5 Signal and Weight Vector Spaces

Objectives	5-1
Theory and Examples	5-2
Linear Vector Spaces	5-2
Linear Independence	5-4
Spanning a Space	5-5
Inner Product	5-6
Norm	5-7
Orthogonality	5-7
Gram-Schmidt Orthogonalization	5-8
Vector Expansions	5-9
Reciprocal Basis Vectors	5-10
Summary of Results	5-14
Solved Problems	5-17
Epilogue	5-26
Further Reading	5-27
Exercises	5-28

Objectives

It is clear from Chapters 3 and 4 that it is very useful to think of the inputs and outputs of a neural network, and the rows of a weight matrix, as vectors. In this chapter we want to examine these vector spaces in detail and to review those properties of vector spaces that are most helpful when analyzing neural networks. We will begin with general definitions and then apply these definitions to specific neural network problems. The concepts that are discussed in this chapter and in Chapter 6 will be used extensively throughout the remaining chapters of this book. They are critical to our understanding of why neural networks work.

Theory and Examples

Linear algebra is the core of the mathematics required for understanding neural networks. In Chapters 3 and 4 we saw the utility of representing the inputs and outputs of neural networks as vectors. In addition, we saw that it is often useful to think of the rows of a weight matrix as vectors in the same vector space as the input vectors.

Recall from Chapter 3 that in the Hamming network the rows of the weight matrix of the feedforward layer were equal to the prototype vectors. In fact, the purpose of the feedforward layer was to calculate the inner products between the prototype vectors and the input vector.

In the single neuron perceptron network we noted that the decision boundary was always orthogonal to the weight matrix (a row vector).

In this chapter we want to review the basic concepts of vector spaces (e.g., inner products, orthogonality) in the context of neural networks. We will begin with a general definition of vector spaces. Then we will present the basic properties of vectors that are most useful for neural network applications.

One comment about notation before we begin. All of the vectors we have discussed so far have been ordered n-tuples (columns) of real numbers and are represented by bold small letters, e.g.,

$$\mathbf{x} = \begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix}^T . \tag{5.1}$$

These are vectors in $\Re^n$, the standard n-dimensional Euclidean space. In this chapter we will also be talking about more general vector spaces than $\Re^n$. These more general vectors will be represented with a script typeface, as in χ. We will show in this chapter how these general vectors can often be represented by columns of numbers.

Linear Vector Spaces

What do we mean by a vector space? We will begin with a very general definition. While this definition may seem abstract, we will provide many concrete examples. By using a general definition we can solve a larger class of problems, and we can impart a deeper understanding of the concepts.

Vector Space

Definition. A linear *vector space*, X, is a set of elements (vectors) defined over a scalar field, F, that satisfies the following conditions:

1. An operation called vector addition is defined such that if $\chi \in X$ (χ is an element of X) and $y \in X$, then $\chi + y \in X$.

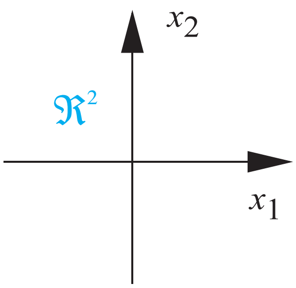
2. $x + y = y + x$.

3. $(x + y) + z = x + (y + z)$.

4. There is a unique vector $0 \in X$, called the zero vector, such that $x + 0 = x$ for all $x \in X$.

5. For each vector $x \in X$ there is a unique vector in X, to be called $-x$, such that $x + (-x) = 0$.

6. An operation, called multiplication, is defined such that for all scalars $a \in F$, and all vectors $x \in X$, $ax \in X$.



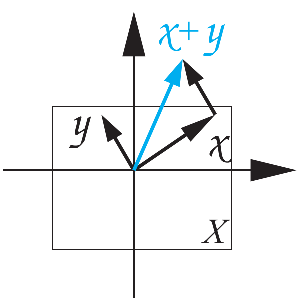
7. For any $x \in X$, $1x = x$ (for scalar 1).

8. For any two scalars $a \in F$ and $b \in F$, and any $x \in X$, $a(bx) = (ab)x$.

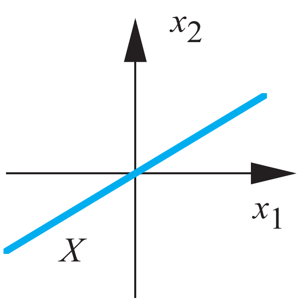
9. $(a + b)x = ax + bx$.

10. $a(x + y) = ax + ay$.

To illustrate these conditions, let's investigate a few sample sets and determine whether or not they are vector spaces. First consider the standard two-dimensional Euclidean space, $\Re^2$, shown in the upper left figure. This is clearly a vector space, and all ten conditions are satisfied for the standard definitions of vector addition and scalar multiplication.



What about subsets of $\Re^2$? What subsets of $\Re^2$ are also vector spaces (subspaces)? Consider the boxed area (X) in the center left figure. Does it satisfy all ten conditions? No. Clearly even condition 1 is not satisfied. The vectors x and y shown in the figure are in X, but $x + y$ is not. From this example it is clear that no bounded sets can be vector spaces.



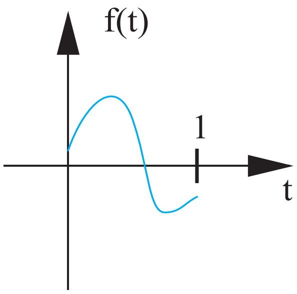
Are there any subsets of $\Re^2$ that are vector spaces? Consider the line (X) shown in the bottom left figure. (Assume that the line extends to infinity in both directions.) Is this line a vector space? We leave it to you to show that indeed all ten conditions are satisfied. Will any such infinite line satisfy the ten conditions? Well, any line that passes through the origin will work. If it does not pass through the origin then condition 4, for instance, would not be satisfied.

In addition to the standard Euclidean spaces, there are other sets that also satisfy the ten conditions of a vector space. Consider, for example, the set P^2 of all polynomials of degree less than or equal to 2. Two members of this set would be

$$\chi = 2 + t + 4t^2$$

$$y = 1 + 5t \,. \tag{5.2}$$

If you are used to thinking of vectors only as columns of numbers, these may seem to be strange vectors indeed. However, recall that to be a vector space, a set need only satisfy the ten conditions we presented. Are these conditions satisfied for the set P^2? If we add two polynomials of degree less than or equal to 2, the result will also be a polynomial of degree less than or equal to 2. Therefore condition 1 is satisfied. We can also multiply a polynomial by a scalar without changing the order of the polynomial. Therefore condition 6 is satisfied. It is not difficult to show that all ten conditions are satisfied, showing that P^2 is a vector space.



Consider the set $C_{[0,\,1]}$ of all continuous functions defined on the interval [0, 1]. Two members of this set would be

$$\chi = \sin(t)$$

$$y = e^{-2t} \,. \tag{5.3}$$

Another member of the set is shown in the figure to the left.

The sum of two continuous functions is also a continuous function, and a scalar times a continuous function is a continuous function. The set $C_{[0,\,1]}$ is also a vector space. This set is different than the other vector spaces we have discussed; it is infinite dimensional. We will define what we mean by dimension later in this chapter.

Linear Independence

Now that we have defined what we mean by a vector space, we will investigate some of the properties of vectors. The first properties are linear dependence and linear independence.

Consider n vectors $\{\chi_1, \chi_2, \ldots, \chi_n\}$. If there exist n scalars $a_1, a_2, \ldots, a_n$, at least one of which is nonzero, such that

$$a_1\chi_1 + a_2\chi_2 + \cdots + a_n\chi_n = 0, \tag{5.4}$$

then the $\{\chi_i\}$ are linearly dependent.

Linear Independence

The converse statement would be: If $a_1\chi_1 + a_2\chi_2 + \cdots + a_n\chi_n = 0$ implies that each $a_i = 0$, then $\{\chi_i\}$ is a set of *linearly independent* vectors.

Note that these definitions are equivalent to saying that if a set of vectors is independent then no vector in the set can be written as a linear combination of the other vectors.

As an example of independence, consider the pattern recognition problem of Chapter 3. The two prototype patterns (*orange* and *apple*) were given by:

$$\mathbf{p}_1 = \begin{bmatrix} 1 \\ -1 \\ -1 \end{bmatrix}, \mathbf{p}_2 = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}. \tag{5.5}$$

Let $a_1\mathbf{p}_1 + a_2\mathbf{p}_2 = \mathbf{0}$, then

$$\begin{bmatrix} a_1 + a_2 \\ -a_1 + a_2 \\ -a_1 + (-a_2) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \tag{5.6}$$

but this can only be true if $a_1 = a_2 = 0$. Therefore $\mathbf{p}_1$ and $\mathbf{p}_2$ are linearly independent.

Consider vectors from the space P^2 of polynomials of degree less than or equal to 2. Three vectors from this space would be

$$\chi_1 = 1 + t + t^2, \ \chi_2 = 2 + 2t + t^2, \ \chi_3 = 1 + t. \tag{5.7}$$

Note that if we let $a_1 = 1$, $a_2 = -1$ and $a_3 = 1$, then

$$a_1\chi_1 + a_2\chi_2 + a_3\chi_3 = 0. \tag{5.8}$$

Therefore these three vectors are linearly dependent.

Spanning a Space

Next we want to define what we mean by the dimension (size) of a vector space. To do so we must first define the concept of a spanning set.

Let X be a linear vector space and let $\{u_1, u_2, \ldots, u_m\}$ be a subset of general vectors in X. This subset spans X if and only if for every vector $\chi \in X$ there exist scalars $x_1, x_2, \ldots, x_n$ such that $\chi = x_1 u_1 + x_2 u_2 + \cdots + x_m u_m$. In other words, a subset spans a space if every vector in the space can be written as a linear combination of the vectors in the subset.

The dimension of a vector space is determined by the minimum number of vectors it takes to span the space. This leads to the definition of a basis set.

Basis Set

A *basis set* for X is a set of linearly independent vectors that spans X. Any basis set contains the minimum number of vectors required to span the space. The dimension of X is therefore equal to the number of elements in the basis set. Any vector space can have many basis sets, but each one must contain the same number of elements. (See [Stra80] for a proof of this fact.)

Take, for example, the linear vector space P^2. One possible basis for this space is

$$u_1 = 1, \; u_2 = t, \; u_3 = t^2. \tag{5.9}$$

Clearly any polynomial of degree two or less can be created by taking a linear combination of these three vectors. Note, however, that *any* three independent vectors from P^2 would form a basis for this space. One such alternate basis is:

$$u_1 = 1, \; u_2 = 1 + t, \; u_3 = 1 + t + t^2. \tag{5.10}$$

Inner Product

From our brief encounter with neural networks in Chapters 3 and 4, it is clear that the inner product is fundamental to the operation of many neural networks. Here we will introduce a general definition for inner products and then give several examples.

Inner Product

Any scalar function of χ and y can be defined as an *inner product*, (χ, y), provided that the following properties are satisfied:

1. $(\chi, y) = (y, \chi)$.
2. $(\chi, a y_1 + b y_2) = a(\chi, y_1) + b(\chi, y_2)$.
3. $(\chi, \chi) \geq 0$, where equality holds if and only if χ is the zero vector.

The standard inner product for vectors in R^n is

$$\mathbf{x}^T \mathbf{y} = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n, \tag{5.11}$$

but this is not the only possible inner product. Consider again the set $C_{[0, 1]}$ of all continuous functions defined on the interval [0, 1]. Show that the following scalar function is man inner product (see Problem P5.6).

$$(\chi, y) = \int_0^1 \chi(t) y(t) dt \tag{5.12}$$

Norm

The next operation we need to define is the norm, which is based on the concept of vector length.

Norm A scalar function $\|\chi\|$ is called a *norm* if it satisfies the following properties:

1. $\|\chi\| \geq 0$.
2. $\|\chi\| = 0$ if and only if $\chi = 0$.
3. $\|a\chi\| = |a|\|\chi\|$ for scalar a.
4. $\|\chi + y\| \leq \|\chi\| + \|y\|$.

There are many functions that would satisfy these conditions. One common norm is based on the inner product:

$$\|\chi\| = (\chi, \chi)^{1/2}. \tag{5.13}$$

For Euclidean spaces, $\Re^n$, this yields the norm with which we are most familiar:

$$\|\mathbf{x}\| = (\mathbf{x}^T\mathbf{x})^{1/2} = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}. \tag{5.14}$$

In neural network applications it is often useful to normalize the input vectors. This means that $\|\mathbf{p}_i\| = 1$ for each input vector.

Angle Using the norm and the inner product we can generalize the concept of angle for vector spaces of dimension greater than two. The *angle* θ between two vectors χ and y is defined by

$$\cos\theta = \frac{(\chi, y)}{\|\chi\|\|y\|}. \tag{5.15}$$

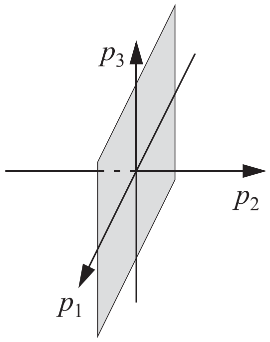
Orthogonality

Now that we have defined the inner product operation, we can introduce the important concept of orthogonality.

Orthogonality Two vectors $\chi, y \in X$ are said to be *orthogonal* if $(\chi, y) = 0$.

Orthogonality is an important concept in neural networks. We will see in Chapter 7 that when the prototype vectors of a pattern recognition problem are orthogonal and normalized, a linear associator neural network can be trained, using the Hebb rule, to achieve perfect recognition.

In addition to orthogonal vectors, we can also have orthogonal spaces. A vector $\chi \in X$ is orthogonal to a subspace X_1 if χ is orthogonal to every vec-



tor in X_1. This is typically represented as $\chi \perp X_1$. A subspace X_1 is orthogonal to a subspace X_2 if every vector in X_1 is orthogonal to every vector in X_2. This is represented by $X_1 \perp X_2$.

The figure to the left illustrates the two orthogonal spaces that were used in the perceptron example of Chapter 3. (See Figure 3.4.) The p_1, p_3 plane is a subspace of $\Re^3$, which is orthogonal to the p_2 axis (which is another subspace of $\Re^3$). The p_1, p_3 plane was the decision boundary of a perceptron network. In Solved Problem P5.1 we will show that the perceptron decision boundary will be a vector space whenever the bias value is zero.

Gram-Schmidt Orthogonalization

There is a relationship between orthogonality and independence. It is possible to convert a set of independent vectors into a set of orthogonal vectors that spans the same vector space. The standard procedure to accomplish this is called Gram-Schmidt orthogonalization.

Assume that we have n independent vectors $y_1, y_2, \ldots, y_n$. From these vectors we want to obtain n orthogonal vectors $v_1, v_2, \ldots, v_n$. The first orthogonal vector is chosen to be the first independent vector:

$$v_1 = y_1. \tag{5.16}$$

To obtain the second orthogonal vector we use y_2, but subtract off the portion of y_2 that is in the direction of v_1. This leads to the equation

$$v_2 = y_2 - a v_1, \tag{5.17}$$

where a is chosen so that v_2 is orthogonal to v_1. This requires that

$$(v_1, v_2) = (v_1, y_2 - a v_1) = (v_1, y_2) - a(v_1, v_1) = 0, \tag{5.18}$$

or

$$a = \frac{(v_1, y_2)}{(v_1, v_1)}. \tag{5.19}$$

Therefore to find the component of y_2 in the direction of v_1, $a v_1$, we need to find the inner product between the two vectors. We call $a v_1$ the *projection* of y_2 on the vector v_1.

Projection

If we continue this process, the kth step will be

$$v_k = y_k - \sum_{i=1}^{k-1} \frac{(v_i, y_k)}{(v_i, v_i)} v_i. \tag{5.20}$$

To illustrate this process, we consider the following independent vectors in $\Re^2$:

$$\mathbf{y}_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \; \mathbf{y}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}. \qquad (5.21)$$

The first orthogonal vector would be

$$\mathbf{v}_1 = \mathbf{y}_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}. \qquad (5.22)$$

The second orthogonal vector is calculated as follows:

$$\mathbf{v}_2 = \mathbf{y}_2 - \frac{\mathbf{v}_1^T \mathbf{y}_2}{\mathbf{v}_1^T \mathbf{v}_1} \mathbf{v}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix} - \frac{\begin{bmatrix} 2 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix}}{\begin{bmatrix} 2 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix}} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} - \begin{bmatrix} 1.6 \\ 0.8 \end{bmatrix} = \begin{bmatrix} -0.6 \\ 1.2 \end{bmatrix}. \qquad (5.23)$$

See Figure 5.1 for a graphical representation of this process.

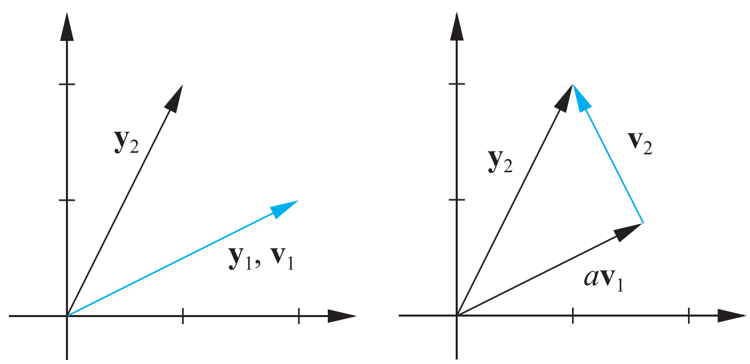


Figure 5.1 Gram-Schmidt Orthogonalization Example

Orthonormal

We could convert $\mathbf{v}_1$ and $\mathbf{v}_2$ to a set of *orthonormal* (orthogonal and normalized) vectors by dividing each vector by its norm.

To experiment with this orthogonalization process, use the Neural Network Design Demonstration Gram-Schmidt (`nnd5gs`).

Vector Expansions

Note that we have been using a script font (χ) to represent general vectors and bold type ($\mathbf{x}$) to represent vectors in $\Re^n$, which can be written as columns of numbers. In this section we will show that general vectors in finite dimensional vector spaces can also be written as columns of numbers and therefore are in some ways equivalent to vectors in $\Re^n$.

Vector Expansion

If a vector space X has a basis set $\{v_1, v_2, \ldots, v_n\}$, then any $\chi \in X$ has a unique *vector expansion*:

$$\chi = \sum_{i=1}^{n} x_i v_i = x_1 v_1 + x_2 v_2 + \cdots + x_n v_n . \tag{5.24}$$

Therefore any vector in a finite dimensional vector space can be represented by a column of numbers:

$$\mathbf{x} = \begin{bmatrix} x_1 & x_2 & \ldots & x_n \end{bmatrix}^T . \tag{5.25}$$

This $\mathbf{x}$ is a representation of the general vector χ. Of course in order to interpret the meaning of $\mathbf{x}$ we need to know the basis set. If the basis set changes, $\mathbf{x}$ will change, even though it still represents the same general vector χ. We will discuss this in more detail in the next subsection.

If the vectors in the basis set are orthogonal ($(v_i, v_j) = 0$, $i \neq j$) it is very easy to compute the coefficients in the expansion. We simply take the inner product of v_j with both sides of Eq. (5.24):

$$(v_j, \chi) = (v_j, \sum_{i=1}^{n} x_i v_i) = \sum_{i=1}^{n} x_i (v_j, v_i) = x_j (v_j, v_j) . \tag{5.26}$$

Therefore the coefficients of the expansion are given by

$$x_j = \frac{(v_j, \chi)}{(v_j, v_j)} . \tag{5.27}$$

When the vectors in the basis set are not orthogonal, the computation of the coefficients in the vector expansion is more complex. This case is covered in the following subsection.

Reciprocal Basis Vectors

If a vector expansion is required and the basis set is not orthogonal, the reciprocal basis vectors are introduced. These are defined by the following equations:

$$\begin{aligned} (r_i, v_j) &= 0 \qquad i \neq j \\ &= 1 \qquad i = j , \end{aligned} \tag{5.28}$$

Reciprocal Basis Vectors

where the basis vectors are $\{v_1, v_2, \ldots, v_n\}$ and the *reciprocal basis vectors* are $\{r_1, r_2, \ldots, r_n\}$.

If the vectors have been represented by columns of numbers (through vector expansion), and the standard inner product is used

$$(r_i, v_j) = \mathbf{r}_i^T \mathbf{v}_j, \tag{5.29}$$

then Eq. (5.28) can be represented in matrix form as

$$\mathbf{R}^T \mathbf{B} = \mathbf{I}, \tag{5.30}$$

where

$$\mathbf{B} = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \ldots & \mathbf{v}_n \end{bmatrix}, \tag{5.31}$$

$$\mathbf{R} = \begin{bmatrix} \mathbf{r}_1 & \mathbf{r}_2 & \ldots & \mathbf{r}_n \end{bmatrix}. \tag{5.32}$$

Therefore $\mathbf{R}$ can be found from

$$\mathbf{R}^T = \mathbf{B}^{-1}, \tag{5.33}$$

and the reciprocal basis vectors can be obtained from the columns of $\mathbf{R}$.

Now consider again the vector expansion

$$\chi = x_1 v_1 + x_2 v_2 + \cdots + x_n v_n. \tag{5.34}$$

Taking the inner product of r_1 with both sides of Eq. (5.34) we obtain

$$(r_1, \chi) = x_1 (r_1, v_1) + x_2 (r_1, v_2) + \cdots + x_n (r_1, v_n). \tag{5.35}$$

By definition

$$\begin{aligned} (r_1, v_2) &= (r_1, v_3) = \cdots = (r_1, v_n) = 0 \\ (r_1, v_1) &= 1. \end{aligned} \tag{5.36}$$

Therefore the first coefficient of the expansion is

$$x_1 = (r_1, \chi), \tag{5.37}$$

and in general

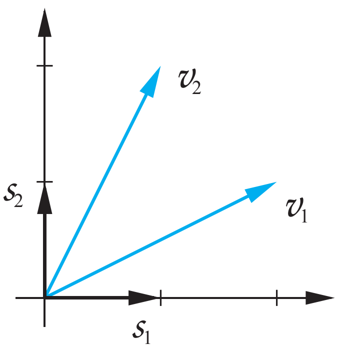
$$x_j = (r_j, \chi). \tag{5.38}$$

As an example, consider the two basis vectors

$$\mathbf{v}_1^s = \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \; \mathbf{v}_2^s = \begin{bmatrix} 1 \\ 2 \end{bmatrix}. \tag{5.39}$$

Suppose that we want to expand the vector

$$\mathbf{x}^s = \begin{bmatrix} 0 \\ \frac{3}{2} \end{bmatrix} \tag{5.40}$$



in terms of the two basis vectors. (We are using the superscript s to indicate that these columns of numbers represent expansions of the vectors in terms of the standard basis in $\Re^2$. The elements of the standard basis are indicated in the adjacent figure as the vectors s_1 and s_2. We need to use this explicit notation in this example because we will be expanding the vectors in terms of two different basis sets.)

The first step in the vector expansion is to find the reciprocal basis vectors.

$$\mathbf{R}^T = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}^{-1} = \begin{bmatrix} \frac{2}{3} & -\frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} \end{bmatrix} \qquad \mathbf{r}_1 = \begin{bmatrix} \frac{2}{3} \\ -\frac{1}{3} \end{bmatrix} \qquad \mathbf{r}_2 = \begin{bmatrix} -\frac{1}{3} \\ \frac{2}{3} \end{bmatrix}. \tag{5.41}$$

Now we can find the coefficients in the expansion.

$$x_1^v = \mathbf{r}_1^T \mathbf{x}^s = \begin{bmatrix} \frac{2}{3} & -\frac{1}{3} \end{bmatrix} \begin{bmatrix} 0 \\ \frac{3}{2} \end{bmatrix} = -\frac{1}{2}$$

$$x_2^v = \mathbf{r}_2^T \mathbf{x}^s = \begin{bmatrix} -\frac{1}{3} & \frac{2}{3} \end{bmatrix} \begin{bmatrix} 0 \\ \frac{3}{2} \end{bmatrix} = 1 \tag{5.42}$$

or, in matrix form,

$$\mathbf{x}^v = \mathbf{R}^T \mathbf{x}^s = \mathbf{B}^{-1} \mathbf{x}^s = \begin{bmatrix} \frac{2}{3} & -\frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} \end{bmatrix} \begin{bmatrix} 0 \\ \frac{3}{2} \end{bmatrix} = \begin{bmatrix} -\frac{1}{2} \\ 1 \end{bmatrix}. \tag{5.43}$$

So that

$$\chi = -\frac{1}{2}v_1 + 1\,v_2\,, \tag{5.44}$$

as indicated in Figure 5.2.

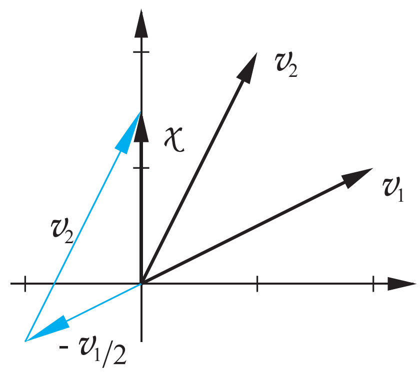


Figure 5.2 Vector Expansion

Note that we now have two different vector expansions for χ, represented by $\mathbf{x}^s$ and $\mathbf{x}^v$. In other words,

$$\chi = 0s_1 + \frac{3}{2}s_2 = -\frac{1}{2}v_1 + 1\,v_2\,. \tag{5.45}$$

When we represent a general vector as a column of numbers we need to know what basis set was used for the expansion. In this text, unless otherwise stated, assume the standard basis set was used.

Eq. (5.43) shows the relationship between the two different representations of χ, $\mathbf{x}^v = \mathbf{B}^{-1}\mathbf{x}^s$. This operation, called a change of basis, will become very important in later chapters for the performance analysis of certain neural networks.

To experiment with the vector expansion process, use the Neural Network Design Demonstration Reciprocal Basis (`nnd5rb`).

Summary of Results

Linear Vector Spaces

Definition. A linear vector space, X, is a set of elements (vectors) defined over a scalar field, F, that satisfies the following conditions:

1. An operation called vector addition is defined such that if $\chi \in X$ and $y \in X$, then $\chi + y \in X$.
2. $\chi + y = y + \chi$.
3. $(\chi + y) + z = \chi + (y + z)$.
4. There is a unique vector $0 \in X$, called the zero vector, such that $\chi + 0 = \chi$ for all $\chi \in X$.
5. For each vector $\chi \in X$ there is a unique vector in X, to be called $-\chi$, such that $\chi + (-\chi) = 0$.
6. An operation, called multiplication, is defined such that for all scalars $a \in F$, and all vectors $\chi \in X$, $a\chi \in X$.
7. For any $\chi \in X$, $1\chi = \chi$ (for scalar 1).
8. For any two scalars $a \in F$ and $b \in F$, and any $\chi \in X$, $a(b\chi) = (ab)\chi$.
9. $(a + b)\chi = a\chi + b\chi$.
10. $a(\chi + y) = a\chi + ay$.

Linear Independence

Consider n vectors $\{\chi_1, \chi_2, \dots, \chi_n\}$. If there exist n scalars $a_1, a_2, \dots, a_n$, at least one of which is nonzero, such that

$$a_1\chi_1 + a_2\chi_2 + \cdots + a_n\chi_n = 0,$$

then the $\{\chi_i\}$ are linearly dependent.

Spanning a Space

Let X be a linear vector space and let $\{u_1, u_2, \ldots, u_m\}$ be a subset of vectors in X. This subset spans X if and only if for every vector $\chi \in X$ there exist scalars $x_1, x_2, \ldots, x_n$ such that $\chi = x_1 u_1 + x_2 u_2 + \cdots + x_m u_m$.

Inner Product

Any scalar function of χ and y can be defined as an inner product, (χ,y), provided that the following properties are satisfied.

1. $(\chi,y) = (y,\chi)$.
2. $(\chi, a y_1 + b y_2) = a(\chi,y_1) + b(\chi,y_2)$.
3. $(\chi,\chi) \geq 0$, where equality holds if and only if χ is the zero vector.

Norm

A scalar function $\|\chi\|$ is called a norm if it satisfies the following properties:

1. $\|\chi\| \geq 0$.
2. $\|\chi\| = 0$ if and only if $\chi = 0$.
3. $\|a\chi\| = |a|\|\chi\|$ for scalar a.
4. $\|\chi + y\| \leq \|\chi\| + \|y\|$.

Angle

The angle θ between two vectors χ and y is defined by

$$\cos\theta = \frac{(\chi,y)}{\|\chi\|\|y\|}.$$

Orthogonality

Two vectors $\chi, y \in X$ are said to be orthogonal if $(\chi,y) = 0$.

Gram-Schmidt Orthogonalization

Assume that we have n independent vectors $y_1, y_2, \ldots, y_n$. From these vectors we will obtain n orthogonal vectors $v_1, v_2, \ldots, v_n$.

$$v_1 = y_1$$

$$v_k = y_k - \sum_{i=1}^{k-1} \frac{(v_i, y_k)}{(v_i, v_i)} v_i \,,$$

where

$$\frac{(v_i, y_k)}{(v_i, v_i)} v_i$$

is the projection of y_k on v_i.

Vector Expansions

$$\chi = \sum_{i=1}^{n} x_i v_i = x_1 v_1 + x_2 v_2 + \cdots + x_n v_n \,.$$

For orthogonal vectors,

$$x_j = \frac{(v_j, \chi)}{(v_j, v_j)}$$

Reciprocal Basis Vectors

$$(r_i, v_j) = 0 \qquad i \neq j$$

$$= 1 \qquad i = j$$

$$x_j = (r_j, \chi) \,.$$

To compute the reciprocal basis vectors:

$$\mathbf{B} = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \ldots & \mathbf{v}_n \end{bmatrix} \,,$$

$$\mathbf{R} = \begin{bmatrix} \mathbf{r}_1 & \mathbf{r}_2 & \ldots & \mathbf{r}_n \end{bmatrix} \,,$$

$$\mathbf{R}^T = \mathbf{B}^{-1} \,.$$

In matrix form:

$$\mathbf{x}^v = \mathbf{B}^{-1} \mathbf{x}^s \,.$$

Solved Problems

P5.1 Consider the single-neuron perceptron network shown in Figure P5.1. Recall from Chapter 3 (see Eq. (3.6)) that the decision boundary for this network is given by $\mathbf{W}\mathbf{p} + b = 0$. Show that the decision boundary is a vector space if $b = 0$.

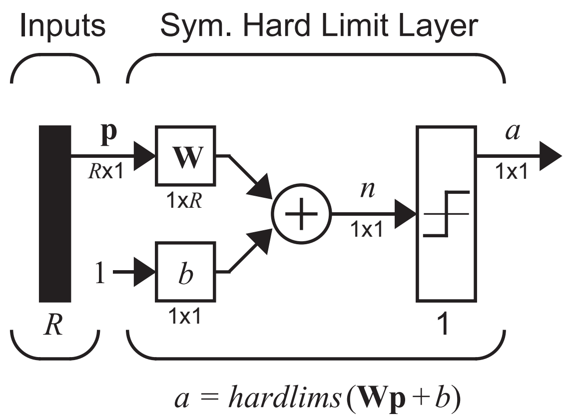


Figure P5.1 Single-Neuron Perceptron

To be a vector space the boundary must satisfy the ten conditions given at the beginning of this chapter. Condition 1 requires that when we add two vectors together the sum remains in the vector space. Let $\mathbf{p}_1$ and $\mathbf{p}_2$ be two vectors on the decision boundary. To be on the boundary they must satisfy

$$\mathbf{W}\mathbf{p}_1 = 0 \qquad \mathbf{W}\mathbf{p}_2 = 0 .$$

If we add these two equations together we find

$$\mathbf{W}(\mathbf{p}_1 + \mathbf{p}_2) = 0 .$$

Therefore the sum is also on the decision boundary.

Conditions 2 and 3 are clearly satisfied. Condition 4 requires that the zero vector be on the boundary. Since $\mathbf{W}\mathbf{0} = 0$, the zero vector is on the decision boundary. Condition 5 implies that if $\mathbf{p}$ is on the boundary, then $-\mathbf{p}$ must also be on the boundary. If $\mathbf{p}$ is on the boundary, then

$$\mathbf{W}\mathbf{p} = 0 .$$

If we multiply both sides of this equation by -1 we find

$$\mathbf{W}(-\mathbf{p}) = 0 .$$

Therefore condition 5 is satisfied.

Condition 6 will be satisfied if for any $\mathbf{p}$ on the boundary $a\mathbf{p}$ is also on the boundary. This can be shown in the same way as condition 5. Just multiply both sides of the equation by a instead of by 1.

$$\mathbf{W}(a\mathbf{p}) = 0$$

Conditions 7 through 10 are clearly satisfied. Therefore the perceptron decision boundary is a vector space.

P5.2 Show that the set Y of nonnegative ($f(t) \geq 0$) continuous functions is not a vector space.

This set violates several of the conditions required of a vector space. For example, there are no negative vectors, so condition 5 cannot be satisfied. Also, consider condition 6. The function $f(t) = |t|$ is a member of Y. Let $a = -2$. Then

$$af(2) = -2|2| = -4 < 0 .$$

Therefore $af(t)$ is not a member of Y, and condition 6 is not satisfied.

P5.3 Which of the following sets of vectors are independent? Find the dimension of the vector space spanned by each set.

i. $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \quad \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \quad \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$

ii. $\sin t \quad \cos t \quad 2\cos\left(t + \frac{\pi}{4}\right)$

iii. $\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \quad \begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \end{bmatrix} \quad \begin{bmatrix} 1 \\ 2 \\ 1 \\ 1 \end{bmatrix}$

i. We can solve this problem several ways. First, let's assume that the vectors are dependent. Then we can write

$$a_1 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + a_2 \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + a_3 \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} .$$

If we can solve for the coefficients and they are not all zero, then the vectors are dependent. By inspection we can see that if we let $a_1 = 2$, $a_2 = -1$ and $a_3 = -1$, then the equation is satisfied. Therefore the vectors are dependent.

Another approach, when we have n vectors in $\Re^n$, is to write the above equation in matrix form:

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 2 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

If the matrix in this equation has an inverse, then the solution will require that all coefficients be zero; therefore the vectors are independent. If the matrix is singular (has no inverse), then a nonzero set of coefficients will work, and the vectors are dependent. The test, then, is to create a matrix using the vectors as columns. If the determinant of the matrix is zero (singular matrix), then the vectors are dependent; otherwise they are independent. Using the Laplace expansion [Brog91] on the first column, the determinant of this matrix is

$$\begin{vmatrix} 1 & 1 & 1 \\ 1 & 0 & 2 \\ 1 & 1 & 1 \end{vmatrix} = 1\begin{vmatrix} 0 & 2 \\ 1 & 1 \end{vmatrix} + (-1)\begin{vmatrix} 1 & 1 \\ 1 & 1 \end{vmatrix} + 1\begin{vmatrix} 1 & 1 \\ 0 & 2 \end{vmatrix} = -2 + 0 + 2 = 0$$

Therefore the vectors are dependent.

The dimension of the space spanned by the vectors is two, since any two of the vectors can be shown to be independent.

ii. By using some trigonometric identities we can write

$$\cos\left(t + \frac{\pi}{4}\right) = \frac{-1}{\sqrt{2}}\sin t + \frac{1}{\sqrt{2}}\cos t .$$

Therefore the vectors are dependent. The dimension of the space spanned by the vectors is two, since no linear combination of $\sin t$ and $\cos t$ is identically zero.

iii. This is similar to part (i), except that the number of vectors is less than the size of the vector space they are drawn from (three vectors in $\Re^4$). In this case the matrix made up of the vectors will not be square, so we will not be able to compute a determinant. However, we can use something called the Gramian [Brog91]. It is the determinant of a matrix whose i, j element is the inner product of vector i and vector j. The vectors are dependent if and only if the Gramian is zero.

For our problem the Gramian would be

$$G = \begin{vmatrix} (\mathbf{x}_1,\mathbf{x}_1) & (\mathbf{x}_1,\mathbf{x}_2) & (\mathbf{x}_1,\mathbf{x}_3) \\ (\mathbf{x}_2,\mathbf{x}_1) & (\mathbf{x}_2,\mathbf{x}_2) & (\mathbf{x}_2,\mathbf{x}_3) \\ (\mathbf{x}_3,\mathbf{x}_1) & (\mathbf{x}_3,\mathbf{x}_2) & (\mathbf{x}_3,\mathbf{x}_3) \end{vmatrix},$$

where

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \qquad \mathbf{x}_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \end{bmatrix} \qquad \mathbf{x}_3 = \begin{bmatrix} 1 \\ 2 \\ 1 \\ 1 \end{bmatrix}.$$

Therefore

$$G = \begin{vmatrix} 4 & 3 & 5 \\ 3 & 3 & 3 \\ 5 & 3 & 7 \end{vmatrix} = 4\begin{vmatrix} 3 & 3 \\ 3 & 7 \end{vmatrix} + (-3)\begin{vmatrix} 3 & 5 \\ 3 & 7 \end{vmatrix} + 5\begin{vmatrix} 3 & 5 \\ 3 & 3 \end{vmatrix} = 48 - 18 - 30 = 0.$$

We can also show that these vectors are dependent by noting

$$2\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} - 1\begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \end{bmatrix} - 1\begin{bmatrix} 1 \\ 2 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

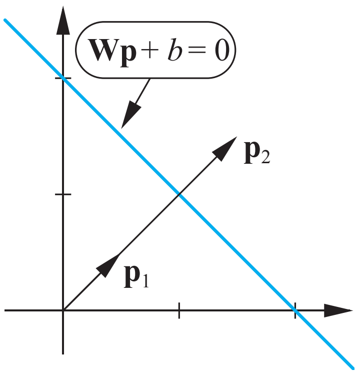
The dimension of the space must therefore be less than 3. We can show that $\mathbf{x}_1$ and $\mathbf{x}_2$ are independent, since

$$G = \begin{vmatrix} 4 & 3 \\ 3 & 3 \end{vmatrix} = 4 \neq 0.$$

Therefore the dimension of the space is 2.

P5.4 Recall from Chapters 3 and 4 that one-layer perceptrons can only be used to recognize patterns that are linearly separable (can be separated by a linear boundary — see Figure 3.3). If two patterns are linearly separable, are they always linearly independent?

No, these are two unrelated concepts. Take the following simple example. Consider the two input perceptron shown in Figure P5.2.



Suppose that we want to separate the two vectors

$$\mathbf{p}_1 = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix} \qquad \mathbf{p}_2 = \begin{bmatrix} 1.5 \\ 1.5 \end{bmatrix}.$$

If we choose the weights and offsets to be $w_{11} = 1$, $w_{12} = 1$ and $b = -2$, then the decision boundary ($\mathbf{Wp} + b = 0$) is shown in the figure to the left. Clearly these two vectors are linearly separable. However, they are not linearly independent since $\mathbf{p}_2 = 3\mathbf{p}_1$.

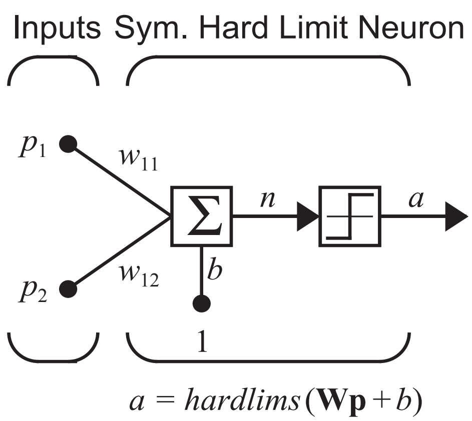


Figure P5.2 Two-Input Perceptron

P5.5 Using the following basis vectors, find an orthogonal set using Gram-Schmidt orthogonalization.

$$\mathbf{y}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \qquad \mathbf{y}_2 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \qquad \mathbf{y}_3 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

Step 1.

$$\mathbf{v}_1 = \mathbf{y}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

Step 2.

$$\mathbf{v}_2 = \mathbf{y}_2 - \frac{\mathbf{v}_1^T\mathbf{y}_2}{\mathbf{v}_1^T\mathbf{v}_1}\mathbf{v}_1 = \begin{bmatrix}1\\0\\0\end{bmatrix} - \frac{\begin{bmatrix}1 & 1 & 1\end{bmatrix}\begin{bmatrix}1\\0\\0\end{bmatrix}}{\begin{bmatrix}1 & 1 & 1\end{bmatrix}\begin{bmatrix}1\\1\\1\end{bmatrix}}\begin{bmatrix}1\\1\\1\end{bmatrix} = \begin{bmatrix}1\\0\\0\end{bmatrix} - \begin{bmatrix}1/3\\1/3\\1/3\end{bmatrix} = \begin{bmatrix}2/3\\-1/3\\-1/3\end{bmatrix}$$

Step 3.

$$\mathbf{v}_3 = \mathbf{y}_3 - \frac{\mathbf{v}_1^T\mathbf{y}_3}{\mathbf{v}_1^T\mathbf{v}_1}\mathbf{v}_1 - \frac{\mathbf{v}_2^T\mathbf{y}_3}{\mathbf{v}_2^T\mathbf{v}_2}\mathbf{v}_2$$

$$\mathbf{v}_3 = \begin{bmatrix}0\\1\\0\end{bmatrix} - \frac{\begin{bmatrix}1 & 1 & 1\end{bmatrix}\begin{bmatrix}0\\1\\0\end{bmatrix}}{\begin{bmatrix}1 & 1 & 1\end{bmatrix}\begin{bmatrix}1\\1\\1\end{bmatrix}}\begin{bmatrix}1\\1\\1\end{bmatrix} - \frac{\begin{bmatrix}2/3 & -1/3 & -1/3\end{bmatrix}\begin{bmatrix}0\\1\\0\end{bmatrix}}{\begin{bmatrix}2/3 & -1/3 & -1/3\end{bmatrix}\begin{bmatrix}2/3\\-1/3\\-1/3\end{bmatrix}}\begin{bmatrix}2/3\\-1/3\\-1/3\end{bmatrix}$$

$$\mathbf{v}_3 = \begin{bmatrix}0\\1\\0\end{bmatrix} - \begin{bmatrix}1/3\\1/3\\1/3\end{bmatrix} - \begin{bmatrix}-1/3\\1/6\\1/6\end{bmatrix} = \begin{bmatrix}0\\1/2\\-1/2\end{bmatrix}$$

P5.6 Consider the vector space of all polynomials defined on the interval [-1, 1]. Show that $(\chi,y) = \int_{-1}^{1} \chi(t)y(t)dt$ **is a valid inner product.**

An inner product must satisfy the following properties.

1. $(\chi,y) = (y,\chi)$

$$(\chi,y) = \int_{-1}^{1} \chi(t)y(t)dt = \int_{-1}^{1} y(t)\chi(t)dt = (y,\chi)$$

2. $(\chi,ay_1 + by_2) = a(\chi,y_1) + b(\chi,y_2)$

$$(\chi, ay_1 + by_2) = \int_{-1}^{1} \chi(t)(ay_1(t) + by_2(t))dt = a\int_{-1}^{1} \chi(t)y_1(t)dt + b\int_{-1}^{1} \chi(t)y_2(t)dt$$
$$= a(\chi, y_1) + b(\chi, y_2)$$

3. $(\chi, \chi) \geq 0$, where equality holds if and only if χ is the zero vector.

$$(\chi, \chi) = \int_{-1}^{1} \chi(t)\chi(t)dt = \int_{-1}^{1} \chi^2(t)\ dt \geq 0$$

Equality holds here only if $\chi(t) = 0$ for $-1 \leq t \leq 1$, which is the zero vector.

P5.7 Two vectors from the vector space described in the previous problem (polynomials defined on the interval [-1, 1]) are $1 + t$ and $1 - t$. Find an orthogonal set of vectors based on these two vectors.

Step 1.

$$v_1 = y_1 = 1 + t$$

Step 2.

$$v_2 = y_2 - \frac{(v_1, y_2)}{(v_1, v_1)} v_1$$

where

$$(v_1, y_2) = \int_{-1}^{1} (1+t)(1-t)dt = \left(t - \frac{t^3}{3}\right)\Bigg|_{-1}^{1} = \left(\frac{2}{3}\right) - \left(-\frac{2}{3}\right) = \frac{4}{3}$$

$$(v_1, v_1) = \int_{-1}^{1} (1+t)^2 dt = \frac{(1+t)^3}{3}\Bigg|_{-1}^{1} = \left(\frac{8}{3}\right) - (0) = \frac{8}{3}.$$

Therefore

$$v_2 = (1 - t) - \frac{4/3}{8/3}(1 + t) = \frac{1}{2} - \frac{3}{2}t.$$

P5.8 Expand $\mathbf{x} = \begin{bmatrix} 6 & 9 & 9 \end{bmatrix}^T$ in terms of the following basis set.

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \qquad \mathbf{v}_2 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \qquad \mathbf{v}_3 = \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix}$$

The first step is to calculate the reciprocal basis vectors.

$$\mathbf{B} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 3 & 2 \end{bmatrix} \qquad \mathbf{B}^{-1} = \begin{bmatrix} \frac{5}{3} & -\frac{1}{3} & -\frac{1}{3} \\ -\frac{1}{3} & -\frac{1}{3} & \frac{2}{3} \\ -\frac{1}{3} & \frac{2}{3} & -\frac{1}{3} \end{bmatrix}$$

Therefore taking the rows of $\mathbf{B}^{-1}$,

$$\mathbf{r}_1 = \begin{bmatrix} 5/3 \\ -1/3 \\ -1/3 \end{bmatrix} \qquad \mathbf{r}_2 = \begin{bmatrix} -1/3 \\ -1/3 \\ 2/3 \end{bmatrix} \qquad \mathbf{r}_3 = \begin{bmatrix} -1/3 \\ 2/3 \\ -1/3 \end{bmatrix}.$$

The coefficients in the expansion are calculated

$$x_1^v = \mathbf{r}_1^T \mathbf{x} = \begin{bmatrix} \frac{5}{3} & \frac{-1}{3} & \frac{-1}{3} \end{bmatrix} \begin{bmatrix} 6 \\ 9 \\ 9 \end{bmatrix} = 4$$

$$x_2^v = \mathbf{r}_2^T \mathbf{x} = \begin{bmatrix} \frac{-1}{3} & \frac{-1}{3} & \frac{2}{3} \end{bmatrix} \begin{bmatrix} 6 \\ 9 \\ 9 \end{bmatrix} = 1$$

$$x_3^v = \mathbf{r}_3^T \mathbf{x} = \begin{bmatrix} \frac{-1}{3} & \frac{2}{3} & \frac{-1}{3} \end{bmatrix} \begin{bmatrix} 6 \\ 9 \\ 9 \end{bmatrix} = 1$$

,

and the expansion is written

$$\mathbf{x} = x_1^v \mathbf{v}_1 + x_2^v \mathbf{v}_2 + x_3^v \mathbf{v}_3 = 4\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + 1\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + 1\begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix}.$$

We can represent the process in matrix form:

$$\mathbf{x}^v = \mathbf{B}^{-1}\mathbf{x} = \begin{bmatrix} \frac{5}{3} & -\frac{1}{3} & -\frac{1}{3} \\ -\frac{1}{3} & -\frac{1}{3} & \frac{2}{3} \\ -\frac{1}{3} & \frac{2}{3} & -\frac{1}{3} \end{bmatrix} \begin{bmatrix} 6 \\ 9 \\ 9 \end{bmatrix} = \begin{bmatrix} 4 \\ 1 \\ 1 \end{bmatrix}.$$

Recall that both $\mathbf{x}^v$ and $\mathbf{x}$ are representations of the same vector, but are expanded in terms of different basis sets. (It is assumed that $\mathbf{x}$ uses the standard basis set, unless otherwise indicated.)

Epilogue

This chapter has presented a few of the basic concepts of vector spaces, material that is critical to the understanding of how neural networks work. This subject of vector spaces is very large, and we have made no attempt to cover all its aspects. Instead, we have presented those concepts that we feel are most relevant to neural networks. The topics covered here will be revisited in almost every chapter that follows.

The next chapter will continue our investigation of the topics of linear algebra most relevant to neural networks. There we will concentrate on linear transformations and matrices.

Further Reading

[Brog91] W. L. Brogan, *Modern Control Theory*, 3rd Ed., Englewood Cliffs, NJ: Prentice-Hall, 1991.

This is a well-written book on the subject of linear systems. The first half of the book is devoted to linear algebra. It also has good sections on the solution of linear differential equations and the stability of linear and nonlinear systems. It has many worked problems.

[Stra76] G. Strang, *Linear Algebra and Its Applications*, New York: Academic Press, 1980.

Strang has written a good basic text on linear algebra. Many applications of linear algebra are integrated into the text.

Exercises

E5.1 Consider again the perceptron described in Problem P5.1. If $b \neq 0$, show that the decision boundary is not a vector space.

E5.2 What is the dimension of the vector space described in Problem P5.1?

E5.3 Consider the set of all continuous functions that satisfy the condition $f(0) = 0$. Show that this is a vector space.

E5.4 Show that the set of 2×2 matrices is a vector space.

E5.5 Consider a perceptron network, with the following weights and bias.

$$\mathbf{W} = \begin{bmatrix} 1 & 0 & -1 \end{bmatrix}, \; b = 0 .$$

i. Write out the equation for the decision boundary.

ii. Show that the decision boundary is a vector space. (Demonstrate that the 10 criteria are satisfied for any point on the boundary.)

iii. What is the dimension of the vector space?

iv. Find a basis set for the vector space.

E5.6 The three parts to this question refer to subsets of the set of real-valued continuous functions defined on the interval [0,1]. Tell which of these subsets are vector spaces. If the subset is not a vector space, identify which of the 10 criteria are not satisfied.

i. All functions such that $f(0.5) = 2$.

ii. All functions such that $f(0.75) = 0$.

iii. All functions such that $f(0.5) = -f(0.75) - 3$.

E5.7 The next three questions refer to subsets of the set of real polynomials defined over the real line (e.g., $3 + 2t + 6t^2$). Tell which of these subsets are vector spaces. If the subset is not a vector space, identify which of the 10 criteria are not satisfied.

i. Polynomials of degree 5 or less.

ii. Polynomials that are positive for positive t.

iii. Polynomials that go to zero as t goes to zero.

E5.8 Which of the following sets of vectors are independent? Find the dimension of the vector space spanned by each set. (Verify your answers to parts (i) and (iv) using the MATLAB function `rank`.)

i. $\begin{bmatrix}1\\2\\3\end{bmatrix} \quad \begin{bmatrix}1\\0\\1\end{bmatrix} \quad \begin{bmatrix}1\\2\\1\end{bmatrix}$

ii. $\sin t \quad \cos t \quad \cos(2t)$

iii. $1+t \quad 1-t$

iv. $\begin{bmatrix}1\\2\\2\\1\end{bmatrix} \quad \begin{bmatrix}1\\0\\0\\1\end{bmatrix} \quad \begin{bmatrix}3\\4\\4\\3\end{bmatrix}$

E5.9 Recall the apple and orange pattern recognition problem of Chapter 3. Find the angles between each of the prototype patterns (*orange* and *apple*) and the test input pattern (*oblong orange*). Verify that the angles make intuitive sense.

$$\mathbf{p}_1 = \begin{bmatrix}1\\-1\\-1\end{bmatrix} (orange) \qquad \mathbf{p}_2 = \begin{bmatrix}1\\1\\-1\end{bmatrix} (apple) \qquad \mathbf{p} = \begin{bmatrix}-1\\-1\\-1\end{bmatrix}$$

E5.10 Using the following basis vectors, find an orthogonal set using Gram-Schmidt orthogonalization. (Check your answer using MATLAB.)

$$\mathbf{y}_1 = \begin{bmatrix}1\\0\\0\end{bmatrix} \qquad \mathbf{y}_2 = \begin{bmatrix}1\\1\\0\end{bmatrix} \qquad \mathbf{y}_3 = \begin{bmatrix}1\\1\\1\end{bmatrix}$$

E5.11 Consider the vector space of all piecewise continuous functions on the interval [0, 1]. The set $\{f_1, f_2, f_3\}$, which is defined in Figure E15.1, contains three vectors from this vector space.

i. Show that this set is linearly independent.

ii. Generate an orthogonal set using the Gram-Schmidt procedure. The inner product is defined to be

$$(f,g) = \int_0^1 f(t)g(t)\,dt \,.$$

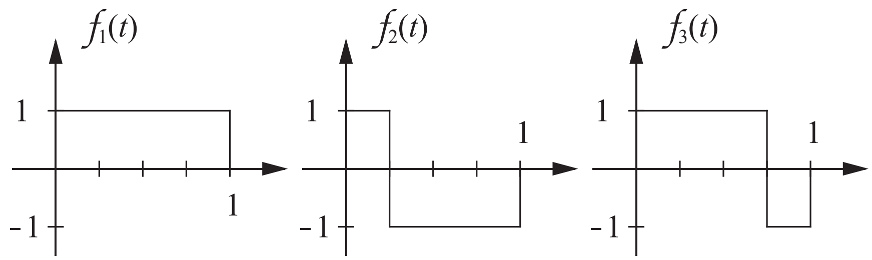


Figure E15.1 Basis Set for Exercise E5.11

E5.12 Consider the vector space of all piece wise continuous functions on the interval [0,1]. The set $\{f_1, f_2\}$, which is defined in Figure E15.2, contains two vectors from this vector space.

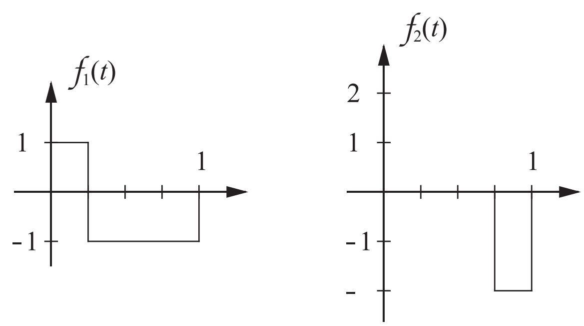


Figure E15.2 Basis Set for Exercise E5.12

i. Generate an orthogonal set using the Gram-Schmidt procedure. The inner product is defined to be

$$(f,g) = \int_0^1 f(t)g(t)\,dt \,.$$

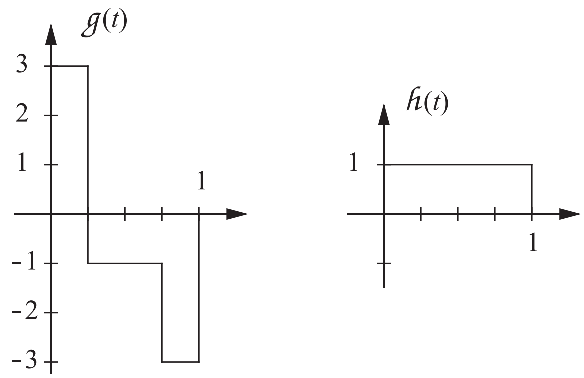


Figure E15.3 Vectors vectors g and h for Exercise E5.12 part ii.

ii. Expand the vectors g and h in Figure E15.3 in terms of the orthogonal set you created in Part 1. Explain any problems you find.

E5.13 Consider the set of polynomials of degree 1 or less. This is a linear vector space. One basis set for this space is

$$\{u_1 = 1, u_2 = t\}$$

Using this basis set, the polynomial $y = 2 + 4t$ can be represented as

$$\mathbf{y}^u = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$$

Consider the new basis set

$$\{v_1 = 1 + t, v_2 = 1 - t\}$$

Use reciprocal basis vectors to find the representation of y in terms of this new basis set.

E5.14 A vector x can be expanded in terms of the basis vectors $\{v_1, v_2\}$ as

$$x = 1v_1 + 1v_2$$

The vectors v_1 and v_2 can be expanded in terms of the basis vectors $\{s_1, s_2\}$ as

$$v_1 = 1s_1 - 1s_2$$

$$v_2 = 1s_1 + 1s_2$$

i. Find the expansion for x in terms of the basis vectors $\{s_1, s_2\}$.

ii. A vector y can be expanded in terms of the basis vectors $\{s_1, s_2\}$ as

$$y = 1s_1 + 1s_2 .$$

Find the expansion of y in terms of the basis vectors $\{v_1, v_2\}$.

E5.15 Consider the vector space of all continuous functions on the interval [0,1]. The set $\{f_1, f_2\}$, which is defined in the figure below, contains two vectors from this vector space.

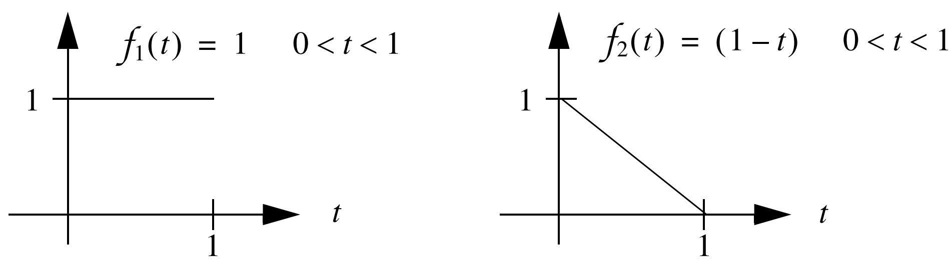


Figure E15.4 Independent Vectors for Exercise E5.15

i. From these two vectors, generate an orthogonal set $\{g_1, g_2\}$ using the Gram-Schmidt procedure. The inner product is defined to be

$$(f,g) = \int_0^1 f(t)g(t)dt .$$

Plot the two orthogonal vectors g_1 and g_2 as functions of time.

ii. Expand the following vector h in terms of the orthogonal set you created in part i., using Eq. (5.27). Demonstrate that the expansion is correct by reproducing h as a combination of g_1 and g_2 .

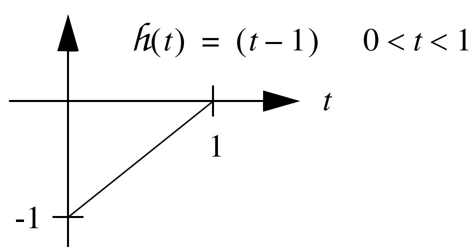


Figure E15.5 Vector h for Exercise E5.15

E5.16 Consider the set of all complex numbers. This can be considered a vector space, because it satisfies the ten defining properties. We can also define an inner product for this vector space $(\chi, y) = Re(\chi)Re(y) + Im(\chi)Im(y)$, where $Re(\chi)$ is the real part of χ, and $Im(\chi)$ is the imaginary part of χ. This leads to the following definition for norm: $\|\chi\| = \sqrt{(\chi, y)}$.

i. Consider the following basis set for the vector space described above: $v_1 = 1 + 2j$, $v_2 = 2 + j$. Using the Gram-Schmidt method, find an orthogonal basis set.

ii. Using your orthogonal basis set from part i., find vector expansions for $u_1 = 1 - j$, $u_2 = 1 + j$, and $\chi = 3 + j$. This will allow you to write χ, u_1, and u_2 as a columns of numbers $\mathbf{x}$, $\mathbf{u}_1$ and $\mathbf{u}_2$.

iii. We now want to represent the vector χ using the basis set $\{u_1, u_2\}$. Use reciprocal basis vectors to find the expansion for χ in terms of the basis vectors $\{u_1, u_2\}$. This will allow you to write χ as a new column of numbers $\mathbf{x}^u$.

iv. Show that the representations for χ that you found in parts ii. and iii. are equivalent (the two columns of numbers $\mathbf{x}$ and $\mathbf{x}^u$ both represent the same vector χ).

E5.17 Consider the vectors defined in Figure E15.6. The set $\{s_1, s_2\}$ is the standard basis set. The set $\{u_1, u_2\}$ is an alternate basis set. The vector χ is a vector that we wish to represent with respect to the two basis sets.

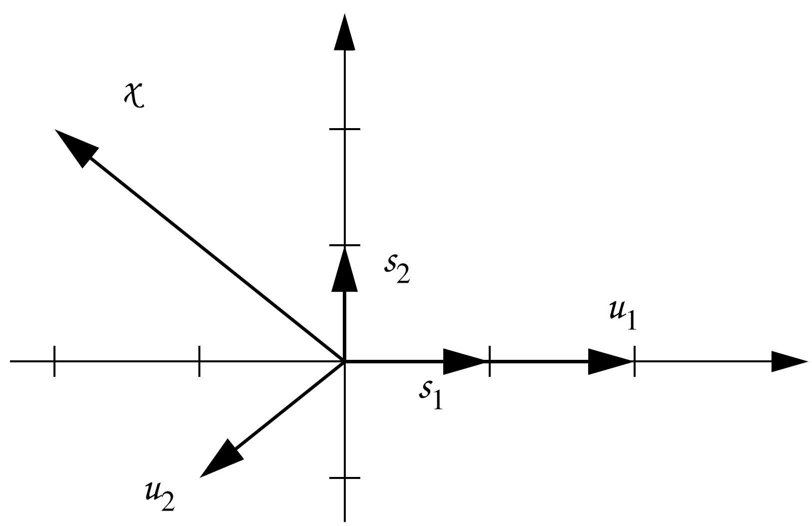


Figure E15.6 Vector Definitions for Exercise E5.17

i. Write the expansion for χ in terms of the standard basis $\{s_1, s_2\}$.

ii. Write the expansions for u_1 and u_2 in terms of the standard basis $\{s_1, s_2\}$

iii. Using reciprocal basis vectors, write the expansion for χ in terms of the basis $\{u_1, u_2\}$.

iv. Draw sketches, similar to Figure 5.2, that demonstrate that the expansions of part i. and part iii. are equivalent.

E5.18 Consider the set of all functions that can be written in the form $A\sin(t+\theta)$. This set can be considered a vector space, because it satisfies the ten defining properties.

i. Consider the following basis set for the vector space described above: $v_1 = \sin(t)$, $v_2 = \cos(t)$. Represent the vector $\chi = 2\sin(t) + 4\cos(t)$ as a column of numbers $\mathbf{x}^v$ (find the vector expansion), using this basis set.

ii. Using your basis set from part i., find vector expansions for $u_1 = 2\sin(t) + \cos(t)$, $u_2 = 3\sin(t)$.

iii. We now want to represent the vector χ of part i., using the basis set $\{u_1, u_2\}$. Use reciprocal basis vectors to find the expansion for χ in terms of the basis vectors $\{u_1, u_2\}$. This will allow you to write χ as a new column of numbers $\mathbf{x}^u$.

iv. Show that the representations for χ that you found in parts i. and iii. are equivalent (the two columns of numbers $\mathbf{x}^v$ and $\mathbf{x}^u$ both represent the same vector χ).

E5.19 Suppose that we have three vectors: $\chi, y, z \in X$. We want to add some multiple of y to χ, so that the resulting vector is orthogonal to z.

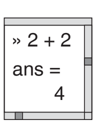
i. How would you determine the appropriate multiple of y to add to χ?

ii. Verify your results in part i. using the following vectors.

$$\mathbf{x} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \qquad \mathbf{y} = \begin{bmatrix} 1 \\ 0.5 \end{bmatrix} \qquad \mathbf{z} = \begin{bmatrix} 0.5 \\ 1 \end{bmatrix}$$

iii. Use a sketch to illustrate your results from part ii.

E5.20 Expand $\mathbf{x} = \begin{bmatrix} 1 & 2 & 2 \end{bmatrix}^T$ in terms of the following basis set. (Verify your answer using MATLAB.)



$$\mathbf{v}_1 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \qquad \mathbf{v}_2 = \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix} \qquad \mathbf{v}_3 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$$

E5.21 Find the value of a that makes $\|\chi - a y\|$ a minimum. (Use $\|\chi\| = (\chi,\chi)^{1/2}$.) Show that for this value of a the vector $z = \chi - a y$ is orthogonal to y and that

$$\|\chi - a y\|^2 + \|a y\|^2 = \|\chi\|^2 .$$

(The vector $a y$ is the projection of χ on y.) Draw a diagram for the case where χ and y are two-dimensional. Explain how this concept is related to Gram-Schmidt orthogonalization.