Lecture 1 - Vector spaces, basis and dimension

1 Introduction

Linear algebra forms the backbone of computer algorithms. It is the language in which mathematical ideas are expressed and executed in a computing environment. Its importance stems from the fact that most data is stored and accessed digitally in the form of arrays or matrices. So the manipulation of data depends on being able to manipulate and derive information from matrices.

The aims of this course are two-fold. We would like to understand and explore the basic numerical algorithms that form the basis of linear algebra algorithms, more informally known as matrix computations. A cursory glance at the LAPACK routines will show you how linear algebra is at the heart of all computing. Secondly, we would like to explore how linear algebra is used directly in learning from data. This is a relatively newer area, in contrast with the earlier classical stuff. We will digress from time-to-time to explore this interesting subject area which is relevant to the data science program in its own right.

We will assume the following background-

- Basic properties of determinant and trace.
- Basic operations on matrices and their properties.

For the most part, we will work with real matrices.

As mentioned earlier, the basic linear algebra part of the course which I will cover in the first semester is meant to prepare you for the deeper numerical linear algebra of the next semester. We will focus more on understanding concepts and being able to do computations. This does not mean, however, that we will not study theorem-proofs at all. While you are not required to memorize proofs at any point of time, a number of proofs in linear algebra contribute immensely to the understanding of the subject. Understanding the proofs also helps in solving problems as the constructive ideas in proofs lend themselves to use in various situations (in short, if you understand proofs, you will do better on your homework and exams). So we will take some pains to go through some of the essential proofs.

2 A short review of matrix multiplication

Recall the following facts about matrix multiplication: for

$$A = (a_{ij}) = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} \text{ and } B = (b_{ij}) = \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \dots & b_{nn} \end{pmatrix},$$

1. the $(i, j)^{th}$ entry of the product $C = AB$ is $c_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj}$ which is obtained by multiplying the i^{th} row of A with the j^{th} column of B;
2. the j^{th} column of AB equals $A \cdot (j^{th}$ column of $B)$;
3. the i^{th} column of AB equals $(i^{th}$ column of $A) \cdot B$;

Thus we may express the product AB as the matrix $(Ab'_1|Ab'_2|\dots|Ab'_n)$ where $b'_1,\dots,b'_n$ are columns of B and also as the matrix $\begin{pmatrix} a_1B \\ \hline a_2B \\ \hline \vdots \\ \hline a_nB \end{pmatrix}$ where $a_1,\dots,a_n$ are the rows of A.

4. The $(i,j)^{th}$ entry of the product ABC is obtained as $(i^{th}$ row of $A) \cdot B \cdot (j^{th}$ column of $C)$.

5. If $A = (a'_1|a'_2|\dots|a'_n)$ and $B = \begin{pmatrix} b_1 \\ \hline b_2 \\ \hline \vdots \\ \hline b_n \end{pmatrix}$ then $AB = a'_1b_1 + a'_2b_2 + \cdots + a'_nb_n$.

6. For a $n \times n$ matrix A, note that
   - if $a_{ij} = 0$ for $i \neq j$ then A is called a *diagonal* matrix.
   - if $a_{ij} = 0$ for $i > j$ then A is called an *upper triangular* matrix.
   - if $a_{ij} = 0$ for $i < j$ then A is called a *lower diagonal* matrix.

3 Vector spaces

Definition 1 *A field $\mathbb{F}$ is a set equipped with two binary operations, which we call* addition, (denoted by $+$) *and* multiplication, (denoted by$*$) *satisfying the following properties:*
w.r.t. addition: for any elements $a, b, c \in \mathbb{F}$,

- *(closure)* $a + b \in \mathbb{F}$
- *(commutativity)* $a + b = b + a$
- *(associativity)* $(a + b) + c = a + (b + c)$
- *(existence of additive identity)* $a + 0 = 0 + a = a$
- *(existence of additive inverse)* $a + (-a) = 0$

w.r.t. multiplication:

- *(closure)* $a * b \in \mathbb{F}$
- *(commutativity)* $a * b = b * a$
- *(associativity)* $(a * b) * c = a * (b * c)$
- *(existence of multiplicative identity)* $a * 1 = 1 * a = a$
- *(existence of multiplicative inverse)* $a * (a^{-1}) = 1$

and distributivity of multiplication over addition: $a * (b + c) = (a * b) + (a * c)$.

Examples:

1. The set of rational numbers $\mathbb{Q}$ is a field.
2. The set of real numbers $\mathbb{R}$ is a field.
3. The set of complex numbers $\mathbb{C}$ is a field.

Now suppose V is a non-empty set, equipped with a binary operation which we will call *addition* (and denote by '+') and a *scalar multiplication* i.e. an operation to interact with elements of the underlying field F.

Definition 2 *V is called a* vector space *over the field $\mathbb{F}$ if the elements of V satisfy:*
w.r.t. addition: for every $x, y, z \in V$,

- *(closure)* $x + y \in V$
- *(commutativity)* $x + y = y + x$
- *(associativity)* $(x + y) + z = x + (y + z)$
- *(existence of identity)* $x + 0 = 0 + x = x$
- *(existence of inverse)* $x + (-x) = 0$

w.r.t. scalar multiplication: for every scalar $c, c_1, c_2 \in \mathbb{F}$,

- *(closure)* $c \cdot x \in V$
- *(respects scalar identity)* $1 \cdot x = x$
- *(distributivity)* $(c_1 + c_2)x = c_1x + c_2x$
- *(distributivity)* $c(x + y) = cx + cy$
- *(associativity)* $(c_1c_2)x = c_1(c_2x)$

Note that this definition is simply saying the following: if V is a vector space over $\mathbb{F}$ then the addition of two elements of V lies in V and that you can expect the addition on V to behave *like it does for real numbers.* You can also expect the scalar multiplication to behave well, in the sense that you don't have to worry about distribution laws and about the order of multiplication.

Examples:

1. $\mathbb{R}$, the set of real numbers is a vector space over itself.
2. $\mathbb{R}^2$ i.e. $\mathbb{R} \times \mathbb{R}$ is a vector space over $\mathbb{R}$.
3. In general, $\mathbb{R}^n$ i.e. $\underbrace{\mathbb{R} \times \mathbb{R} \times \cdots \times \mathbb{R}}_{n \text{ times}}$ is a vector space over $\mathbb{R}$. The elements of this vector space are n-tuples written as $(x_1, x_2, \ldots, x_n)$, where each $x_i \in \mathbb{R}$.
4. Subspaces: a *subspace* is a subset of a vector space which is a vector space in its own right.
   (a) For every $n \leq m$, $\mathbb{R}^n$ is a subspace of $\mathbb{R}^m$.
   (b) Any line passing through the origin in $\mathbb{R}^2$: this is expressed in set notation as $L = \{(x, y) \mid y = mx\}$. L is a subspace of $\mathbb{R}^n$ for any $n \geq 2$.
   (c) Notice that the set of points on the line $L' : y = mx + c$, where $c \neq 0$, is not a subspace; L' does not contain the additive identity, so its not a vector space.
   (d) Set of points on the plane $ax_1 + bx_2 + cx_3 = 0$ is a subspace of $\mathbb{R}^3$ (or any $\mathbb{R}^n$ for $n \geq 3$).

We will see more examples of vector spaces and subspaces later.

3.1 Linear independence and dependence

From now on, V is a vector space over the field $\mathbb{F}$. In most examples we will see, $\mathbb{F}$ will be either $\mathbb{R}$ or $\mathbb{C}$.

Definition 3 *An expression of the form $c_1v_1 + c_2v_2 + \cdots + c_nv_n$, where c_i are scalars in $\mathbb{F}$ and v_i are vectors in V*

Definition 4 *A set of non-zero vectors $\{v_1, v_2, \ldots, v_n\}$ in a vector space v is called linearly independent if $c_1v_1 + c_2v_2 + \cdots + c_nv_n = 0$ implies $c_1 = c_2 = \cdots = c_n = 0$.*

In other words, the only way a linear combination of linearly independent vectors can be zero is if the coefficients themselves are all zero OR a linear combination of linearly independent vectors cannot be zero if the coefficients are not all zero. Look what will happen if even a single coefficient is non-zero: suppose $c_1v_1+c_2v_2+\cdots+c_nv_n = 0$, with only $c_1 \neq 0$, then the expression becomes $c_1v_1 = 0$, which will imply $v_1 = 0$, a contradiction. If two coefficients, say c_1 and c_2 are non-zero, then the above expression becomes $c_1v_1 + c_2v_2 = 0$ i.e. $v_1 = -\frac{c_2}{c_1}v_2$, i.e. v_1 is a scalar multiple of v_2. In this case we will say that v_1 and v_2 are *linearly dependent*. More generally, if in a given set of vectors, some of them can be expressed in terms of (i.e. as a linear combination of) the remaining ones, we say that the set is linearly dependent. To be precise:

Definition 5 *A set of vectors $\{v_1, v_2, \ldots, v_n\}$ is said to be* linearly dependent *if it is not linearly independent, i.e. there exist scalars $c_1, \ldots, c_n$, not all zero, such that $c_1v_1 + c_2v_2 + \cdots + c_nv_n = 0$.*

Examples: done in class.

3.2 Basis and dimension

Definition 6 *Let $S = \{v_1, v_2, \ldots, v_n\}$ be a set of vectors. Then the set of all linear combinations of elements of S, given by $T = \{c_1v_1 + c_2v_2 + \cdots + c_nv_n \mid c_i \in \mathbb{F}\}$, is called the* linear span of S.

In this case we will say that S spans T or that T is the linear span of S. When it is necessary to stress which field the scalars come from, we will say T is the $\mathbb{F}$-*linear span of* S.

Definition 7 *A set of vectors $\mathcal{B}$ is said to be a basis for a vector space V if*

- *$\mathcal{B}$ is a linearly independent set and*
- *$\mathcal{B}$ spans V.*

In other words, $\mathcal{B}$ is a basis for V and only if every element of V can be uniquely written as a linear combination of elements of $\mathcal{B}$.

(More explanation in class.)

Theorem 1 *Let V be a vector space. Then any two bases of V have the same cardinality.*

(Proof in class)

Definition 8 *The* dimension *of a vector space V is the number of elements in a basis of V.*

Corollary 1 *Any linearly independent subset of V has cardinality less than any spanning set of V. In particular, every linearly independent subset of a finite-dimensional vector space V is finite.*

(Proof in class)

More examples of vector spaces and subspaces (details in class - work out a basis and the dimension for each)

1. $\mathbb{F}^n$ for any field $\mathbb{F}$ is a n-dimensional vector space over $\mathbb{F}$.

2. $\mathbb{F}[x]$ = set of all polynomials in the variable x with coefficients from $\mathbb{F}$ is a vector space over $\mathbb{F}$

3. $\mathbb{F}^{m\times n}$ = set of $m \times n$ matrices with entries in $\mathbb{F}$ is a vector space over $\mathbb{F}$ w.r.t. the usual addition and scalar multiplication of matrices.

4. The following subsets are all subspaces of $\mathbb{F}^{n\times n}$:

   (a) the set of all symmetric matrices

   (b) the set of all skew-symmetric matrices

   (c) the set of all diagonal matrices

   (d) the set of all upper triangular matrices

   (e) the set of all trace zero matrices

Proposition 1 *Let V be a finite dimensional vector space over field $\mathbb{F}$. Then the following statements are equivalent:*

(Proof in class)

Proposition 2 *Any linearly independent subset of a finite-dimensional vector space can be extended to form a basis of the vector space.*

(Proof in class)

Corollary 2 *Any set of $n + 1$ vectors in a vector space of dimension n is linearly independent.*

(Proof in class)

Corollary 3 *If $\dim V = n$, and X is a linearly independent subset of V consisting of n elements, then X is a basis of V.*

(Proof in class)

Proposition 3 *Let V be a finite-dimensional vector space and W be a subspace of V. Then $\dim W \leq \dim V$, and $V = W$ if and only if $\dim V = \dim W$.*

(Proof in class)