CS 435 : Linear Optimization Fall 2008

Lecture 3: Linear Algebra Basics

Lecturer: *Sundar Vishwanathan*
Computer Science & Engineering Indian Institute of Technology, Bombay

Last time we saw how to solve a system of linear equations. To understand what the set of all solutions look like, we need some vocabulary. This comes from Linear Algebra. We need to wade through a few definitions first. Most abstract algebraic objects have sets and operations defined on them. Our object of interest is a vector space, which comes with two sets (both infinite) with operations. To most of you mention of a vector space would conjure a set with objects which look like $(3, 4.55, .76, 0, \ldots, 4)^T$. We do not wish to disturb this. Indeed, we will get to this picture very soon. We wish to point out that the subject of linear algebra can be developed without explicit co-ordinates and we will tread this path initially.

1 Vector Space

A *vector space* is defined as a set of vectors $\mathbf{V}$ and the real numbers $\mathbf{R}$ (called *scalars*) with the following operations defined:

- **Vector Addition:** $\mathbf{V} \times \mathbf{V} \rightarrow \mathbf{V}$, represented as $\mathbf{u} + \mathbf{v}$, where $\mathbf{u}, \mathbf{v} \in \mathbf{V}$.
- **Scalar Multiplication:** $\mathbf{R} \times \mathbf{V} \rightarrow \mathbf{V}$, represented as $a.\mathbf{u}$, where $a \in \mathbf{R}$ and $\mathbf{u} \in \mathbf{V}$.

The operations follow the following laws. First we deal with the operation on vectors, addition.

- **Abelian Group laws:**
  1. **Associativity:** $\mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w}$
  2. **Identity:** $\exists$ a zero vector $\overline{\mathbf{0}}$ which is the group identity element, i.e. $\overline{\mathbf{0}} + \mathbf{u} = \mathbf{u}$
  3. **Inverse:** $\forall \mathbf{u} \in \mathbf{V}$, there exists the additive inverse $-\mathbf{u}$ s.t. $\mathbf{u} + (-\mathbf{u}) = \overline{\mathbf{0}}$
  4. **Commutativity:** $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$

  Now we connect the other set, the reals, with the set of vectors.

- **Scalar multiplication laws:**
  1. **Multiplication by** 0**:** $0.\mathbf{u} = \overline{\mathbf{0}}$
  2. **Multiplication by -1:** $(-1).\mathbf{u} = -\mathbf{u}$
  3. **Identity multiplication:** $1.\mathbf{u} = \mathbf{u}$
  4. **Distributivity of vector sum:** $a.(\mathbf{u} + \mathbf{v}) = a.\mathbf{u} + a.\mathbf{v}$, where $a \in \mathbf{R}$ and $\mathbf{u}, \mathbf{v} \in \mathbf{V}$
  5. **Distributivity of scalar sum:** $(a + b).\mathbf{u} = a.\mathbf{u} + b.\mathbf{u}$
  6. **Associativity of scalar multiplication:** $a.(b.\mathbf{u}) = (ab).\mathbf{u}$

Check that the vectors we are familiar with obey these laws.

2 Subspace

The next important definition is that of a subspace. This is just a subset of the vectors that is also a vector space. The laws of vector addition and scalar multiplication are inherited, and a subset is a vector space in its own right provided for any two vectors u and v in the subspace and α any real, $u+v$ and αu are also in the subspace. The details of the proof are left to the reader.

$\mathbf{U} \subseteq \mathbf{V}$ is a *subspace* of $\mathbf{V}$ if $\mathbf{U}$ itself is a vector space, i.e. for all $\mathbf{u}_1, \mathbf{u}_2 \in \mathbf{U}$ and $\alpha \in \mathbf{R}$, $\mathbf{u}_1 + \mathbf{u}_2 \in \mathbf{U}$ and $\alpha.\mathbf{u}_1 \in \mathbf{U}$. For example, if $\mathbf{u} \in \mathbf{V}$, then $\mathbf{U} = \{\alpha.\mathbf{u}\ \forall \alpha \in \mathbf{R}\}$ is a subspace.
For $\alpha = -1$, $-\mathbf{u}_1 \in \mathbf{U}$ whenever $\mathbf{u}_1 \in \mathbf{U}$. Hence, $-\mathbf{u}_1 + \mathbf{u}_1 \in \mathbf{U}$. Therefore, $\overline{\mathbf{0}}$ is always a member of any subspace.

Consider the familiar vector space in two dimensions; what are the subspaces?

EXAMPLE 1 In a 2-dimensional space, any line passing through the origin is a subspace. If there is any vector in $\mathbf{U}$ that does not lie on this line, then $\mathbf{U}$ has to be the entire plane. Of course, the origin by itself, is a subspace. So there are three different types of subspaces in 2 dimensions.

3 Linear Dependence, Independence and Basis

The next definition is crucial to the development of this subject.

DEFINITION 1 *Vectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ are linearly dependent if there exist $\alpha_1, \ldots, \alpha_n \in \mathbf{R}$, not all zero, such that*

$$\sum_{i=1}^{n} \alpha_i.\boldsymbol{v}_i = \overline{\boldsymbol{0}}$$

The important point is the not all zero clause.

DEFINITION 2 *Vectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ are linearly independent if they are not linearly dependent, i.e. for $\alpha_1, \ldots, \alpha_n \in \mathbf{R}$*

$$\sum_{i=1}^{n} \alpha_i.\boldsymbol{v}_i = \overline{\boldsymbol{0}} \quad \Rightarrow \quad \alpha_i = 0,\ \forall i$$

DEFINITION 3 *Vectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ form the basis of a vector space $\mathbf{V}$ it iff:*

1. *they are linearly independent.*
2. *every other vector $\boldsymbol{w}$ which belongs to $\mathbf{V}$ can be written as*

$$\boldsymbol{w} = \sum_{i=1}^{n} \beta_i.\boldsymbol{v}_i$$

Alternatively, $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ form the basis of the vector space $\mathbf{V}$ if they are linearly independent and on adding any other vector $\boldsymbol{w} \in \mathbf{V}$ to this set, the set becomes linearly dependent.

The crucial fact is that though there can be multiple basis for the same vector space, all of them will have the same size. We will prove this beautiful fact soon. This seems obvious once you have a picture in your mind. Certainly in two and three dimensions. However, this does require a proof and the proof technique is used often in mathematics and algorithm design.

To restate this important fact, if $\mathbf{v}_1, \ldots, \mathbf{v}_n$ is a basis, and so is $\mathbf{u}_1, \ldots, \mathbf{u}_m$, then $m = n$. The number of vectors in a basis is called the *dimension* of the vector space.

One other point. There are vector spaces where the dimension is infinite. We will, however, only deal with vector spaces where the dimension is finite.

We first give a proof of the result stated above. If two sets of vectors form a basis for a vector space V, their sizes are the same.

We recall that a set X of vectors is a basis for a space V if:

1. the vectors in X are linearly independent and
2. any vector in V can be expressed as a linear combination of vectors in X.

We will prove the result by contradiction. So consider two bases S and T. We assume for a contradiction that $|S|$ is strictly smaller than $|T|$. We will then obtain a set $S' \subset T$ which is also a basis, with the property $|S'| = |S|$. Now, $|S'| = |S| < |T|$ and $S' \subset T$. Hence the vectors in $T \setminus S'$ can be expressed in terms of those in S', contradicting the fact that T is a basis.

The idea behind the proof is this. Suppose $S = \{u_1, u_2, \ldots, u_m\}$ and $T = \{v_1, v_2, \ldots, v_n\}$ with $m < n$. We will repeatedly replace elements of T with elements of S, each time maintaining the invariant that the set remains a basis. After m such substitutions, we have our desired contradiction.

For a set of vectors, the *span* of a set of vectors is the set of all vectors that can be obtained by taking linear combinations of the vectors in the set. The vectors in the set may or may not be independent. In other words, for a set of vectors $v_1, \ldots, v_k$, $span(v_1, \ldots, v_k)$ is the set $\{\sum_{i=1}^{n} \alpha_i v_i : \alpha_i \in \mathbb{R}\}$.

Exercise: Prove that for a set of vectors $v_1, \ldots, v_k$; $span(v_1, \ldots, v_k)$ is a subspace.

The following lemma shows how to replace one vector in a basis with another so that the new set is also a basis.

LEMMA 1 *Suppose $S = \{v_1, \ldots, v_n\}$ and let $x = \sum_{i=1}^{n} \alpha_i v_i$ with $\alpha_1 \neq 0$. Let $S' = \{x, v_2, \ldots, v_n\}$. Then, $span(S) = span(S')$.*

PROOF: There are two directions to prove. First observe that since $\alpha_1 \neq 0$, we can write v_1 as a linear combination of $x, v_2, \ldots, v_n$. Thus, any vector in $span(S)$ is also in $span(S')$. For the other direction, since x is a combination of v_is, every vector in $span(S')$ is also in $span(S)$. $\square$

We did something like this during the proof that Gaussian Elimination works. Do go over it.

THEOREM 1 *Suppose $S = \{u_1, u_2, \ldots, u_m\}$ and $T = \{v_1, v_2, \ldots, v_n\}$ be two sets of vectors such that each is a basis for the vector space V. Then $m = n$.*

PROOF: The proof will follow the outline above. Suppose without loss of generality, that $m < n$. Starting with the set $S_0 = S$, we do the following: replace one of the vectors of S_0 by a vector from T such that the new set, say S_1, still spans all of V. Also, if v_i was the vector in T which was added to S_0, we set $T_1 = T \setminus \{v_i\}$.

We repeat this step m times. Further, we ensure that at each step, the element removed from S_i is one of the u_j's and not the v_j's that have been added. For the generic step, assume that we have two sets S_i and T_i (with $i < m$). We maintain the invariants that S_i is a basis and $T_i = T \setminus S_i$. We show that it is always possible to obtain a set S_{i+1} such that:

1. S_{i+1} is obtained from S_i by removing one of the u_j's from S_i and adding one of the v_j's (which is from T_i) (call it x) to it.
2. The span of S_{i+1} is the same as the span of S_i.

Further, we set $T_{i+1} = T_i \setminus \{x\}$, to maintain the invariant.

We assume that we have re-numbered the u_i's and v_i's such that $S_i = \{u_{i+1}, \ldots, u_m, v_1, \ldots, v_i\}$ and $T_i = \{v_{i+1}, \ldots, v_n\}$.[1] Since S_i spans the whole of V, we have

$$v_{i+1} = \sum_{j=i+1}^{m} \alpha_j u_j + \sum_{j=1}^{i} \beta_j v_j \tag{1}$$

[1]If $i = 0$, $S_0 = S$, $T_0 = T$, and we interpret summations of the form $(\sum_{j=1}^{i} \ldots)$ as zero.

Now, at least one of α_j's must be non-zero, since otherwise we will have a non-trivial combination of v_i's yielding zero, which is not possible since T is a basis. So assume without loss of generality that α_{i+1} is non-zero. Then, by the lemma above, replacing u_{i+1} by v_{i+1} yields another basis. This is the required S_{i+1}.

After repeating this process m times, the resulting set, S_m will have m vectors from T and they span the whole of V. This is the contradiction we desire. $\square$

We have proved that given a vector space V, any basis for it will have the same size. Thus, this number is a property of V alone, and it is called the *dimension* of V.