Minimal spanning and maximal independent sets, 
Basis and Dimension

February 23, 2010

Let $S$ be a set of real $n$-vectors. In particular, $S$ can be a linear space, or its subspace. Yet, $S$ can also be just a finite set of vectors, $S = \{x^1, \ldots, x^m\}$, where $x^i = (x^i_1, \ldots, x^i_n)$ and $i = 1, \ldots, m$. For example $S$ is the set of columns of an $\langle m \times n \rangle$ real matrix $A$.

Given $k$ vectors $x^1, \ldots, x^k \in S$ and $k$ real numbers $c_1, \ldots, c_k \in \mathbb{R}$, the expression $x = c_1x^1 + \ldots + c_kx^k$ (and the obtained vector $x \in \mathbb{R}^n$ itself) is called the linear combination of $x^1, \ldots, x^k$. Note that if $S$ is a linear space then $x \in S$, too, otherwise $x$ is not necessarily is an element of $S$.

The linear combination is called trivial if every its coefficients is 0, that is $c_1 = \ldots = c_k = 0 \in \mathbb{R}$. In this case obviously $x = c_1x^1 + \ldots + c_kx^k = 0 \in \mathbb{R}^n$.

Set of vectors $I = \{x^1, \ldots, x^k\}$ is called independent if the inverse holds, too, that is if $x = c_1x^1 + \ldots + c_kx^k = 0$ implies $c_1 = \ldots = c_k = 0$. In other words, set $\{x^1, \ldots, x^k\} \subseteq S$ is linearly dependent if and only if there exists a non-trivial linear combination $x = c_1x^1 + \ldots + c_kx^k = 0$.

**Lemma 1.** Vectors $\{x^1, \ldots, x^k\} \subseteq V$ are linearly dependent if and only if ONE of them is a linear combination of the others.

**Proof.** See the textbook. \qed

**Remark 1.** Note that “One of them” but not necessarily EACH ONE. For example, let $3.5x^1 + 0x^2 + 5x^3 = 0$. Then $x^1$ and $x^3$ are such linear combinations, $x^1 = -(5/3.5)x^3 = -(10/7)x^3$ and $x^3 = -(3.5/5)x^1 = -0.7x^1$ BUT NOT $x^2$.

In general, given a non-trivial linear combination $c_1x^1 + \ldots + c_kx^k = 0$, its vector $x^i$ is a linear combination of the others if and only if the corresponding coefficient $c_i$ is not 0.

The above lemma provides an equivalent definition for the notion of linear dependence.

Standardly, a linearly dependent set is called MINIMAL if every its proper subset is linearly independent. A minimal dependent set $C \subseteq S$ sometimes (in Matroid Theory) is called a CIRCUIT.
Lemma 2. For a circuit $C = \{x^1, \ldots, x^k\}$ there exists a UNIQUE (up to a multiplier) non-trivial linear combination $c_1x^1 + \ldots + c_kx^k$ which is equal to 0; NONE of its coefficient is 0.

Proof (hint). Both claims hold, since otherwise, the corresponding linearly dependent set is not minimal.

The last property of circuits can be reformulated in several ways:
(i) EACH vector of a circuit is a linear combination of others.
(ii) If $c_1x^1 + \ldots + c_kx^k = 0$ then either each or none of the coefficients is 0.

From these we derive the following fundamental property of the circuits.

Lemma 3. Given two circuits $C_1, C_2 \subseteq S$, and a vector $x \in C_1 \cap C_2$, there exists a circuit $C_3$ such that $C_3 \subseteq C_1 \cup C_2 \setminus \{x\}$.

Proof. Try yourself, I only give you the following example as a hint. Let two circuits $\{x^1, x^2, x^3, x^4\}$ and $\{x^1, x^2, x^3, x^6\}$ be defined by the equalities $x^1 + 3x^2 + 4x^3 + 7x^4 = 0$ and $2x^1 + 7x^2 + 5x^3 + 8x^6 = 0$ respectively. Let us multiply, the first equality by $-2$ and add to the second one to get rid of the vector $x^1$.
(Or alternatively, we could multiply the equalities by $-7$ and $3$ respectively and add them to get rid of $x^2$.) In such a way we can easily prove that for each $x \in C_1 \cap C_2$ the set of vectors $C'_3 \subseteq C_1 \cup C_2 \setminus \{x\}$ is linearly dependent, though not necessarily minimal. Yet obviously, every dependent set contains a minimal dependent subset, that is a circuit. □

Analogously, a linearly independent set is called MAXIMAL if every its proper superset is linearly dependent.

One more important notion. If $x = c_1x^1 + \ldots + c_kx^k$ we say that vector $x$ is spanned by $\{x^1, \ldots, x^k\}$. Analogously, given set of vectors $S$ and its (finite) subset $S' = \{x^1, \ldots, x^k\} \subseteq S$, we say that $S$ is spanned by $S'$ if every vector $x \in S$ is spanned by $S$. In particular, $S$ can be a linear space or its subspace but also $S$ can be a finite set of vectors, for example the columns (or the rows) of a given matrix.

Lemma 4 If $I$ is an independent set and a vector $x$ is not spanned by $I$ then we can extend $I$ by $x$. In other words, the set $I' = I \cup \{x\}$ is also independent.

Proof. We got two equivalent definitions and can use each one. Let us try the first one. Consider a linear combination over $I'$ which is 0. Can it be non-trivial? The coefficient associated to $x$ must be 0, because otherwise $x$ would be spanned by $I$. All other coefficients (associated to vectors of $I$) must also be 0, because $I$ is independent. □

Try the second definition yourself.

Obviously, every set $S$ spans itself. The following lemma shows that the relation of spanning is transitive.
Lemma 5 If $S$ spans $S'$ and $S'$ spans $S''$ then $S$ spans $S''$.

Proof. Any superposition of linear combinations is a linear combination again. So we can express vectors of $S''$ via vectors of $S'$ then express vectors of $S'$ via vectors of $S$ and substitute. Do it accurately. □

Standardly, a spanning subset $S' \subseteq S$ is called MINIMAL if every proper subset $S'' \subset S'$ does not span $S$.

Lemma 6 Given a set of vectors $S$ and a subset $S' \subseteq S$, the following three properties are equivalent:

- $S'$ is independent and spans $S$,
- $S'$ is a minimal spanning subset of $S$,
- $S'$ is a maximal independent subset of $S$,
- $S'$ is a maximal subset of $S$ which contains no circuit.

Proof (sketch). Verify accurately the following three implications yourself. If $S$ is independent and spanning then it is maximal independent (since it is spanning) and it is minimal spanning (since it is independent). If $S$ is maximal independent then it is spanning, otherwise it would not be maximal. If $S$ is minimal spanning then it is independent, otherwise it would not be minimal. The last two claims are obviously equivalent just by definition of a circuit. □

Definition 1 A subset $B = \{x^1, ..., x^k\} \subseteq S$ is called a BASIS or BASE if $B$ satisfies the above four equivalent properties.

In linear algebra they usually assume that $S$ is a linear space or its subspace. Yet, in LP we will use the above definition for an arbitrary set $S$ which can be infinite (for example a linear space or its subspace), or finite (for example $S$ is a set of columns of a given matrix).

One of the most fundamental results of linear algebra is

Theorem 1 All bases of a set $S$ have the same cardinality.

(The cardinality of a finite set $S'$ is defined just as the number of its elements; notation: $|S'|$.) We shall prove Theorem 1.

Definition 2 This number is called the DIMENSION of $S$.

In principle, a linear space may contain an infinite independent set. For example, the space $C$ of all continuous functions. Yet, we will see soon that for our main “working space” $\mathbb{R}^n$ there exists no independent set of cardinality greater than $n$. Moreover, every maximum independent set is of cardinality exactly $n$. 

3
Remark 2 Given five vectors, assume that the family of all circuits is

$$\{(x_1, x^2), (x^2, x^3, x^4), (x^4, x^5)\}$$

then the family of all bases (all maximal sets of vectors which contains none of the above 3 sets) is

$$\{(x_1, x^3, x^5), (x^1, x^2, x^4), (x^2, x^3, x^5), (x^2, x^4)\}$$

There is one “basis” of cardinality 2 and three others are of cardinality 3, in contradiction with Theorem 1. How could it happen?

Explanation is simple. Our assumption was not realistic: the family of all circuits never can be $$\{(x_1, x^2), (x^2, x^3, x^4), (x^4, x^5)\}$$ since for this family the circuit axiom (see Lemma 3) does not hold.

As a simple example, let us consider four vectors

$$x^1 = (1, 2), x^2 = (1, 3), x^3 = (0, 1), x^4 = (3, 6) \in \mathbb{R}^2.$$ 

Verify that three sets of vectors

$$\{(x^1, x^4), (x^1, x^2, x^3), (x^2, x^3, x^4)\}$$

form the family of circuits and that the circuit axiom (Lemma 3) holds. Now, let us construct all minimal sets of vectors that contain no circuit. We obtain

$$\{(x^1, x^2), (x^1, x^3), (x^2, x^3), (x^2, x^4), (x^3, x^4)\},$$

that is all pairs, except $$(x^1, x^4)$$. Thus, there are five bases, all of cardinality 2.

Remark 3 Theorem 1 is not too easy to prove. The following two alternative approach may look tempting. We may try to call “basis” (i) an independent set of maximum cardinality, or (ii) a spanning set of minimum cardinality. Such definitions would automatically imply that all bases have the same cardinalities.

Moreover, for (i) it is easy to prove that every “basis” B spans the whole set S. Indeed, suppose there exists a vector $$x \in S$$ which is not spanned by B, that is $$x$$ is not a linear combination of vectors of B. Then according to Lemma 4 we could just extend B by x and get a larger independent set which contradict the status of B.

Analogously, from (ii) it is easy to derive that every “basis” B is independent. Prove yourself.

Yet both above approaches have an important disadvantage. Can you see it? Maybe there is a spanning and independent subset in S of cardinality less than |B| in the first case, or of cardinality greater than |B|, in the second case. We want to prove that this is impossible. The next Lemma will do.

Lemma 7 Given a linear space V in which I ⊆ V is an independent set and B ⊆ V is a basis. (Both subsets are finite.) Then |I| ≤ |B|
This Lemma implies immediately Theorem 1. Indeed, every basis is an independent set. Hence its cardinality can not be more than the cardinality of another basis. Hence all these cardinalities are equal.

**Proof** of Lemma 7 is based on so-called exchange property (which is important in theory of matroids). We will substitute vectors of the difference \(I \setminus B\) by vectors of \(B \setminus I\) one by one. All the obtained sets \(I, I_1, I_2, \ldots\) will be still independent and of the same cardinality as \(I\), and finally we will get \(I_l \subseteq B\). Thus \(|I| = |I_l| \leq |B|\).

The procedure is based on the above Lemmas. If already \(I \subseteq B\) then obviously \(|I| \leq |B|\) and there is nothing to prove. Otherwise, chose an arbitrary vector \(x^1 \in I \setminus B\) and delete it from \(I\). The obtained set \(I'_1 = I \setminus \{x^1\}\) can NOT be spanning in \(S\). In particular it can not span \(x^1\). Why? Because \(I\) was independent. Hence \(I'_1\) can not span \(B\) either. Why? See Lemma 5 about transitivity of spanning. Hence there is a vector \(y^1 \in B\) which is not spanned by \(I'_1\). Why? Let us add \(y^1\) to \(I'_1\) and get \(I_1\). Clearly, \(|I_1| = |I|\), because we have just exchanged \(x^1\) by \(y^1\). Further, \(I_1\) is independent again. Why? Now we can repeat the same with \(I_1\) instead of \(I\), etc. Finally we will get \(I_l \subseteq B\).

The above Lemmas imply also the following important claims:

(i) Every maximal (non-extendable) independent set is a basis.

(ii) Every independent set can be extended to a basis.

Prove these claims yourself.

Now let us consider the previous Lemma for the special case when \(I\) itself is a basis. We obtain the following property of the bases.

Given two bases \(B_1, B_2 \subseteq S\) and a vector \(b_1 \in B_1 \setminus B_2\) then there exists a vector \(b_2 \in B_2 \setminus B_1\) such that \(B_3 = (B_1 \setminus \{b_1\}) \cup \{b_2\}\) is a new basis of \(S\).

In Matroid Theory they call it the Exchange Property.

Verify that it holds for five bases from the above example.

In LP the exchange property is important, too. We will see later that the Simplex Method in fact successively exchanges the bases.

Furthermore, the exchange property immediately implies Theorem 1, that is all bases are of the same cardinality. Indeed, we proved that \(|I| \leq |B|\). But \(I\) itself can be an arbitrary basis. That implies \(|B_1| \leq |B_2|\) for every two bases \(B_1\) and \(B_2\) in \(S\). This in turn means that \(|B_1| = |B_2|\) for every two bases \(B_1\) and \(B_2\).

**Proposition 1** In a linear space \(V\) of DIMENSION \(k\) every vector \(x \in V\) is spanned by a basis \(B = \{x^1, \ldots, x^k\}\) in a UNIQUE way: \(x = c_1 x^1 + \ldots + c_k x^k \in V\).

These \(k\) numbers \(c_1, \ldots, c_k\) are referred to as the coordinates of the vector \(x\) in basis \(B\).

**Proof**. See the textbook.

Examples. In \(\mathbb{R}^3\) vectors \((001), (010), (100)\) form a canonical basis \(B\). Vectors \((110), (101), (011)\) form another basis \(B'\). Represent vectors of \(B\) as linear
combinations of vectors of $B'$ and vice versa. Every such a representation is unique. Why?

In the linear space $P_k$ of polynomials of degree $\leq k$ vectors (polynomials) $1, z, z^2, \ldots, z^k$ form a basis. Why? What is the dimension of $P_k$?

The following properties of a $n \times n$ matrix $A$ are equivalent:

1) The rows of $A$ are independent vectors in $\mathbb{R}^n$;
1') The columns of $A$ are independent vectors $\mathbb{R}^n$;
2) The rows of $A$ form a basis in $\mathbb{R}^n$;
2') The columns of $A$ form a basis in $\mathbb{R}^n$;
3) $A$ is non-singular;
4) $\det A \neq 0$;
5) there exists the inverse matrix $A^{-1}$

More generally, given a rectangular $m \times n$ real matrix $A$ and a square $k \times k$ submatrix $A'$, the following claims are equivalent:

1) the corresponding $k$ rows of $A$ are linearly independent,
1') the corresponding $k$ columns of $A$ are linearly independent,
2) $A'$ is non-singular.

Rank of $A$ is defined as the maximum number $k$ for which $A$ contains a non-singular $k \times k$ submatrix. The following claims are equivalent:

1) the corresponding $k$ rows of $A$ form a basis in the set of all rows,
1') the corresponding $k$ columns of $A$ form a basis in the set of all columns,
2) $A'$ is non-singular.

Thus for each matrix $A$ the following 3 numbers are equal: the dimension of the rows, the dimension of the columns, and $rkA$. 

6