompositions, direct sums I

In this section we want to decompose a given vector space into smaller subspaces n such a way that the action of a given linear transformation T on each of these subspaces has a simple form.
We need to start by understanding the notion of subspace decomposition.
We recall that we define subspaces by their spanning sets. We should think of a space W_i as the span of a certain set of vectors S_i.
Example: $R^3 = W_1 + W_2 + W_3$, where W_1 is the (x, y) plane, W_2 is the (x, z) plane, W_3 is the (y, z) plane is a subspace decomposition.
A subspace decomposition is called to be a direct sum, is a given vector X can be uniquely decomposed as

$$X = X_1 + X_2 + X_k$$

with each X_i belonging to W_i.

Subspace decompositions, direct sums II

Example: Example: $R^3 = W_1 + W_2 + W_3$, where W_1 is the x axis, W_2 is the y axis, W_3 is the z axis is a direct sum decomposition. Remark: This example should not be misleading. The direct sum decomposition does not imply any notion of orthogonality. We have not yet defined orthogonality. For example

$$R^2 = W_1 + W_2$$

where W_1 is the span of e_1 and W_2 is the span of $e_1 + e_2$ is a direct sum decomposition. We denote direct sums as

$$V = W_1 \oplus W_2 \oplus \cdots \oplus W_k.$$

Invariant subspaces I

Invariant subspaces: A subspace W of V is said to be invariant under A if the image of every vector in W under A lies in W.
The kernel of A and the eigenspaces are invariant subspaces.
Remark: If W is invariant under A it is also invariant under any polynomial in A.
We will be interested in finding invariant direct sums,

$$V = W_1 \oplus W_2 \oplus \cdots \oplus W_k.$$

such that $TW_i \in W_i$.
The Primary Decomposition Theorem: Let T be a linear operator on a finite dimensional vector space V and let p be the minimal polynomial of T

$$p(x) = (x - a_1)^{r_1} \dots (x - a_k)^{r_k}$$

Invariant subspaces II

Let W_i be the kernel of $(T - a_1)^{r_1}$. Then

$$V = W_1 \oplus W_2 \oplus \cdots \oplus W_k$$

where $TW_i \in W_i$. (That means we have an invariant direct sum decomposition). Furthermore, the operators T_i that are the restrictions of T to W_i have minimal polynomials $(x - a_i)^{r_i}$. The proof of this theorem relies on writing T as a sum of projections and it can be omitted at a first reading.
The importance of this therom lies in the fact that, without loss of generality, we can work with linear transformations with minimal polynomials $(x - a)^r$.

Nilpotent and diagonal matrices I

A linear operator T is called nilpotent, if $T^k = 0$ for some k.
Theorem: Let T be a linear operator on a finite dimensional vector space V and the minimal polynomial of T is a product of linear factor (which is always the case when $F = C$. Then, there is a nilpotent operator N and a diagonal operator D such that

$$T = D + N, \quad DN = ND,$$

and D and N are uniquely determined and they are polynomials in T.
This theorem tells us essentially the existence of the Jordan canonical form J, for an operator T with a single eigenvalue a, where J is the sum of a diagonal matrix and a matrix with ones and/or zeros below the main diagonal

Nilpotent and diagonal matrices II

The rational canonical form: A vector X is called to T^iX form a basis for V. The existence of a cyclic operator leads to the "rational canonșcal form

Examples I

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

Characteristic and minimal polynomials are the same:
$k(x) = (x-1)(x-2)^2$. $AX = X$ gives

$$3a + b - c = a, \quad 2a + 2b - c = b, \quad 2a + 2b = c.$$

We get $b = 0$, $c = 2a$. Thus $X = (1, 0, 2)^t$.
$AY = 2Y$ gives

$$3a + b - c = 2a, \quad 2a + 2b - c = 2b, \quad 2a + 2b = 2c.$$

Examples II

We get $b = 2a$, $c = 2a$. Thus $Y = (1,2,2)^t$. We found only one eigenvector corresponding to the eigenvalue 2. We should find Z such that $AZ = 2Z + Y$. Equivalently

$$(A - 2I)Z = Y$$

This is an inhomogeneous system in which the coefficient matrix is non-invertible. The existence of solutions is not guaranteed a priori for such systems. The condition is that the rank of the coefficient matrix and the rank of the augmented system should be equal.

$$rank((A - 2I)) = rank([(A - 2I)Y]).$$

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

This condition is satisfied.

Examples III

The fact that this system has always a solution is ensured by the existence theorem of the Jordan canonical form.

Day 3 Inner Product Spaces I

Let F be the field of real or complex numbers and let V be a vector space over F. An inner product on V is a function that assigns to each pair of vectors X and Y an element of F, denoted by (X, Y), such that for all vectors X, Y, Z, and for all scalars c,

- $(X + Y, Z) = (X, Z) + (Y, Z)$
- $(cX, Y) = c(X, Y)$
- $(Y, X) = (X, \bar{Y})$
- $(X, X) > 0$ if $X \neq 0$.

Examples:

- The standard inner product on R^n i.e, the dot product is a typical example.
- On the space of matrices

$$(A, B) = trace(A\bar{B}^t = \sum_{i,j} A_{ij}\bar{B}_{ij}$$

- If Q is an invertible matrix,

$$(X, Y) = (\bar{(}Y)^t \bar{Q}^t Q X$$

is an inner product.

- If V is the vector space of functions,

$$(f, g) = \int_0^1 f(t) g\bar{(}t)\ dt$$

is an inner product.

Norms I

The positive square root of (X, X) is called the norm of X and it is denoted by $||X||$.

The properties of the inner products lead to the following properties for norms

- $||cX|| = |c| \; ||X||$,
- $||X|| > 0$ for $X \neq 0$.
- $|(X, Y)| \leq ||X||| \; ||Y||$, (Cauchy-Schwarz inequality)
- $||X + Y|| \leq ||X|| + ||Y||$.

Gram-Schmidt orthogonalization I

Given any set of linearly independent vectors, one can construct an orthogonal set of vectors.
Example: $\{X_1, X_2, X_3\}$.

- Start by $Y_1 = X_1$
- Define $Y_2 = X_2 - (X_2, Y_1)/||Y_1||\ Y_1$. Then $(Y_2, Y_1) = 0$.
- Define $Y_3 = X_3 - (X_3, Y_2)/||Y_2||\ Y_2 - (X_3, Y_1)/||Y_1||\ Y_1$.
  Then $(Y_3, Y_1) = (Y_3, Y_2) = 0$.
- Define $Y_m =$
  $X_m - (X_m, Y_{m-1})/||Y_{m-1}||\ Y_{m-1} - \cdots - (X_3, Y_1)/||Y_1||\ Y_1$.
  Then $(Y_m, Y_1) = \ldots (Y_m, Y_{m-1}) = 0$.

After normalization we get an orthonormal set.
Theorem: Every inner product space has an orthonormal basis.

Orthogonal direct sum decompositions I

If S is a set in a vector space, the orthogonal complement of S is the set of vectors that are orthogonal to the vectors in S.
A vector Y in W is a best approximation to X by vectors in W such that $||X - Y|| \leq ||X - Z||$ for all Z in W.
Given a subspace W of an inner product space V and a vector X in V, a vector Y in W, is the best approximation for X in W, if and only if $X - Y$ is orthogonal to every vector in W. This is the orthogonal projection of X on W.

The adjoint of a linear transformation I

Let T be a linear operator on an inner product space. We say that T has an adjoint on V if there is a linear operator T^*on V such that

$$(TX, Y) = (X, T^*Y)$$

For example if A is an $n \times n$ matrix,

$$(TX, Y) = (TX)^t Y = X^t T^t Y = (X, T^t Y).$$

Thus the adjoint of a linear transformation is its transpose.
A linear operator is self adjoint if

$$T^* = T.$$

Normal Operators I

- An operator on a finite dimensional inner product space is called normal, if it commutes with its adjoint,

$$TT^* = T^*T.$$

- : Let V be a finite dimensional inner product space. Every self adjoint operator T has a (nonzero) eigenvector.
  The proof is non elementary.
- Theorem: Let T be a self adjoint operator on V. Eigenvalues of T are real. Let $TX = cX$. $c(X, X) = (cX, X) = (TX, X) = (X, T^*X) = (X, TX) = (X, cX) = \bar{c}(X, X)$ Thus $c = \bar{c}$.
- Theorem: Let T be a self adjoint operator on V. Eigenvectors associated with different eigenvalues are orthogonal. Let $TX = cX$, $TY = dY$. $c(X, Y) = (cX, Y) = (TX, Y) = (X, T^*Y) = (X, TY) = (X, dY) = \bar{d}(X, Y) = d(X, Y)$ If $c \neq d$, then $X, Y) = 0$.

Normal Operators II

- Theorem. If T is a self-adjoint operator, then it is diagonalizable, with respect to an orthonormal basis.

Diagonalization I

A matrix is called
Theorem. If A is a normal matrix, then there is a unitary matrix P such that

$$P = P^{-1}AP$$

is diagonal. (Unitary means $\bar{P}^t P = I$. Orthogonal means $P^t P = I$