Linear Algebra and Matrix Theory

Part 2 - Vector Spaces

1. References

(1) S. Friedberg, A. Insel and L. Spence, *Linear Algebra*, Prentice-Hall.

(2) M. Golubitsky and M. Dellnitz, *Linear Algebra and Differential Equations Using Matlab*, Brooks-Cole.

(3) K. Hoffman and R. Kunze, *Linear Algebra*, Prentice-Hall.

(4) P. Lancaster and M. Tismenetsky, *The Theory of Matrices*, Academic Press.

2. Definition and Examples

Briefly, a vector space consists of a set of objects called vectors along with a set of objects called scalars. The vector space axioms concern the algebraic relationships among the vectors and scalars. They may be found in any of the above references. Informally, they are as follows. Vectors can be added to form new vectors. Vectors can also be multiplied by scalars to form new vectors. There is one vector, called the zero vector, that acts as an identity element with respect to addition of vectors. Each vector has a negative associated with it. The sum of a vector and its negative is the zero vector. Addition of vectors is associative and commutative. In summary, the set of vectors with the operation of vector addition is an abelian group.

The set of scalars is an algebraic field, such as the field of real numbers or the field of complex numbers. Multiplication of scalars by vectors distributes over addition of vectors and also over addition of scalars. The product of the unit element 1 of the field with any vector is that vector. Finally, multiplication of vectors by scalars is associative.

Generically, we denote the set of vectors by $\mathcal{V}$ and the field of scalars by $\mathbb{F}$. It is common to say that $\mathcal{V}$ is a vector space over $\mathbb{F}$. Individual vectors will be denoted by lower case Latin letters and individual scalars by lower case Greek letters. We use the same symbol $+$ to denote both addition of scalars and addition of vectors. Multiplication of scalars by scalars or scalars by vectors is indicated simply by conjoining the symbols for the multiplicands.

Example 2.1. *Let $\mathbb{F}$ be either $\mathbb{R}$ or $\mathbb{C}$ and let $\mathcal{V} = \mathbb{F}^{m\times n}$. Two vectors (matrices) are added by adding corresponding entries and a matrix is multiplied by a scalar by multiplying each entry by that scalar. The zero matrix is the matrix all of whose entries are 0. The negative of a matrix is obtained by multiplying it by -1. If $m = 1$ the vectors of this space are called row vectors and if $n = 1$ they are called column vectors.*

Example 2.2. *Two directed line segments in the Euclidean plane are equivalent if they have the same length and the same direction. Let $\mathcal{V}$ be the set of equivalence classes. If u and v are vectors (i.e., elements of $\mathcal{V}$), choose representative line segments such that the representative of v begins where the representative of u ends. Then $u + v$ is represented by the line segment* *from the initial point of* u *to the terminal point of* v. *To multiply a vector* u *by a real number* α, *let* αu *be represented by a line segment whose length is* $|\alpha|$ *times the length of* u. *Let* αu *have the same direction as* u *if* $\alpha > 0$ *and the opposite direction if* $\alpha < 0$. *These are the familiar geometric vectors of elementary mathematics. We leave it to you to explain how the zero vector and negatives of vectors are defined and to verify the other properties of a vector space.*

Example 2.3. *Let* $\mathbb{F} = \mathbb{R}$ *and let* $\mathcal{V} = C^k(0,1)$, *the set of all functions* $u : (0,1) \to \mathbb{R}$ *with at least* k *continuous derivatives on* $(0,1)$. *Vector addition and scalar multiplication are defined pointwise:* $(u+v)(t) = u(t) + v(t)$ *and* $(\alpha u)(t) = \alpha u(t)$ *for each* $t \in (0,1)$.

3. Subspaces

Let $\mathcal{V}$ be a vector space over $\mathbb{F}$ and let $\mathcal{W}$ be a nonempty subset of $\mathcal{V}$. We say that $\mathcal{W}$ is a *subspace* of $\mathcal{V}$ if it is a vector space over $\mathbb{F}$ with the operations of vector addition and scalar multiplication inherited from $\mathcal{V}$. This means that $\mathcal{W}$ is closed under vector addition and scalar multiplication and that it contains the zero vector of $\mathcal{V}$. There is a simple test for when a subset of $\mathcal{V}$ is a subspace.

Theorem 3.1. $\mathcal{W}$ *is a subspace of* $\mathcal{V}$ *if and only if* $\alpha u + \beta v \in \mathcal{W}$ *for all* $u, v \in \mathcal{W}$ *and all* $\alpha, \beta \in \mathbb{F}$.

Example 3.1. *In Example 2.3, let* $\mathcal{W} = C^r(0,1)$, *where* $r > k$.

Definition 3.1. *Let $\mathcal{V}$ be a vector space over $\mathbb{F}$ and let S be a nonempty subset of $\mathcal{V}$. The* span *of S is the subspace*

$$sp(S) = \{\alpha_1 u_1 + \cdots + \alpha_k u_k | k \in \mathbb{N}; \alpha_1, \cdots, \alpha_k \in \mathbb{F}; u_1, \cdots, u_k \in S\}$$

We also say that $sp(S)$ is the subspace spanned by *the elements of S.*

Definition 3.2. *Let $A \in \mathbb{F}^{m\times n}$. The* row space *of A is the subspace of $\mathbb{F}^{1\times n}$ spanned by the rows of A. The* column space *of A is the subspace of $\mathbb{F}^{m\times 1}$ spanned by the columns of A. The null space of A is the set of all solutions $x \in \mathbb{F}^{n\times 1}$ of the homogeneous system $Ax = 0$.*

There are short ways of denoting these subspaces. The row space is $\{yA|y \in \mathbb{F}^{1\times m}\}$, the column space is $\{Ax|x \in \mathbb{F}^{n\times 1}\}$, and the null space is $\{x \in \mathbb{F}^{n\times 1}|Ax = 0\}$.

Theorem 3.2. *Row equivalent matrices have the same row space and null space.*

Subspaces can be combined to form new subspaces. One way is by taking their intersection. Another is by forming their sum, defined as

Definition 3.3. *Let $\mathcal{U}$ and $\mathcal{W}$ be subspaces of a vector space $\mathcal{V}$. Their sum is*

$$\mathcal{U} + \mathcal{W} = \{u + w|u \in \mathcal{U}, w \in \mathcal{W}\}.$$

Theorem 3.3. *$\mathcal{U} \cap \mathcal{W}$ and $\mathcal{U} + \mathcal{W}$ are subspaces of $\mathcal{V}$.*

Solved Problems:

1. Tell whether the following sets S are subspaces of $\mathcal{V}$ or not. If the answer is no, explain.

(1) $\mathcal{V} = \mathbb{R}^{1\times 2}$, $S = \{(x_1, x_2) | x_1 = 0 \text{ or } x_2 = 0\}$.

(2) $\mathcal{V} = \mathbb{F}^{n\times 1}$, $S = \{x | Ax = b\}$, where $A \in \mathbb{F}^{m\times n}$ and $b \in \mathbb{F}^{m\times 1}$ are given, $b \neq 0$.

(3) $\mathcal{V} = C^2(0,1)$, $S =$All real solutions y on $(0,1)$ of the homogeneous differential equation $y'' - 2y' + y = 0$.

(4) $\mathcal{V} = \mathbb{C}^{4\times 4}$, $\mathbb{F} = \mathbb{C}$, $S = \{H \in \mathcal{V} | H^* = H\}$, where H^* is the conjugate of H^t.

(5) Same $\mathcal{V}$ and S as in the preceding problem, but $\mathbb{F} = \mathbb{R}$.

Solution:

(1) Not a subspace because it is not closed under vector addition. For example, $(1,0)$ and $(0,1)$ are both in S but $(1,1)$ is not.

(2) Not a subspace because it is not closed under vector addition, not closed under scalar multiplication, does not contain the zero vector, and other reasons.

(3) S is a subspace of $\mathcal{V}$.

(4) Not a subspace because it is not closed under scalar multiplication by elements of $\mathbb{C}$.

(5) S is a subspace. Incidentally, a complex matrix satisfying $H^* = H$ is called *hermitian*.

2. Describe in the simplest possible terms the row space of the matrix

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

Solution: The reduced row echelon form has the same row space as the given matrix. It is

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

The row space is the span of the rows of this matrix, i.e., the set of row vectors of the form

$$(\alpha_1, \alpha_2, -2\alpha_1 + \alpha_2, \alpha_1)$$

where α_1 and α_2 are arbitrary elements of $\mathbb{F}$.

3. Describe in simplest possible terms the null space of the same matrix.

Solution: The reduced row echelon form has the same null space. A vector $x = (x_1, x_2, x_3, x_4)^t$ is in the null space if and only if

$$x_1 - 2x_3 + x_4 = 0$$

$$x_2 + x_3 = 0$$

You may think of two of the variables, say x_1 and x_2, as having arbitrary values and the other two variables as being determined by them according to the last set of equations. After substituting, the null space is the set of all vectors of the form

$$(x_1, x_2, -x_2, -x_1 - 2x_2)$$

where x_1 and x_2 are arbitrary elements of $\mathbb{F}$.

4. Describe in simplest terms the column space of the same matrix.

Solution: A vector y is in the column space of A if and only if there is a solution of the linear system $Ax = y$. In the present case the augmented matrix is

$$\begin{pmatrix} 1 & 2 & 0 & 1 & | & y_1 \\ 0 & 1 & 1 & 0 & | & y_2 \\ 1 & 2 & 0 & 1 & | & y_3 \end{pmatrix}$$

Now row-reduce the augmented matrix to get the coefficient part in row echelon form, with symbolic calculations in the last column. The result is

$$\begin{pmatrix} 1 & 0 & -2 & 1 & | & y_1 - 2y_2 \\ 0 & 1 & 1 & 0 & | & y_2 \\ 0 & 0 & 0 & 0 & | & y_3 - y_1 \end{pmatrix}$$

Obviously, this system has a solution if and only if $y_3 - y_1 = 0$. The column space is the set of all vectors $y = (y_1, y_2, y_3)^t$ with $y_3 = y_1$.

Unsolved Problems:

1. Describe the row space, the null space and the column space of the following matrix.

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

2. Show that the intersection of two subspaces of a vector space is a subspace.

3. Let S_1 and S_2 be nonempty subsets of a vector space $\mathcal{V}$. Show that $sp(S_1) + sp(S_2) = sp(S_1 \cup S_2)$.

4. Let $A \in \mathbb{F}^{m\times n}$. Show that if $y \in \mathbb{F}^{1\times n}$ is in the row space of A and $x \in \mathbb{F}^{n\times 1}$ is in the null space of A, then $yx = 0$.

4. Linear Independence, Bases and Coordinates

Definition 4.1. *Vectors* $v_1, \cdots, v_m$ *in a vector space* $\mathcal{V}$ *over* $\mathbb{F}$ *are* linearly dependent *if there are scalars* $\alpha_1, \cdots, \alpha_m$*, not all zero, such that* $\alpha_1 v_1 + \cdots + \alpha_m v_m = 0$*. Otherwise,* $v_1, \cdots, v_m$ *are* linearly independent*.*

An informal way of expressing linear dependence is to say that there is a non-trivial linear combination of the given vectors which is equal to the zero vector.

Definition 4.2. *A vector space* $\mathcal{V}$ *is* finite-dimensional *if there is a finite linearly independent set of vectors in* $\mathcal{V}$ *which spans* $\mathcal{V}$*. Such a set of vectors is called a* basis *for* $\mathcal{V}$*.*

Example 4.1. *In* $\mathbb{F}^{1\times n}$*, let* e_i *denote the vector whose entries are all 0 except the* i^{th}*, which is 1. Then* $\{e_1, \cdots, e_n\}$ *is a basis for* $\mathbb{F}^{1\times n}$ *called the* standard basis. *In* $\mathbb{F}^{m\times n}$*, let* $E_{i,j}$ *denote the matrix all of whose entries are 0 except the* i, j^{th}*, which is 1. Then* $\{E_{i,j} | 1 \leq i \leq m; 1 \leq j \leq n\}$ *is a basis for* $\mathbb{F}^{m\times n}$*.*

There is never just one basis for a finite-dimensional vector space. If $v_1, \cdots, v_m$ is a basis then so, for example, is $2v_1, \cdots, 2v_m$. However, we have the following theorem.

Theorem 4.1. *Any two bases for a finite-dimensional vector space have the same number of elements. The set* $\{v_1, \cdots, v_m\}$ *is a basis for* $\mathcal{V}$ *if and only if for each vector* $u \in \mathcal{V}$ *there are unique scalars* $\alpha_1, \cdots, \alpha_m$ *such that* $u = \alpha_1 v_1 + \cdots + \alpha_m v_m$*.*

Definition 4.3. *The* dimension *of a finite-dimensional vector space is the number of elements in a basis. The dimension of* $\mathcal{V}$ *is denoted by* $dim(\mathcal{V})$*.*

If $dim(\mathcal{V}) = m$ and $v_1, \cdots, v_k$ are linearly independent vectors in $\mathcal{V}$, then $k \leq m$ and if $k < m$ we may extend this set of vectors to form a basis $v_1, \cdots, v_k, v_{k+1}, \cdots, v_m$. Similarly, if $u_1, \cdots, u_r$ spans $\mathcal{V}$ then $r \geq m$ and we may select a subset of m of the given vectors which forms a basis for $\mathcal{V}$.

Definition 4.4. *Let* $\mathcal{B} = \{v_1, \cdots, v_m\}$ *be an ordered basis for* $\mathcal{V}$ *and let* $u \in \mathcal{V}$*. The unique m-tuple of scalars* $(\alpha_1, \cdots, \alpha_m)$ *such that* $\sum_{i=1}^{m} \alpha_i v_i = u$

is called the coordinate vector of u relative to the basis $\mathcal{B}$. It is usually represented as a column vector in $\mathbb{F}^{m\times 1}$

$$[u]_{\mathcal{B}} = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_m \end{pmatrix}$$

Solved Problems:

1. Show that any set of more than m vectors in $\mathbb{F}^{m\times 1}$ is linearly dependent.

Solution: Suppose $n > m$ vectors are given. Let $A \in \mathbb{F}^{m\times n}$ be a matrix whose columns are the given vectors. These vectors are linearly dependent if and only if the homogeneous linear system $Ax = 0$ has a solution other than $x = 0$. Since row equivalent matrices have the same null space, we may assume that A is in reduced row echelon form. There are at least $n - m$ columns of A that do not contain leading 1's. The variables corresponding to these columns may be assigned arbitrary values. The variables corresponding to the columns with leading 1's are determined once these are specified.

2. Show that any set of fewer than m vectors in $\mathbb{F}^{m\times 1}$ does not span $\mathbb{F}^{m\times 1}$.

Solution: Again, let A be an $m \times n$ matrix whose columns are the given vectors and $n < m$. These vectors fail to span $\mathbb{F}^{m\times 1}$ if and only if there is a vector $y \in \mathbb{F}^{m\times 1}$ such that the system $Ax = y$ has no solution. For such a vector y and for any invertible matrix Q, the system $QAx = Qy$ has no solution. Thus, we may suppose that A is in reduced row echelon form to begin with. Since A has more rows than columns, its last row must be all zeros. Therefore, if we choose $y_m \neq 0$ the system has no solution.

3. In $\mathbb{F}^{3\times 1}$, let $v_1 = (1,0,0)^t$, $v_2 = (1,-1,0)^t$, and $v_3 = (1,0,-1)^t$. Show that $\mathcal{B} = \{v_1, v_2, v_3\}$ is a basis and find the coordinate vector of $u = (0,1,0)^t$ relative to this basis.

Solution: Let A be a matrix with columns v_1, v_2, and v_3.

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

$\mathcal{B}$ is a basis if and only if it is a linearly independent spanning set of vectors, which is equivalent to invertibility of A. Clearly, A is invertible because its determinant is 1.

To find the coordinates of u, we must find scalars α_1, α_2, and α_3 such that $u = \alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3$. Then $[u]_\mathcal{B} = (\alpha_1, \alpha_2, \alpha_3)^t$ is the solution of $A[u]_\mathcal{B} = u$. The augmented matrix is

$$\begin{pmatrix} 1 & 1 & 1 & | & 0 \\ 0 & -1 & 0 & | & 1 \\ 0 & 0 & -1 & | & 0 \end{pmatrix}$$

The solution is

$$[u]_\mathcal{B} = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$$

Unsolved Problems:

1. Let $\mathcal{V}$ be the set of all polynomials of degree ≤ 3 with coefficients in $\mathbb{F}$. Constant functions are included as polynomials of degree 0. With the obvious definitions of addition and multiplication by elements of $\mathbb{F}$, $\mathcal{V}$ is a vector space over $\mathbb{F}$ of dimension 4. Find a basis for this vector space. With respect to this basis, find the coordinates of the vector $v(z) = -1 + 2z^2$.

2. Let $\mathcal{V}$ be the real vector space of all 3×3 hermitian matrices with complex entries. Find a basis for $\mathcal{V}$.

5. Change of Coordinates Under Change of Basis

Let $\mathcal{B}_1 = \{v_1, \cdots, v_n\}$ be a basis for $\mathcal{V}$. Let $Q = (q_{i,j}) \in \mathbb{F}^{n \times n}$ and define a new set of vectors $\mathcal{B}_2 = \{u_1, \cdots, u_n\}$ by the equations

$$u_j = \sum_{i=1}^{n} q_{i,j} v_i$$

for $j = 1, \cdots, n$.

Theorem 5.1. *$\mathcal{B}_2$ is a basis if and only if the matrix Q is nonsingular. If so, then the coordinate vectors $[w]_1$ and $[w]_2$ with respect to the bases $\mathcal{B}_1$ and $\mathcal{B}_2$ of a vector w are related by*

$$[w]_1 = Q[w]_2$$

The matrix Q is called the *transition matrix* from one basis to the other. If two bases are given, then the transition matrix is completely determined by expressing the elements of one basis as linear combinations of the elements of the other basis.