Chapter 3. Vector spaces

Lecture notes for MA1111

P. Karageorgis

pete@maths.tcd.ie

Linear combinations

- Suppose that $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n$ and $\boldsymbol{v}$ are vectors in $\mathbb{R}^m$.

Definition 3.1 – Linear combination

We say that $\boldsymbol{v}$ is a linear combination of $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n$, if there exist scalars $x_1, x_2, \ldots, x_n$ such that $\boldsymbol{v} = x_1\boldsymbol{v}_1 + x_2\boldsymbol{v}_2 + \ldots + x_n\boldsymbol{v}_n$.

- Geometrically, the linear combinations of a nonzero vector form a line. The linear combinations of two nonzero vectors form a plane, unless the two vectors are collinear, in which case they form a line.

Theorem 3.2 – Expressing a vector as a linear combination

Let B denote the matrix whose columns are the vectors $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n$. Expressing $\boldsymbol{v} = x_1\boldsymbol{v}_1 + x_2\boldsymbol{v}_2 + \ldots + x_n\boldsymbol{v}_n$ as a linear combination of the given vectors is then equivalent to solving the linear system $B\boldsymbol{x} = \boldsymbol{v}$.

Linear combinations: Example

- We express $\boldsymbol{v}$ as a linear combination of $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3$ in the case that

$$\boldsymbol{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}, \qquad \boldsymbol{v}_2 = \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix}, \qquad \boldsymbol{v}_3 = \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix}, \qquad \boldsymbol{v} = \begin{bmatrix} 3 \\ 2 \\ 4 \end{bmatrix}.$$

- Let B denote the matrix whose columns are $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3$ and consider the linear system $B\boldsymbol{x} = \boldsymbol{v}$. Using row reduction, we then get

$$[B\ \boldsymbol{v}] = \begin{bmatrix} 1 & 1 & 3 & 3 \\ 1 & 2 & 1 & 2 \\ 2 & 4 & 2 & 4 \end{bmatrix} \longrightarrow \begin{bmatrix} \mathbf{1} & 0 & 5 & 4 \\ 0 & \mathbf{1} & -2 & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

- In particular, the system has infinitely many solutions given by

$$x_1 = 4 - 5x_3, \qquad x_2 = -1 + 2x_3.$$

- Let us merely pick one solution, say, the one with $x_3 = 0$. Then we get $x_1 = 4$ and $x_2 = -1$, so $\boldsymbol{v} = x_1\boldsymbol{v}_1 + x_2\boldsymbol{v}_2 + x_3\boldsymbol{v}_3 = 4\boldsymbol{v}_1 - \boldsymbol{v}_2$.

Linear independence

- Suppose that $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n$ are vectors in $\mathbb{R}^m$.

Definition 3.3 – Linear independence

We say that $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n$ are linearly independent, if none of them is a linear combination of the others. This is actually equivalent to saying that $x_1\boldsymbol{v}_1 + x_2\boldsymbol{v}_2 + \ldots + x_n\boldsymbol{v}_n = 0$ implies $x_1 = x_2 = \ldots = x_n = 0$.

Theorem 3.4 – Checking linear independence in $\mathbb{R}^m$

The following statements are equivalent for each $m \times n$ matrix B.

1. The columns of B are linearly independent.
2. The system $B\boldsymbol{x} = 0$ has only the trivial solution $\boldsymbol{x} = 0$.
3. The reduced row echelon form of B has a pivot in every column.

Linear independence: Example

- We check $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3, \boldsymbol{v}_4$ for linear independence in the case that

$$\boldsymbol{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}, \qquad \boldsymbol{v}_2 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \qquad \boldsymbol{v}_3 = \begin{bmatrix} 4 \\ 5 \\ 9 \end{bmatrix}, \qquad \boldsymbol{v}_4 = \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix}.$$

- Letting B be the matrix whose columns are these vectors, we get

$$B = \begin{bmatrix} 1 & 1 & 4 & 3 \\ 1 & 2 & 5 & 1 \\ 2 & 3 & 9 & 2 \end{bmatrix} \longrightarrow \begin{bmatrix} \mathbf{1} & 0 & 3 & 0 \\ 0 & \mathbf{1} & 1 & 0 \\ 0 & 0 & 0 & \mathbf{1} \end{bmatrix}.$$

- Since the third column does not contain a pivot, we conclude that the given vectors are not linearly independent.
- On the other hand, the 1st, 2nd and 4th columns contain pivots, so the vectors $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_4$ are linearly independent. As for $\boldsymbol{v}_3$, this can be expressed as a linear combination of the other three vectors.

Span and complete sets

- Suppose that $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n$ are vectors in $\mathbb{R}^m$.

Definition 3.5 – Span and completeness

The set of all linear combinations of $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n$ is called the span of these vectors and it is denoted by $\mathrm{Span}\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$. If the span is all of $\mathbb{R}^m$, we say that $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n$ form a complete set for $\mathbb{R}^m$.

Theorem 3.6 – Checking completeness in $\mathbb{R}^m$

The following statements are equivalent for each $m \times n$ matrix B.

1. The columns of B form a complete set for $\mathbb{R}^m$.
2. The system $B\boldsymbol{x} = \boldsymbol{y}$ has a solution for every vector $\boldsymbol{y} \in \mathbb{R}^m$.
3. The reduced row echelon form of B has a pivot in every row.

Bases of $\mathbb{R}^m$

Definition 3.7 – Basis of $\mathbb{R}^m$

A basis of $\mathbb{R}^m$ is a set of linearly independent vectors which form a complete set for $\mathbb{R}^m$. Every basis of $\mathbb{R}^m$ consists of m vectors.

Theorem 3.8 – Basis criterion

The following statements are equivalent for each $m \times m$ matrix B.

1. The columns of B form a complete set for $\mathbb{R}^m$.
2. The columns of B are linearly independent.
3. The columns of B form a basis of $\mathbb{R}^m$.
4. The matrix B is invertible.

Theorem 3.9 – Reduction/Extension to a basis

A complete set of vectors can always be reduced to a basis and a set of linearly independent vectors can always be extended to a basis.

Subspaces of $\mathbb{R}^m$

Definition 3.10 – Subspace of $\mathbb{R}^m$

A subspace V of $\mathbb{R}^m$ is a nonempty subset of $\mathbb{R}^m$ which is closed under addition and scalar multiplication. In other words, it satisfies

1. $\boldsymbol{v} \in V$ and $\boldsymbol{w} \in V$ implies $\boldsymbol{v} + \boldsymbol{w} \in V$,
2. $\boldsymbol{v} \in V$ and $c \in \mathbb{R}$ implies $c\boldsymbol{v} \in V$.

- Every subspace of $\mathbb{R}^m$ must contain the zero vector. Moreover, lines and planes through the origin are easily seen to be subspaces of $\mathbb{R}^m$.

Definition 3.11 – Basis and dimension

A basis of a subspace V is a set of linearly independent vectors whose span is equal to V. If a subspace has a basis consisting of n vectors, then every basis of the subspace must consist of n vectors. We usually refer to n as the dimension of the subspace.

Column space and null space

When it comes to subspaces of $\mathbb{R}^m$, there are three important examples.

1. **Span.** If we are given some vectors $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n$ in $\mathbb{R}^m$, then their span is easily seen to be a subspace of $\mathbb{R}^m$.
2. **Null space.** The null space of a matrix A is the set of all vectors $\boldsymbol{x}$ such that $A\boldsymbol{x} = 0$. It is usually denoted by $\mathcal{N}(A)$.
3. **Column space.** The column space of a matrix A is the span of the columns of A. It is usually denoted by $\mathcal{C}(A)$.

Definition 3.12 – Standard basis

Let $\boldsymbol{e}_i$ denote the vector whose ith entry is equal to 1, all other entries being zero. Then $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_m$ form a basis of $\mathbb{R}^m$ which is known as the standard basis of $\mathbb{R}^m$. Given any matrix A, we have

$$A\boldsymbol{e}_i = i\text{th column of } A.$$

Finding bases for the column/null space

Theorem 3.13 – A useful characterisation

A vector $\boldsymbol{x}$ is in the column space of a matrix A if and only if $\boldsymbol{x} = A\boldsymbol{y}$ for some vector $\boldsymbol{y}$. It is in the null space of A if and only if $A\boldsymbol{x} = 0$.

- To find a basis for the column space of a matrix A, we first compute its reduced row echelon form R. Then the columns of R that contain pivots form a basis for the column space of R and the corresponding columns of A form a basis for the column space of A.
- The null space of a matrix A is equal to the null space of its reduced row echelon form R. To find a basis for the latter, we write down the equations for the system $R\boldsymbol{x} = 0$, eliminate the leading variables and express the solutions of this system in terms of the free variables.
- The dimension of the column space is equal to the number of pivots and the dimension of the null space is equal to the number of free variables. The sum of these dimensions is the number of columns.

Column space: Example

- We find a basis for the column space of the matrix

$$A = \begin{bmatrix} 1 & 2 & 4 & 5 & 4 \\ 3 & 1 & 7 & 2 & 3 \\ 2 & 1 & 5 & 1 & 5 \end{bmatrix}.$$

- The reduced row echelon form of this matrix is given by

$$R = \begin{bmatrix} \mathbf{1} & 0 & 2 & 0 & 0 \\ 0 & \mathbf{1} & 1 & 0 & 7 \\ 0 & 0 & 0 & \mathbf{1} & -2 \end{bmatrix}.$$

- Since the pivots of R appear in the 1st, 2nd and 4th columns, a basis for the column space of A is formed by the corresponding columns

$$\boldsymbol{v}_1 = \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix}, \qquad \boldsymbol{v}_2 = \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}, \qquad \boldsymbol{v}_4 = \begin{bmatrix} 5 \\ 2 \\ 1 \end{bmatrix}.$$

Null space: Example, page 1

- To find a basis for the null space of a matrix A, one needs to find a basis for the null space of its reduced row echelon form R. Suppose, for instance, that the reduced row echelon form is

$$R = \begin{bmatrix} \mathbf{1} & 0 & -2 & 3 \\ 0 & \mathbf{1} & -4 & 1 \end{bmatrix}.$$

- Its null space consists of the vectors $\boldsymbol{x}$ such that $R\boldsymbol{x} = 0$. Writing the corresponding equations explicitly, we must then solve the system

$$x_1 - 2x_3 + 3x_4 = 0, \qquad x_2 - 4x_3 + x_4 = 0.$$

- Once we now eliminate the leading variables, we conclude that

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 2x_3 - 3x_4 \\ 4x_3 - x_4 \\ x_3 \\ x_4 \end{bmatrix}.$$

Null space: Example, page 2

- This determines the vectors $\boldsymbol{x}$ which lie in the null space of R. They can all be expressed in terms of the free variables by writing

$$\boldsymbol{x} = \begin{bmatrix} 2x_3 - 3x_4 \\ 4x_3 - x_4 \\ x_3 \\ x_4 \end{bmatrix} = x_3 \begin{bmatrix} 2 \\ 4 \\ 1 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -3 \\ -1 \\ 0 \\ 1 \end{bmatrix} = x_3 \boldsymbol{v} + x_4 \boldsymbol{w}$$

  for some particular vectors $\boldsymbol{v}, \boldsymbol{w}$. Since the variables x_3, x_4 are both free, this means that the null space of R is the span of $\boldsymbol{v}, \boldsymbol{w}$.
- To show that $\boldsymbol{v}, \boldsymbol{w}$ form a basis for the null space, we also need to check linear independence. Suppose that $x_3\boldsymbol{v} + x_4\boldsymbol{w} = 0$ for some scalars x_3, x_4. We must then have $\boldsymbol{x} = 0$ by above. Looking at the last two entries of $\boldsymbol{x}$, we conclude that $x_3 = x_4 = 0$.

Definition 3.14 – Vector space

A set V is called a vector space, if it is equipped with the operations of addition and scalar multiplication in such a way that the usual rules of arithmetic hold. The elements of V are generally regarded as vectors.

- We assume that addition is commutative and associative with a zero element 0 which satisfies $0 + \boldsymbol{v} = \boldsymbol{v}$ for all $\boldsymbol{v} \in V$. We also assume that the distributive laws hold and that $1\boldsymbol{v} = \boldsymbol{v}$ for all $\boldsymbol{v} \in V$.
- The rules of arithmetic are usually either obvious or easy to check.
- We call V a real vector space, if the scalars are real numbers. The scalars could actually belong to any field such as $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$.
- A field is equipped with addition and multiplication, it contains two elements that play the roles of 0 and 1, while every nonzero element has an inverse. For instance, $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$ are fields, but $\mathbb{Z}$ is not.

Examples of vector spaces

1. The zero vector space $\{0\}$ consisting of the zero vector alone.
2. The vector space $\mathbb{R}^m$ consisting of all vectors in $\mathbb{R}^m$.
3. The space M_{mn} of all $m \times n$ matrices.
4. The space of all (continuous) functions.
5. The space of all polynomials.
6. The space P_n of all polynomials of degree at most n.

. . . .

- The set of all matrices is not a vector space.
- The set of polynomials of degree n is not a vector space.
- Concepts such as linear combination, span and subspace are defined in terms of vector addition and scalar multiplication, so one may naturally extend these concepts to any vector space.

Vector space concepts

- **Linear combination:** $\boldsymbol{v} = \sum x_i \boldsymbol{v}_i$ for some scalars x_i.
- **Span:** The set of all linear combinations of some vectors.
- **Complete set for $\boldsymbol{V}$:** A set of vectors whose span is equal to V.
- **Subspace:** A nonempty subset which is closed under addition and scalar multiplication. For instance, the span is always a subspace.
- **Linearly independent:** $\sum x_i \boldsymbol{v}_i = 0$ implies that $x_i = 0$ for all i.
- **Basis of $\boldsymbol{V}$:** A set of linearly independent vectors that span V.
- **Basis of a subspace $\boldsymbol{W}$:** Linearly independent vectors that span W.
- **Dimension of a subspace:** The number of vectors in a basis.

. . . .

- The zero element 0 could stand for a vector, a matrix or a polynomial. In these cases, we require that all entries/coefficients are zero.

Example: Linear combinations in P_2

- We express $f = 2x^2 + 4x - 5$ as a linear combination of

$$f_1 = x^2 - 1, \qquad f_2 = x - 1, \qquad f_3 = x.$$

- This amounts to finding scalars c_1, c_2, c_3 such that

$$f = c_1 f_1 + c_2 f_2 + c_3 f_3.$$

In other words, we need to find scalars c_1, c_2, c_3 such that

$$2x^2 + 4x - 5 = c_1 x^2 - c_1 + c_2 x - c_2 + c_3 x.$$

- Comparing coefficients, we end up with the system

$$c_1 = 2, \qquad c_2 + c_3 = 4, \qquad -c_1 - c_2 = -5.$$

This has a unique solution, namely $c_1 = 2$, $c_2 = 3$ and $c_3 = 1$. One may thus express f as the linear combination $f = 2f_1 + 3f_2 + f_3$.

Example: Linear independence in P_2

- We show that f_1, f_2, f_3 are linearly independent in the case that

$$f_1 = x^2 - x, \qquad f_2 = x^2 - 1, \qquad f_3 = x + 1.$$

- Suppose $c_1 f_1 + c_2 f_2 + c_3 f_3 = 0$ for some scalars c_1, c_2, c_3. Then

$$c_1 x^2 - c_1 x + c_2 x^2 - c_2 + c_3 x + c_3 = 0$$

  and we may compare coefficients to find that

$$c_1 + c_2 = 0, \qquad c_3 - c_1 = 0, \qquad c_3 - c_2 = 0.$$

- This gives a system that can be easily solved. The last two equations give $c_1 = c_3 = c_2$, so the first equation reduces to $2c_1 = 0$. Thus, the three coefficients are all zero and f_1, f_2, f_3 are linearly independent.

Example: Basis for a subspace in P_2

- Let U be the subset of P_2 which consists of all polynomials $f(x)$ such that $f(1) = 0$. To find the elements of this subset, we note that

$$f(x) = ax^2 + bx + c \quad \Longrightarrow \quad f(1) = a + b + c.$$

- Thus, $f(x) \in U$ if and only if $a + b + c = 0$. Using this equation to eliminate a, we conclude that all elements of U have the form

$$\begin{aligned} f(x) = ax^2 + bx + c &= (-b - c)x^2 + bx + c \\ &= b(x - x^2) + c(1 - x^2). \end{aligned}$$

- Since b, c are arbitrary, we have now expressed U as the span of two polynomials. To check those are linearly independent, suppose that

$$b(x - x^2) + c(1 - x^2) = 0$$

for some constants b, c. One may then compare coefficients to find that $b = c = 0$. Thus, the two polynomials form a basis of U.

Example: Basis for a subspace in M_{22}

- Let U be the subset of M_{22} which consists of all 2×2 matrices whose diagonal entries are equal. Then every element of U has the form

$$A = \begin{bmatrix} a & b \\ c & a \end{bmatrix}, \qquad a, b, c \in \mathbb{R}.$$

- The entries a, b, c are all arbitrary and we can always write

$$A = a \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + b \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + c \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}.$$

- This equation expresses U as the span of three matrices. To check these are linearly independent, suppose that

$$a \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + b \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + c \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} = 0$$

  for some scalars a, b, c. Then the matrix A above must be zero, so its entries a, b, c are all zero. Thus, the three matrices form a basis of U.

Coordinate vectors in a vector space

Definition 3.15 – Coordinate vectors

Suppose V is a vector space that has $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n$ as a basis. Then every element $\boldsymbol{v} \in V$ can be expressed as a linear combination

$$\boldsymbol{v} = x_1\boldsymbol{v}_1 + x_2\boldsymbol{v}_2 + \ldots + x_n\boldsymbol{v}_n$$

for some unique coefficients $x_1, x_2, \ldots, x_n$. The unique vector $\boldsymbol{x} \in \mathbb{R}^n$ that consists of these coefficients is called the coordinate vector of $\boldsymbol{v}$ with respect to the given basis.

- By definition, the coordinate vector of $\boldsymbol{v}_i$ is the standard vector $\boldsymbol{e}_i$.
- Identifying each element $\boldsymbol{v} \in V$ with its coordinate vector $\boldsymbol{x} \in \mathbb{R}^n$, one may identify the vector space V with the vector space $\mathbb{R}^n$.
- Coordinate vectors depend on the chosen basis. If we decide to use a different basis, then the coordinate vectors will change as well.

Coordinate vectors in $\mathbb{R}^n$

Theorem 3.16 – Coordinate vectors in $\mathbb{R}^n$

Suppose that $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n$ form a basis of $\mathbb{R}^n$ and let B denote the matrix whose columns are these vectors. To find the coordinate vector of $\boldsymbol{v}$ with respect to this basis, one needs to solve the system $B\boldsymbol{x} = \boldsymbol{v}$. Thus, the coordinate vector of $\boldsymbol{v}$ is given explicitly by $\boldsymbol{x} = B^{-1}\boldsymbol{v}$.

- In practice, the explicit formula is only useful when the inverse of B is easy to compute. This is the case when B is 2×2, for instance.
- The coordinate vector of $\boldsymbol{w}$ is usually found using the row reduction

$$\begin{bmatrix} \boldsymbol{v}_1 & \boldsymbol{v}_2 & \cdots & \boldsymbol{v}_n & \boldsymbol{w} \end{bmatrix} \longrightarrow \begin{bmatrix} I_n & \boldsymbol{x} \end{bmatrix}.$$

  One may similarly deal with several vectors $\boldsymbol{w}_i$ using the row reduction

$$\begin{bmatrix} \boldsymbol{v}_1 & \boldsymbol{v}_2 & \cdots & \boldsymbol{v}_n & \boldsymbol{w}_1 & \cdots & \boldsymbol{w}_m \end{bmatrix} \longrightarrow \begin{bmatrix} I_n & \boldsymbol{x}_1 & \cdots & \boldsymbol{x}_m \end{bmatrix}.$$

- The vector $\boldsymbol{v}$ is generally different from its coordinate vector $\boldsymbol{x}$. These two coincide, however, when one is using the standard basis of $\mathbb{R}^n$.