REU APPRENTICE CLASS #2

INSTRUCTOR: LÁSZLÓ BABAI
SCRIBE: ASILATA BAPAT

Tuesday, June 28, 2011

1. Vector spaces

Definition 1.1. A *vector space* over a field F is a set V along with two operations, namely *addition* (denoted by $+$) from $V \times V$ to V, and *scalar multiplication* (denoted by $\cdot$) from $F \times V$ to V , that satisfy the following properties.

(1) Addition is associative: $(u + v) + w = u + (v + w)$ for every $u, v, w \in V$.
(2) There is an additive identity, denoted by 0.
(3) For every $v \in V$, there is an element $v^* \in V$ such that $v + v^* = 0$. The element v^* is usually denoted by $-v$.
(4) The addition operation is commutative.
(5) Pseudo-associativity holds: $\alpha \cdot (\beta \cdot v) = (\alpha\beta) \cdot v$ for every $\alpha, \beta \in F$ and $v \in V$.
(6) For every $\alpha, \beta \in F$ and $v \in V$, we have $(\alpha + \beta) \cdot v = \alpha \cdot v + \beta \cdot v$.
(7) For all $\alpha \in F$ and $u, v \in V$, we have $\alpha \cdot (u + v) = \alpha \cdot u + \alpha \cdot v$.
(8) Normalisation axiom holds: $1 \cdot v = v$ for every $v \in V$.

Proposition 1.2. *For every $v \in V$, the equation $0 \cdot v = 0$ holds.*

Proof. Let $w = 0 \cdot v$. Observe that

$$w = 0 \cdot v = (0 + 0) \cdot v = 0 \cdot v + 0 \cdot v = w + w.$$

Now adding $-w$ to both sides, we get $0 = w$. □

Exercise 1.3. For every $\alpha \in F$, the equation $\alpha \cdot 0 = 0$ holds.

The converse of these last two statement also holds:

Exercise 1.4. $(\forall \alpha \in F)(\forall v \in V)(\alpha v = 0 \Leftrightarrow \alpha = 0 \text{ or } v = 0)$

Recall that a list of vectors $v_1, \ldots, v_k$ is called *linearly independent*, if there is no nontrivial linear combination of the vectors that evaluates to the zero vector.

Definition 1.5. If $v_1, \ldots, v_k \in V$ is a list of vectors, then their *span*, denoted $\text{Span}\{v_1, \ldots, v_k\}$, is the set $\{\sum_{i=1}^{k} \alpha_i v_i \mid \alpha_i \in F\}$. A vector u is said to be *dependent* on the vectors $v_1, \ldots, v_k$ if $u \in \text{Span}\{v_1, \ldots, v_k\}$.

The span of an infinite list is the set of all linear combinations of all finite sublists.

Exercise 1.6. If $v_1, \ldots, v_k$ is a linearly independent list of vectors, then so is every sublist.

Observe that the list consisting only of $0 \in V$ is linearly dependent, because $1 \cdot 0 = 0$. However for all $v \neq 0$, the list consisting only of v is linearly independent.

Definition 1.7. Let V be a vector space over a field F. Then $U \subseteq V$ is said to be a subspace if U is a vector space over F under the same operations as that of V. We indicate this circumstance by the notation $U \leq V$.

Proposition 1.8. *A subset $U \subseteq V$ is a subspace if and only if it has the following properties:*

(1) *The set U is nonempty. (Equivalently, $0 \in U$.)*
(2) *The set U is closed under addition.*
(3) *The set U is closed under multiplication by scalars.*

The *dimension* of a subspace $U \leq V$ is defined to be the rank of the set U.

A set of vectors whose span is equal to the entire vector space is known as a *set of generators* of the vector space.

Definition 1.9. A list $(b_1, \ldots, b_\ell)$ of vectors is called a *basis* if it is a linearly independent set of generators.

We have the following examples of bases:

(1) The standard basis in $\mathbb{R}^n$, consisting of vectors $e_i = (0, \ldots, 0, 1, 0, \ldots, 0)$, where the 1 is in the ith position.
(2) The space of polynomials with real coefficients ($\mathbb{R}[x]$) has $1, x, x^2, \ldots$ as a basis.
(3) As a vector space over $\mathbb{R}$, the set $\{1, i\}$ forms a basis of $\mathbb{C}$.
(4) The real numbers form vector space over the rational numbers. This space has uncountably large dimension. It has a basis, but this basis has uncountably many elements and we cannot describe it explicitly. This is called a Hamel basis.

Theorem 1.10. *Every vector space has a basis.*

We will prove this theorem via two lemmas.

Lemma 1.11. *If $v_1, \ldots, v_k$ are linearly independent, and $v_1, \ldots, v_k, w$ are linearly dependent, then w depends on $w_1, \ldots, w_k$.*

Proof. Since $v_1, \ldots, v_k, w$ are linearly dependent, there exists a non-trivial linear combination that evaluates to 0. That is, there are elements $\alpha_1, \ldots, \alpha_k, \beta \in F$ such that:

$$\alpha_1 v_1 + \cdots + \alpha_k v_k + \beta w = 0.$$

If $\beta \neq 0$, then we can express w as the following linear combination:

$$w = -\beta^{-1}\alpha_1 v_1 - \cdots - \beta^{-1}\alpha_k v_k,$$

which proves the lemma.

If $\beta = 0$, then we see that $\sum_{i=1}^{k} \alpha_i v_i = 0$. However since $v_1, \ldots, v_k$ are linearly independent, we have $\alpha_i = 0$ for every i. This is a contradiction, since we assumed that at least one of the $\alpha_1, \ldots, \alpha_k, \beta$ is non-zero. Thus the proof is complete. □

Lemma 1.12. *A maximal linearly independent set in a vector space is always a basis.*

Proof. Consider some maximal linearly independent set S. If S is not a basis, then $\mathrm{Span}(S)$ is strictly smaller than V. So there is a vector $w \in V$ such that $w \notin \mathrm{Span}(S)$. In particular, w does not depend on S. By the contrapositive statement of the previous lemma, we see that $S \cup \{w\}$ is linearly independent. This is a contradiction because S was assumed to be a maximal linearly independent set. □

Idea of proof of the theorem. By using the two lemmas, we observe that it is enough to exhibit a maximal linearly independent set of vectors.

If a given set of linearly independent vectors is not maximal, we can add a vector and still keep it linearly independent, by the definition of (non)maximality. Repeating this will give us a basis if there is a finite upper bound on the size of linearly independent sets.

If there is no such upper bound, we refer to Zorn's lemma from set theory which implies that any linearly independent set can be extended to a maximal one. □

Exercise 1.13. Show that $v_1, \ldots, v_k$ is linearly independent if and only if for every i, the vector v_i does not depend on the rest.

2. The Fibonacci space

Consider the vector space V consisting of sequences $(a_0, a_1, \ldots)$ of real numbers. Consider the elements $f_i = (0, \ldots, 0, 1, 0, \ldots)$, which has a 1 in the ith position. The set of all f_i appears to be a basis for V, but it is not a basis. In fact, the span of $\{f_1, f_2, \ldots\}$ consists of all sequences that are "eventually zero". That is, all sequences that have finitely many non-zero terms.

Say that $s \in V$ is a *Fibonacci-type* sequence if for every $n \in \mathbb{N}$, we have $s_{n+2} = s_{n+1} + s_n$. For example, the Fibonacci numbers form a Fibonacci-type sequence.

Exercise 2.1. The set Fib, consisting of all the Fibonacci-type sequences, is a subspace of V with dimension 2. Show that the sequences s and t form a basis of Fib, where $s_0 = 1$, $s_1 = 0$, $t_0 = 0$ and $t_1 = 1$.

Problem 22. Find a basis of Fib consisting of two geometric progressions: $(1, r, r^2, \ldots)$ and $(1, s, s^2, \ldots)$. Then if F_n is the nth Fibonacci number, we will have $F_n = \alpha r^n + \beta s^n$. Find r, s, α, β.

3. The First Miracle of Linear Algebra

Theorem 3.1 (First miracle of linear algebra)**.** *If $v_1, \ldots, v_k$ are linearly independent, and $v_1, \ldots, v_k \in Span\{w_1, \ldots, w_l\}$, then $k \leq \ell$.*

We prove this via the following lemma.

Lemma 3.2 (Steinitz exchange principle)**.** *If $v_1, \ldots, v_k \in Span\{w_1, \ldots, w_\ell\}$, then for every i there exists a j such that $v_1, \ldots, v_{i-1}, w_j, v_{i+1}, \ldots, v_k$ is linearly independent.*

Proof. (Proof by contradiction.) Suppose this is not true. Then there is some i such that none of the w_j can be exchanged for v_i. Therefore w_j depends on $v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_k$ for every j. Since $v_i \in \text{Span}\{w_1, \ldots, w_\ell\}$, we infer $v_i \in \text{Span}\{v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_k\}$. This is a contradiction. □

Proof of the first miracle of linear algebra. Use the Steinitz exchange principle successively on $v_1, \ldots, v_k$ to get a new list, namely $w_{j_1}, \ldots, w_{j_k}$, which is linearly independent. Therefore $w_{j_a} \neq w_{j_b}$ for $a \neq b$. From this it follows that $k \leq \ell$. □

As a corollary to this theorem, we deduce the fundamental fact that **all bases have equal size.** This is equivalent to the Firsat Miracle.

If $S \subset V$ is an arbitrary subset, then a basis of this set is a set $B \subset S$ such that $S \subseteq \text{Span}(B)$ and B is linearly independent.

Exercise 3.3. All bases of a set of vectors have equal size.

4. Matrices; the Second Miracle of Linear Algebra

Let $A = (\alpha_{ij})_{k \times n}$ be a $k \times n$, with the entries α_{ij} lying in a fixed field F.

Definition 4.1. The *column-rank* of the matrix A is the rank of the list of its column vectors, which lie in F^k. The *row-rank* of the matrix A is the rank of the list of its row vectors, which lie in F^n.

Theorem 4.2 (Second miracle of linear algebra)**.** *The row-rank of a matrix is equal to its column-rank.*

Definition 4.3. An *elementary column operation* for a matrix $A = (\underline{a}_1, \ldots, \underline{a}_n)$ (where $\underline{a}_i$ is the ith column) replaces some $\underline{a}_i$ by $\underline{a}_i - \lambda \underline{a}_j$, where $\lambda \in F$ and $i \neq j$. Elementary row operations are defined analogously for rows.

Problem 23. Elementary row and column operations do not change either the row-rank or the column-rank of the matrix.

The *column space* is defined as the span of the column vectors inside F^k, while the *row space* is defined as the span of the row vectors inside F^n. In fact, elementary row operations do not change the column space, and elementary column operations do not change the row space.

Exercise 4.4. The column rank is exactly the dimension of the column space.

Proof sketch of the second miracle of linear algebra. We will demonstrate the process of *Gaussian elimination* via an example, on which we carry out column operations:

$$A = \begin{pmatrix} 1 & 2 & 3 & 3 \\ 2 & 3 & 5 & 4 \\ 3 & 7 & 0 & 3 \end{pmatrix}$$

$$\mapsto \begin{pmatrix} 1 & 0 & 0 & 0 \\ 2 & -1 & -1 & -2 \\ 3 & 1 & -9 & -6 \end{pmatrix} \quad \text{[subtracted multiples of the first column from other columns]}$$

$$\mapsto \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 5 & 1 & -10 & -8 \end{pmatrix} \quad \text{[subtracted multiples of the second column from other columns]}$$

$$\mapsto \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -10 & 0 \end{pmatrix} \quad \text{[subtracted multiples of the third column from other columns]}$$

At the last step it is clear that the row rank is equal to the column rank. This algorithm works in general. □

Definition 4.5. The *transpose* of a matrix A, denoted A^t, is defined by $(A^t)_{ij} = A_{ij}$.

The row rank of a matrix A is equal to the column rank of A^t and vice-versa. Therefore the rank of A is equal to the rank of A^t.

5. Systems of Linear Equations

Consider a system of k linear equations in n unknowns: $\alpha_{i1}x_1 + \cdots + \alpha_{in}x_n = \beta_i$ for $1 \leq i \leq k$. We can consider the matrix $A = (\alpha_{ij})$ with columns $\underline{a}_i$, and the column vector $\underline{b}$ consisting of β_i.

Then the given system of equations is equivalent to the following vector equation:

$$x_1\underline{a}_1 + \cdots + x_n\underline{a}_n = \underline{b}.$$

Therefore a solution exists if and only if $\underline{b}$ is in the span of the column vectors of A.

First consider the homogeneous case, when $\beta_i = 0$ for every i. Then a nontrivial solution exists if and only if the column vectors $\underline{a}_i$ are linearly dependent, which is true if and only if $\operatorname{rank}(A) < n$.

Definition 5.1. A $k \times n$ matrix A has *full row rank* if $\operatorname{rank}(A) = k$. Similarly A has *full column rank* if $\operatorname{rank}(A) = n$.

Exercise 5.2. A matrix has full column-rank if and only if the columns are linearly independent.

So our system of homogeneous linear equations has a nontrivial solution if and only if the matrix does not have full column rank.

Problem 24. The rows of a matrix are linearly independent if and only if the columns span F^k. Similarly the columns are linearly independent if and only if the rows span F^n.

Exercise 5.3. A solution to the equation $\sum_{i=1}^{n} x_i\underline{a}_i = \underline{b}$ exists if and only if

$$\operatorname{rank}(A) = \operatorname{rank}(A|\underline{b}).$$