AFCS / Spring 2014

Vector Spaces

Prof.Dr. Ferucio Laurenţiu Ţiplea

"Al. I. Cuza" University of Iaşi,

Department of Computer Science,

Iasi 740083, Romania,

URL: http://www.infoiasi.ro/~fltiplea

E-mail **fltiplea@mail.dntis.ro**

Contents

1. Definitions and examples

Definition 1 Let $(F, +, -, 0, \circ, ', e)$ be a field. A vector space over F is an algebraic system $(V, \oplus, \ominus, \mathbf{0}, \cdot)$ which consists of a commutative group $(V, \oplus, \ominus, \mathbf{0})$ and a function $\cdot : F \times V \to V$ such that:

1. $\alpha \cdot (x \oplus y) = \alpha \cdot x \oplus \alpha \cdot y$, for any $\alpha \in F$ and $x, y \in V$;
2. $(\alpha + \beta) \cdot x = \alpha \cdot x \oplus \beta \cdot x$, for any $\alpha, \beta \in F$ and $x \in V$;
3. $(\alpha \circ \beta) \cdot x = \alpha \cdot (\beta \cdot x)$, for any $\alpha, \beta \in F$ and $x \in V$;
4. $e \cdot x = x$, for any $x \in V$.

The elements of V are called vectors, the elements of F are called scalars, and F is called the field of scalars of V. The operation $\oplus$ is called the vector addition and the operation $\cdot$ is called the scalar multiplication.

1. Definitions and examples

Remark 1 To simplify the notation, we will denote the operations of F by $(F, +, -, 0, \cdot, ', 1)$ and the operations of V by $(V, +, -, 0, \cdot)$. Moreover, the symbol of the operation $\cdot$ will be mostly omitted. Therefore, the axioms of V can be rewritten as follows:

1. $\alpha(x + y) = \alpha x + \alpha y$, for any $\alpha \in F$ and $x, y \in V$;
2. $(\alpha + \beta)x = \alpha x + \beta x$, for any $\alpha, \beta \in F$ and $x \in V$;
3. $(\alpha\beta)x = \alpha(\beta x)$, for any $\alpha, \beta \in F$ and $x \in V$;
4. $1x = x$, for any $x \in V$.

Vector subtraction is defined by $x - y = x + (-y)$, for any $x, y \in V$.

The vector space which consists of the only element 0 is called the trivial vector space (it is unique up to isomorphism).

1. Definitions and examples

Example 1

1. Let F be a field and $n \geq 1$. Denote by F^n the set of all n-dimensional vectors over F. Define vector addition by

$$(a_1, \ldots, a_n) + (b_1, \ldots, b_n) = (a_1 + b_1, \ldots, a_n + b_n)$$

   and scalar multiplication by

$$b(a_1, \ldots, a_n) = (ba_1, \ldots, ba_n),$$

   for any $(a_1, \ldots, a_n), (b_1, \ldots, b_n) \in F^n$ and $b \in F$.
   With these operations, F^n is a vector space over F. If we identify F^1 with F, then F can be viewed as a vector space over itself.

2. The set of all $m \times n$ matrices over F, denoted ${}^mF^n$, can be organized as a vector space over F. Vector addition is matrix addition, and scalar multiplication is the usual multiplication with scalars.

1. Definitions and examples

Example 2

1. $\mathbf{Q}^n$, $\mathbf{R}^n$, and $\mathbf{C}^n$ are vector spaces.

2. $\mathbf{C}$ can be viewed as a vector space over $\mathbf{R}$, and both $\mathbf{C}$ and $\mathbf{R}$ can be viewed as vector spaces over $\mathbf{Q}$.

3. The set of all functions from $\mathbf{R}$ to $\mathbf{R}$, together with the addition $f+g$ and scalar multiplication αf ($(\alpha f)(x) = \alpha f(x)$, for any x), form a vector space over $\mathbf{R}$.

1. Definitions and examples

Proposition 1 Let V be a vector space over a field F. Then, for any $x, y \in V$ and $\alpha, \beta \in F$, the following properties hold

1. $0x = 0$;
2. $(-1)x = -x$;
3. $(-\alpha)x = \alpha(-x) = -\alpha x$;
4. $\alpha 0 = 0$;
5. if $\alpha x = 0$, then $\alpha = 0$ or $x = 0$;
6. if $\alpha x = \alpha y$, then $\alpha = 0$ or $x = y$;
7. if $\alpha x = \beta x$, then $\alpha = \beta$ or $x = 0$.

1. Definitions and examples

Definition 2 Let V and U be vector spaces over a field F. We say that U is a subspace of V, denoted $U \leq V$, if $U \subseteq V$ and the restriction of V's operations to U coincide with U's operations.

Example 3

1. If V is a vector space over F, then $\{0\}$ and V are subspaces of V.
2. Let F be a field and $n \geq 1$. The set U of all vectors of F^n whose first coordinate is 0 is a subspace of F^n. When $n \geq 2$, this subspace can be identified with F^{n-1}.

1. Definitions and examples

Let V be a vector space over a field F, $x_1, \ldots, x_k \in V$, and $\alpha_1, \ldots, \alpha_k \in F$, where $k \geq 1$. An expression

$$\alpha_1 x_1 + \ldots + \alpha_k x_k$$

is called a linear combination of $x_1, \ldots, x_k$.

The set of all linear combinations of $x_1, \ldots, x_k$ forms a subspace of V; this subspace is called the subspace generated by $x_1, \ldots, x_k$. It is usually denoted by $\langle x_1, \ldots, x_k \rangle_V$ or $\langle x_1, \ldots, x_k \rangle$. Therefore,

$$\langle x_1, \ldots, x_k \rangle = \{\alpha_1 x_1 + \cdots + \alpha_k x_k | \alpha_1, \ldots, \alpha_k \in F\}.$$

If $x = \sum \alpha_i x_i$ then we say that x is a linear combination of $x_1, \ldots, x_k$ or that x is linearly dependent of $x_1, \ldots, x_k$.

1. Definitions and examples

Definition 3 Let V be a vector space over a field F. The vectors $x_1, \ldots, x_k$ from V are called linearly dependent if there exist $\alpha_1, \ldots, \alpha_k \in F$, not all 0, such that $\sum \alpha_i x_i = 0$.

If $x_1, \ldots, x_k$ are not linearly dependent, then they are called linearly independent. That is, $x_1, \ldots, x_k$ are linearly independent if for any $\alpha_1, \ldots, \alpha_k \in F$, the relation $\sum \alpha_i x_i = 0$ leads to $\alpha_1 = \cdots = \alpha_k = 0$.

Remark 2 Let V be a vector space over a field F.

1. $x \in V$ is linearly independent iff $x \neq 0$.
2. If $x_1, \ldots, x_k \in V$ are linearly independent, then $x_i \neq 0$, for any i. Moreover, $x_i \neq x_j$, for any $i \neq j$.

Proposition 2 Let V be a vector space over a field F. $x_1, \ldots, x_k$ from V are linearly dependent iff there exists $1 \leq i \leq k$ such that x_i is a linear combination of the other vectors.

2. Basis and dimension

Definition 4 Let V be a non-trivial vector space over a field F. A finite subset $B \subseteq V$ is called a basis of V if it is linearly independent and generates V (each element in V is a linear combination of vectors in B).

Remark 3

- If $x_1, \ldots, x_k$ form a basis for V, then $x_i \neq x_j$, for any $i \neq j$. Therefore, $\{x_1, \ldots, x_k\}$ has exactly k vectors.
- We have considered only finite basis. There are approaches for infinite basis too.

2. Basis and dimension

Example 4

1. Let F be a field and $n \geq 1$. The vector space F^n can be generated by

$$\begin{array}{rcl}
\mathbf{e_1} & = & (1, 0, 0, \ldots, 0, 0) \\
\mathbf{e_2} & = & (0, 1, 0, \ldots, 0, 0) \\
& \ldots & \\
\mathbf{e_n} & = & (0, 0, 0, \ldots, 0, 1).
\end{array}$$

2. Let F be a field and $m, n \geq 1$. The vector space ${}^mF^n$ can be generated by E_{ij}, where

$$E_{ij}(u, v) = \begin{cases} 1, & \text{if } u = i \text{ and } v = j \\ 0, & \text{otherwise,} \end{cases}$$

for any $i, u \in \{1, \ldots, m\}$ and $v, j \in \{1, \ldots, n\}$.

2. Basis and dimension

Theorem 1 Let V be a vector space over a field F. $B = \{x_1, \ldots, x_k\} \subseteq V$ is a basis of V iff any $x \in V$ can be uniquely written as a linear combination of vectors in B.

Corollary 1 If A and B are finite linearly independent sets that generate a vector space V, then $|A| = |B|$.

Definition 5 Let V be a vector space over a field F. V is called finite dimensional if there exists a (finite) basis B for V. In this case, $|B|$ is called the dimension of V, denoted $dim(V)$. If V is not finite dimensional then it is called infinite dimensional.

Example 5

1. $dim(F^n) = n$ and $dim({}^mF^n) = mn$.
2. $F^{\mathbf{N}}$ is an infinite dimensional.

3. Applications: error detecting and correcting codes

Entities involved in information transmission:

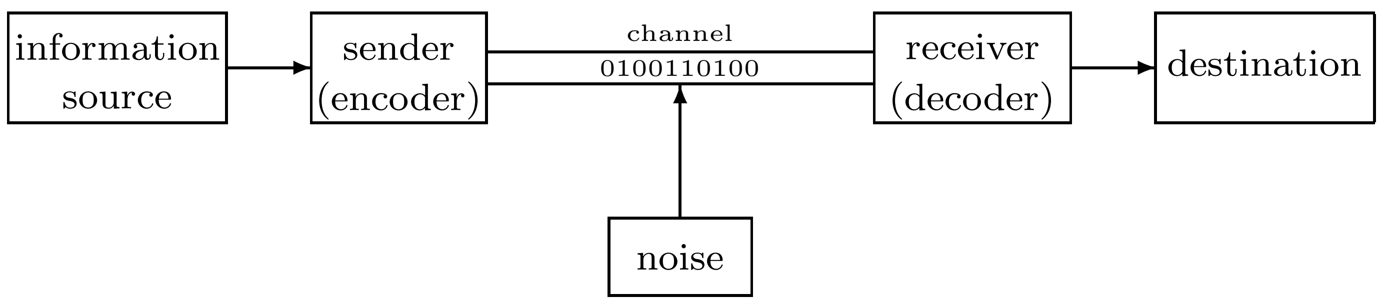
- sender (encoder);
- receiver (decoder);
- channel.

Examples of entities involved in information transmission:

- satellite station, Earth station, atmosphere;
- emission device, reception device, telephone cable.

3. Applications: error detecting and correcting codes

Main problem: noise



Main question: develop codes capable of error detection and correction

3. Applications: error detecting and correcting codes

We will use only bloc binary codes.

Transmission channels can be classified into:

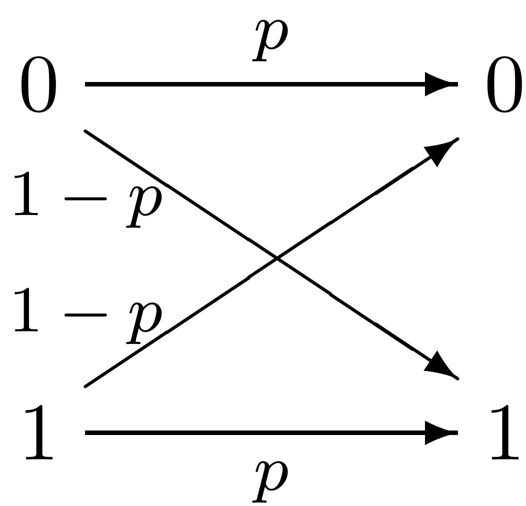
- noiseless channels (also called perfect channels);
- noise channels, which can be
  - symmetric – the probability that a bit is (correctly) received is the same for both bits;
  - asymmetric – it is not symmetric.

We will use only binary symmetric channels (BSC). Basic assumptions about them:

- BSCs do not change the length of the binary sequence transmitted through them;
- receiving order of the bits = sending order of the bits.

3. Applications: error detecting and correcting codes

The reliability of a BSC is a real number $p \in (0, 1)$ which gives the probability that the bit b received is the bit b sent.



We may consider only BSCs with reliability $1/2 < p < 1$.

3. Applications: error detecting and correcting codes

Let $C_1 = \{00, 01, 10, 11\}$. With such a code, no error can be detected (but they may occur).

Let $C_2 = \{000, 011, 101, 110\}$ (obtained from C_2 by adding the parity bit). With such a code, any singular error is detected.

Definition 6 The information ratio of a code C of length n is

$$ri(C) = \frac{log_2|C|}{n}.$$

$ri(C_1) = 1$ and $ri(C_2) = 2/3$.

3. Applications: error detecting and correcting codes

Case analysis: channel reliability $p = 1 - 10^{-8}$, transmission rate 10^7 bits/sec:

- Let $C = \{0,1\}^{11}$. A simple computation shows that

$$\frac{11}{10^8} \cdot \frac{10^7}{11} = 0.1 \text{ code words/sec}$$

  with exact one undetected error will be transmitted. This means 8640 code words/day !!!

- Let C' be obtained from C by adding the parity bit. A simple computation shows that

$$\frac{66}{10^{16}} \cdot \frac{10^7}{12} \approx \frac{5.5}{10^9} \text{code words/sec}$$

  with undetected errors will be transmitted. This means a code word/2000 days !!!

3. Applications: error detecting and correcting codes

Let C be a code of length n, $w \in \{0,1\}^n$ and $v \in C$. Let d be the number of positions on which w and v disagree. Then, the probability that v was sent when w was received is

$$\phi_p(v,w) = p^{n-d}(1-p)^d,$$

where p is the channel reliability.

In practice, we know w but we do not know v. Usually, we choose v such that the probability

$$\phi_p(v,w) = max\{\phi_p(u,w) | u \in C\}$$

is minimized. Of course, v might not be unique.

3. Applications: error detecting and correcting codes

Theorem 2 Let C be a code of length n, $v_1, v_2 \in C$, and $w \in \{0,1\}^n$, and d_1 (d_2) be the number of positions on which v_1 and w (v_2 and w, respectively), disagree. Then,

$$\phi_p(v_1, w) \leq \phi_p(v_2, w) \quad \Leftrightarrow \quad d_1 \geq d_2$$

(it is assumed that the channel reliability satisfies $1/2 < p < 1$).

3. Applications: error detecting and correcting codes

We will work exclusively with the vector space F_2^n, where $F_2 = \mathbf{Z}_2$.
Vector addition and scalar multiplication are given by:

- $x_1 \cdots x_n + y_1 \cdots y_n = (x_1 + y_1) \cdots (x_n + y_n)$;
- $\alpha(x_1 \cdots x_n) = (\alpha \cdot x_1) \cdots (\alpha \cdot x_n)$,

where $\alpha, x_i, y_i \in F_2$, $x_i + y_i$ is the addition modulo 2, and $\alpha \cdot x_i$ is given by

$$0 \cdot 0 = 0 \cdot 1 = 1 \cdot 0 = 0 \text{ and } 1 \cdot 1 = 1.$$

Definition 7 Let $v \in \{0,1\}^*$. The Hamming weigth of v, denoted $Hw(v)$, is the number of 1s in v.

Definition 8 Let $v, w \in \{0,1\}^n$, for some n. The Hamming distance of v and w, denoted $Hd(v,w)$, is $Hd(v,w) = Hw(v+w)$.

3. Applications: error detecting and correcting codes

Proposition 3 The following properties hold true:

(1) $0 \leq Hw(v) \leq n$;

(2) $Hw(v) = 0$ iff $v = 0$;

(3) $0 \leq Hd(v, w) \leq n$;

(4) $Hd(v, w) = 0$ iff $v = w$;

(5) $Hd(v, w) = Hd(w, v)$;

(6) $Hw(v + w) \leq Hw(v) + Hw(w)$;

(7) $Hd(v, w) \leq Hd(v, u) + Hd(u, w)$;

(8) $Hw(av) = aHw(v)$;

(9) $Hd(av, aw) = aHd(v, w)$,

for any $u, v, w \in \{0, 1\}^n$ and $a \in \{0, 1\}$, where $n \geq 1$.

3. Applications: error detecting and correcting codes

Definition 9 Let C be a code of length n.

(1) C detects the error $u \in \{0,1\}^n - \{0^n\}$ if $v + u \notin C$, for any $v \in C$.

(2) C is a t-detector code if C detects any error with Hamming weight at most t, but there exists an error with Hamming weight $t+1$ that cannot be detected by C.

Definition 10 Let C be a code. The distance of C, denoted $d(C)$, is

$$d(C) = min\{Hd(v,w) | v, w \in C,\ v \neq w\}.$$

Theorem 3 Let C be a code of length n and distance d. Then,

(1) C detects all errors $u \in \{0,1\}^n - \{0^n\}$ with $Hw(u) \leq d-1$;

(2) there exists at least one error $u \in \{0,1\}^n - \{0^n\}$ with $Hw(u) = d$ that cannot be detected by C.

3. Applications: error detecting and correcting codes

Definition 11 Let C be a code of length n.

(1) C corrects the error $u \in \{0,1\}^n - \{0^n\}$ if
$Hd(v+u, v) < Hd(v+u, w)$, for any $v \in C$ şi $w \in C - \{v\}$.

(2) C is a t-corrector code if C corrects all errors with Hamming weight at most t, but there exists at least one error with Hamming weight $t+1$ that cannot be corrected by C.

Theorem 4 Let C be a code of length n and distance d. Then,

(1) C corrects all errors $u \in \{0,1\}^n - \{0^n\}$ with $Hw(u) \leq \lfloor (d-1)/2 \rfloor$;

(2) there exists at least one error $u \in \{0,1\}^n - \{0^n\}$ with
$Hw(u) = \lfloor (d-1)/2 \rfloor + 1$ that cannot be corrected by C.