1. Vector spaces

Definition 1.1. Let F be a set in which two operations $+$ and $\cdot$, called *addition* and *multiplication*, are defined. That is, to each pair (x, y) of elements of F, there correspond elements $x + y$ and $x \cdot y$ of F. Then F is a *field* if the following axioms are satisfied:

(1) $(x + y) + z = x + (y + z)$ for all $x, y, z \in F$,
(2) $x + y = y + x$ for all $x, y \in F$,
(3) there is a unique element $0 \in F$ such that $x + 0 = x$ for all $x \in F$,
(4) for every element $x \in F$ there is a unique element $-x \in F$ such that $x + (-x) = 0$,
(5) $(xy)z = x(yz)$ for all $x, y, z \in F$,
(6) $xy = yx$ for all $x, y \in F$,
(7) there is a unique element $1 \in F \setminus \{0\}$ such that $x1 = x$ for all $x \in F$,
(8) for every element $x \in F \setminus \{0\}$ there is a unique element $x^{-1} \in F$ such that $xx^{-1} = 1$, and
(9) $x(y + z) = xy + xz$ for all $x, y, z \in F$.

Examples 1.2. (1) Number fields: $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$, $\mathbb{Q}(\sqrt{2}) = \{a + b\sqrt{2} : a, b \in \mathbb{Q}\}$.

(2) The field $\mathbb{F}_p = \{0, 1, \ldots, p - 1\}$ of integers modulo a prime number p.

Definition 1.3. Let V be a set and let F be a field. Elements of V will be denoted by lowercase Latin letters $(a, b, c, \ldots)$ and referred to as *vectors*. Elements of F will be denoted by lowercase Greek letters $(\alpha, \beta, \gamma, \ldots)$ and referred to as *scalars*. Suppose that in V two operations, called *vector addition* and *scalar multiplication*, are defined. That is, to each pair (a, b) of vectors there corresponds a vector $a + b$, and to each pair (α, a), where α is a scalar and a is a vector, there corresponds a vector αa. Then V is a *vector space* over F if the following axioms are satisfied:

(1) $(a + b) + c = a + (b + c)$ for all $a, b, c \in V$,
(2) $a + b = b + a$ for all $a, b \in V$,
(3) there is a unique vector $\theta \in V$ such that $a + \theta = a$ for all $a \in V$,
(4) for every vector $a \in V$ there is a unique vector $-a \in V$ such that $a + (-a) = \theta$,
(5) $\alpha(\beta a) = (\alpha\beta)a$ for all $\alpha, \beta \in F$ and $a \in V$,
(6) $1a = a$ for all $a \in V$,
(7) $(\alpha + \beta)a = \alpha a + \beta a$ for all $\alpha, \beta \in F$ and $a \in V$, and
(8) $\alpha(a + b) = \alpha a + \alpha b$ for all $\alpha \in F$ and $a, b \in V$.

Examples 1.4. (1) The *coordinate space* F^n. The vectors of F^n are n-tuples $a = (\alpha_1, \ldots, \alpha_n)$ of elements of F. The vector addition and scalar multiplication are coordinatewise, that is, $a+b = (\alpha_1+\beta_1, \ldots, \alpha_n+\beta_n)$ and $\alpha a = (\alpha\alpha_1, \ldots, \alpha\alpha_n)$. Important partial examples are $\mathbb{R}^n$ and $\mathbb{C}^n$.

(2) The space F^X of mappings $f : X \to F$ with pointwise operations, that is, $(f + g)(x) = f(x) + g(x)$ and $(\alpha f)(x) = \alpha(f(x))$. An important partial example is $\mathbb{R}^{[a,b]}$.

(3) The space $F[x]$ of polynomials in x over F, in particular, $\mathbb{R}[x]$.

(4) The space $\mathcal{P}(X)$ of subsets of X over $\mathbb{F}_2$. The vector addition is the symmetric difference, $A\Delta B = (A \setminus B) \cup (B \setminus A)$, and the scalar multiplication is defined by $0 \cdot A = \emptyset$ and $1 \cdot A = A$.

(5) The space E/F of elements of a field E over a subfield F of E, in particular, $\mathbb{R}/\mathbb{Q}$.

Exercise 1.5. Deduce from the axioms of a vector space that for any scalar α and vector a, the following statements hold:

(i) $0a = \theta$,
(ii) $-a = (-1)a$,
(iii) $\alpha\theta = \theta$, and
(iv) if $\alpha a = \theta$, then either $\alpha = 0$ or $a = \theta$.

2. Linear independence and bases

Let V be a vector space over a field F.

Definition 2.1. A *linear combination* of vectors $a_1, \ldots, a_n \in V$ with coefficients $\alpha_1, \ldots, \alpha_n \in F$ is the expression $\alpha_1 a_1 + \cdots + \alpha_n a_n$. If $\alpha_1 = \ldots = \alpha_n = 0$, then the linear combination is called *trivial*. If $\alpha_i \neq 0$ for at least one $i = 1, \ldots, n$, then the linear combination is called *nontrivial*. Obviously, the trivial linear combination is equal to θ. A system $\{a_1, \ldots, a_n\}$ of vectors of V is *linearly independent* if only trivial linear combination of $a_1, \ldots, a_n$ is equal to θ (that is, the equality $\alpha_1 a_1 + \cdots + \alpha_n a_n = \theta$ implies $\alpha_1 = \ldots = \alpha_n = 0$). A system $\{a_1, \ldots, a_n\}$ is *linearly dependent* if there is a nontrivial linear combination of $a_1, \ldots, a_n$ equal to θ.

Exercise 2.2. Check whether a system $\{a_1, \ldots, a_n\}$ of vectors in F^m is linearly dependent.

Solution. Let A denote the $m \times n$ matrix with columns $a_1, \ldots, a_n$. Then $\{a_1, \ldots, a_n\}$ is linearly dependent if and only if the homogeneous system of linear equations $Ax = 0$ has a nontrivial solution. □

Proposition 2.3. *Let $\{a_1, \ldots, a_n\}$ be a system of vectors of V. Then the following statements are equivalent:*

(i) *$\{a_1, \ldots, a_n\}$ is linearly dependent,*
(ii) *at least one of the vectors $a_1, \ldots, a_n$ is a linear combination of the remaining ones,*
(iii) *at least one of the vectors $a_1, \ldots, a_n$ is a linear combination of the preceding ones.*

Proof. (i)⇒(iii) We have that $\sum_{i=1}^{n} \alpha_i a_i = \theta$ for some $\alpha_i \in F$ and not all of them are zeros. Let $j = \max\{i : \alpha_i \neq 0\}$. Then $\sum_{i=1}^{j} \alpha_i a_i = \theta$, whence $\alpha_i a_j = \sum_{i=1}^{j-1} (-\alpha_i) a_i$, and so $a_j = \sum_{i=1}^{j-1} (-\frac{\alpha_i}{\alpha_j}) a_i$.

(iii)⇒(ii) is obvious.

(ii)⇒(i) There is $j \in \{1, \ldots, n\}$ such that $a_j = \sum_{i \neq j} \alpha_i a_i$ for some $\alpha_i \in F$. Put $\alpha_j = -1$. Then $\sum_{i=1}^{n} \alpha_i a_i = \theta$ and the linear combination $\sum_{i=1}^{n} \alpha_i a_i$ is nontrivial. □

Definition 2.4. For every subset $S \subseteq V$, the *span* of S, denoted $\langle S \rangle$, is the set of all vectors of V which can be expressed as a linear combination of vectors from S, that is, $\langle S \rangle = \{\sum_{i=1}^{n} \alpha_i a_i : a_i \in S, \alpha_i \in F, n \in \mathbb{N}\}$.

Exercise 2.5. Prove that

(i) $\theta \in \langle S \rangle$,
(ii) $S \subseteq \langle S \rangle$,
(iii) if $S \subseteq T$, then $\langle S \rangle \subseteq \langle T \rangle$,
(iv) $\langle \langle S \rangle \rangle = \langle S \rangle$,

(v) the span of S remains the same if we add to S a vector which is a linear combination of vectors from S or remove from S a vector which is a linear combination of the remaining vectors from S, and

(vi) a system $\{a_1, \ldots, a_n, a_{n+1}\}$ of vectors in V is linearly independent iff the system $\{a_1, \ldots, a_n\}$ is linearly independent and $a_{n+1} \notin \langle a_1, \ldots, a_n \rangle$.

Definition 2.6. A system $\{a_1, \ldots, a_n\}$ of vectors in V is a *basis* if it is linearly independent and $\langle a_1, \ldots, a_n \rangle = V$.

Equivalently, a basis is a maximal linearly independent system of vectors in V.

Proposition 2.7. *Let S be a finite subset of V such that $\langle S \rangle = V$. Then V has a basis $\{a_1, \ldots, a_n\} \subseteq S$.*

Proof. It suffices to choose a linearly independent system $\{a_1, \ldots, a_n\}$ of vectors from S such that $S \subseteq \langle a_1, \ldots, a_n \rangle$. Then $\{a_1, \ldots, a_n\}$ would be a basis.

Pick a nonzero vector $a_1 \in S$. The system $\{a_1\}$ is linearly independent. If $S \subseteq \langle a_1 \rangle$, we are done. Otherwise pick $a_2 \in S \setminus \langle a_1 \rangle$. The system $\{a_1, a_2\}$ is linearly independent. If $S \subseteq \langle a_1, a_2 \rangle$, we are done. Otherwise pick $a_3 \in S \setminus \langle a_1, a_2 \rangle$ and so on. In $\leq |S|$ steps we obtain a required system $\{a_1, \ldots, a_n\}$. □

Definition 2.8. A vector space V is *finite-dimensional* if there is a finite subset $S \subseteq V$ such that $\langle S \rangle = V$.

It is immediate from Proposition 2.11 that

Corollary 2.9. *Every finite-dimensional vector space has a basis.*

Now we are going to show that all bases of a finite-dimensional vector space have the same number of vectors.

Lemma 2.10 (Two Systems Lemma)**.** *Let $A = \{a_1, \ldots, a_n\}$ and $B = \{b_1, \ldots, b_m\}$ be two systems of vectors in a vector space. Suppose that B is linearly independent and $B \subseteq \langle A \rangle$ (= every vector b_j is a linear combination of vectors $a_1, \ldots, a_n$). Then $n \geq m$.*

Proof. Assume on the contrary that $n < m$. Let A_1 denote the system obtained from A by adding the vector b_1 from the left, that is, $A_1 = \{b_1, a_1, \ldots, a_n\}$. Since $b_1 \in \langle A \rangle$, A_1 is linearly dependent. Remove from A_1 any vector which is a linear combination of its predecessors. This vector cannot be b_1. So we obtain a system $C_1 = \{b_1, \ldots\}$ consisting of n vectors and such that $\langle C_1 \rangle = \langle A \rangle$. Next, let A_2 denote the system obtained from C_1 by adding the vector b_2 from the left. Since $b_2 \in \langle A \rangle$, A_2 is linearly dependent. Remove from A_2 any vector which is a linear combination of its predecessors. This vector cannot be neither b_2 nor b_1, since B is linearly independent. So we obtain a system $C_2 = \{b_2, b_1, \ldots\}$ consisting of n vectors and such that $\langle C_2 \rangle = \langle A \rangle$. Continuing this process, after n steps we obtain the system $C_n = \{b_n, b_{n-1}, \ldots, b_1\}$ with $\langle C_n \rangle = \langle A \rangle$. But then $b_{n+1} \in \langle b_n, b_{n-1}, \ldots, b_1 \rangle$, which is a contradiction. □

Theorem 2.11. *In a finite-dimensional vector space all bases have the same number of vectors.*

Proof. Let A and B be two bases in V. Since B is linearly independent and $B \subseteq V = \langle A \rangle$, we obtain by Lemma 2.10 that $|A| \geq |B|$. Similarly $|B| \geq |A|$. Hence $|A| = |B|$. □

Theorem 2.11 justifies the following definition.

Definition 2.12. Let V be a finite-dimensional vector space over a field F. Then the *dimension* of V, denoted $\dim V$, is the number of vectors in a basis of V.

Examples 2.13. (1) Consider the space F^n. For each $i = 1, \dots, n$, let e_i denote the vector whose i-th coordinate is 1 and others are 0-s. Then $\{e_1, \dots, e_n\}$ is a basis of F^n, and so $\dim F^n = n$.

To see that $\{e_1, \dots, e_n\}$ is linearly independent, let $\alpha_1 e_1 + \cdots + \alpha_n e_n = \theta$. Since $\alpha_1 e_1 + \cdots + \alpha_n e_n = (\alpha_1, \dots, \alpha_n)$, it follows that $\alpha_1 = \ldots = \alpha_n = 0$.

To see that $\langle e_1, \dots, e_n \rangle = F^n$, let $x \in F^n$, say $x = (\alpha_1, \dots, \alpha_n)$. Then $x = \alpha_1 e_1 + \cdots + \alpha_n e_n$.

The basis $\{e_1, \dots, e_n\}$ is called the *standard basis* of F^n.

(2) Consider the space $F_n[x]$ of polynomials over a filed F of degree $\leq n$. It is easy to see that $\{1, x, \dots, x^n\}$ is a basis of $F_n[x]$, and so $\dim F_n[x] = n + 1$.

3. Coordinates

Let V be a finite-dimensional vector space over a field F.

Lemma 3.1. *Let $\{a_1, \dots, a_n\}$ be a basis of V. Then every vector $x \in V$ can be uniquely written in the form $x = \alpha_1 a_1 + \cdots + \alpha_n a_n$ where all $\alpha_i \in F$.*

Proof. That x can be written in this form follows from the fact that $\langle a_1, \dots, a_n \rangle = V$. To see uniqueness, suppose that $x = \alpha_1 a_1 + \cdots \alpha_n a_n$ and $x = \beta_1 a_1 + \cdots \beta_n a_n$ for some $\alpha_i, \beta_i \in F$. Subtracting the second equality from the first gives us that $(\alpha_1 - \beta_1) a_1 + \cdots + (\alpha_n - \beta_n) a_n = 0$. Since $\{a_1, \dots, a_n\}$ is linearly independent, it follows that $\alpha_i - \beta_i = o$ and so $\alpha_i = \beta_i$. $\square$

Definition 3.2. The column vector $[x]_S = (\alpha_1, \dots, \alpha_n)^T \in F^n$ is called the *coordinate vector* of x relative to S.

Exercise 3.3. Given vectors $a_1, \dots, a_n, b \in F^n$, determine whether $S = \{a_1, \dots, a_n\}$ is a basis, and if yes, find $[b]_S$.

Solution. Let A denote the $n \times n$ matrix with columns $a_1, \dots, a_n$. Applying elementary row operations to the augmented matrix $(A|b)$, find its reduced row echelon form. Then S is a basis iff we obtain a matrix $(E|c)$, where E is the identity matrix, and in this case, $[b]_S = c$. $\square$

Exercise 3.4. Let $S = \{a_1, \dots, a_n\}$ be a basis of V. Check that

(i) $[x + y]_S = [x]_S + [y]_S$ for all $x, y \in V$, and
(ii) $[\alpha x]_S = \alpha [x]_S$ for all $\alpha \in F$ and $x \in V$.

Definition 3.5. Let V and W be vector spaces over a field F. An *isomorphism* between V and W is any bijection $\mathcal{A} : V \to W$ such that

(1) $\mathcal{A}(x + y) = \mathcal{A}(x) + \mathcal{A}(y)$ for all $x, y \in V$, and
(2) $\mathcal{A}(\alpha x) = \alpha \mathcal{A}(x)$ for all $\alpha \in F$ and $x \in V$.

The next theorem tells us that if V is a finite-dimensional vector space over a field F and $n = \dim V$, then V is isomorphic to the coordinate space F^n.

Theorem 3.6. *Let V be a finite-dimensional vector space over a field F and let $n = \dim V$. Then V is isomorphic to F^n.*

Proof. Let S be a basis of V. Define $\mathcal{A} : V \to F^n$ by $\mathcal{A}(x) = [x]_S$. Clearly, $\mathcal{A}$ is a bijection. Then by Exercise 3.4 and Definition 3.5, $\mathcal{A}$ is an isomorphism. □

Now let $S = \{a_1, \dots, a_n\}$ and $S' = \{b_1, \dots, b_n\}$ be two bases of V and let $[x]_S = (\alpha_1, \dots, \alpha_n)^T$ and $[x]_{S'} = (\beta_1, \dots, \beta_n)^T$. How is the vector $[x]_S$ related to the vector $[x]_{S'}$?

To answer this question, express vectors of S' via vectors of S:

$$b_j = \sum_{i=1}^{n} \gamma_{ij} a_i.$$

Then on the one hand,

$$x = \sum_{j=1}^{n} \beta_j b_j = \sum_{j=1}^{n} \beta_j \sum_{i=1}^{n} \gamma_{ij} a_i = \sum_{i=1}^{n} (\sum_{j=1}^{n} \gamma_{ij} \beta_j) a_i,$$

and on the other hand,

$$x = \sum_{i=1}^{n} \alpha_i a_i.$$

Consequently,

$$\alpha_i = \sum_{j=1}^{n} \gamma_{ij} \beta_j.$$

Definition 3.7. The matrix $T = (\gamma_{ij})_{i,j=1}^{n}$ is called the *transition matrix* from S to S'.

Thus, we have showed that

$$[x]_S = T[x]_{S'}.$$

Exercises 3.8. (i) Let $S = \{e_1, e_2\}$ be the standard basis of the plane $\mathbb{R}^2$ and let $S' = \{e'_1, e'_2\}$ be the basis obtained from S by rotating the plane counterclockwise about the origin through an angle φ. Find the transition matrix from S to S'.

(ii) Find the transition matrix from the basis $S = \{1, x, \dots, x^n\}$ of $F_n[x]$ to the basis $S' = \{1, x + \alpha, (x + \alpha)^2, \dots, (x + \alpha)^n\}$.

(iii) Given bases $S = \{a_1, \dots, a_n\}$ and $S' = \{b_1, \dots, b_n\}$ of F^n, find the transition matrix from S to S'.

Solution. Let A and B denote the $n \times n$ matrices with columns $a_1, \dots, a_n$ and $b_1, \dots, b_n$, respectively. Then the transition matrix T from S to S' is the solution of the matrix equation $AX = B$, that is $T = A^{-1}B$. To find $A^{-1}B$, apply elementary row operations to the matrix $(A|B)$ so that to obtain its reduced row echelon form $(E|C)$. Then $A^{-1}B = C$. □

Remark 3.9. The elementary row operations method is based on the following simple fact:

Let A be an $m \times n$ matrix and let B be the matrix obtained from A by an elementary row operation. Then $B = FA$ where F is the matrix obtained from the identity matrix by the same elementary row operation.

Now to justify the method, let $\tau_1, \dots, \tau_k$ be the sequence of elementary row operations reducing $(A|B)$ to $(E|C)$ and let $F_1, \dots, F_k$ be the sequence of matrices obtained from E by $\tau_1, \dots, \tau_k$. We then have that $F_k \cdots F_1 A = E$ and $F_k \cdots F_1 B =$

C. It follows from the first equality that $F_k \cdots F_1 = A^{-1}$. Then the second equality gives us that $C = A^{-1}B$.

4. Subspaces and direct sums

Let V be a vector space over a field F.

Definition 4.1. A nonempty subset $W \subseteq V$ is a *subspace* if

(1) for all $x, y \in W$, one has $x + y \in W$, and
(2) for all $x \in W$ and $\alpha \in F$, one has $\alpha x \in W$.

Note that every subspace of V is itself a vector space over F.

Examples 4.2. (1) Every space V has *trivial* subspaces: $\{0\}$ and V. All others are called *nontrivial* or *proper*.

(2) The proper subspaces of $\mathbb{R}^3$ are precisely the lines and the planes through the origin.

(3) $F_n[x]$ is a subspace of $F[x]$.

(4) The space $C[0,1]$ of continuous real-valued functions defined on the interval $[0,1]$ is a subspace of the space $\mathbb{R}^{[0,1]}$ of all real-valued function on $[0,1]$.

(5) For every subset $S \subseteq V$, the span of S is the smallest subspace of V containing S.

(6) Let A be an $m \times n$ matrix over F and consider the homogeneous system of linear equations $Ax = 0$. Its solutions form a subspace of F^n, called the *solution space*.

Exercises 4.3. (i) Prove that for any family $\{W_i : i \in I\}$ of subspaces of V, $\bigcap_{i \in I} W_i$ is a subspace of V.

(ii) Assuming that V is finite-dimensional, prove that for every subspace $W \subseteq V$, one has

$$0 \leq \dim W \leq \dim V.$$

Furthermore, $\dim W = 0$ iff $W = \{0\}$, and $\dim W = \dim V$ iff $W = V$.

(iii) Check that for any subspaces $W_1, \ldots, W_n \subseteq V$, their sum

$$W_1 + \cdots + W_n = \{x_1 + \cdots + x_n : x_i \in W_i \text{ for each } i\}$$

is a subspace of V.

(iv) Assuming that V is finite-dimensional, prove that for any subspaces $W_1, W_2 \subseteq V$, one has

$$\dim(W_1 + W_2) = \dim W_1 + \dim W_2 - \dim(W_1 \cap W_2).$$

Proof. Pick a basis $a_1, \ldots, a_m$ of $W_1 \cap W_2$ and extend it to bases $a_1, \ldots, a_m, b_1, \ldots, b_k$ of W_1 and $a_1, \ldots, a_m, c_1, \ldots, c_l$ of W_2. We claim that $a_1, \ldots, a_m, b_1, \ldots, b_k, c_1, \ldots, c_l$ is a basis of $W_1 + W_2$. Clearly these vectors span $W_1 + W_2$. To see that they are linearly independent, let

$$\alpha_1 a_1 + \ldots + \alpha_m a_m + \beta_1 b_1 + \ldots + \beta_k b_k + \gamma_1 c_1 + \ldots + \gamma_l c_l = 0.$$

Write this equality as

$$\alpha_1 a_1 + \ldots + \alpha_m a_m + \beta_1 b_1 + \ldots + \beta_k b_k = -\gamma_1 c_1 - \ldots - \gamma_l c_l.$$

Since the left-hand side is a vector from W_1 and the right-hand side is a vector from W_2, we obtain that

$$\alpha_1 a_1 + \ldots + \alpha_m a_m + \beta_1 b_1 + \ldots + \beta_k b_k = \delta_1 a_1 + \ldots + \delta_m a_m$$

for some $\delta_1, \dots, \delta_m \in F$, whence

$$(\alpha_1 - \delta_1)a_1 + \ldots + (\alpha_m - \delta_m)a_m + \beta_1 b_1 + \ldots + \beta_k b_k = 0.$$

It follows that $\beta_1 = \ldots = \beta_k = o$. Similarly $\gamma_1 = \ldots = \gamma_l = o$. Then also $\alpha_1 = \ldots = \alpha_m = o$. $\square$

Definition 4.4. Let $V_1, \dots, V_n$ be vector spaces over a field F. The *direct sum* of $V_1, \dots, V_n$, denoted $V_1 \oplus \cdots \oplus V_n$, is the vector space over F whose vectors are n-tuples $(x_1, \dots, x_n)$, where each $x_i \in V_i$, and the operations are coordinatewise, that is, $(x_1, \dots, x_n) + (y_1, \dots, y_n) = (x_1 + y_1, \dots, x_n + y_n)$ and $\alpha(x_1, \dots, x_n) = (\alpha x_1, \dots, \alpha x_n)$.

Example 4.5. F^n is the direct sum of n copies of F.

Theorem 4.6. *Let V be a vector space over a field F and let $W_1, \dots, W_n$ be subspaces of V. Suppose that every $x \in V$ can be uniquely written in the form $x = x_1 + \cdots + x_n$ where each $x_i \in W_i$. Then V is isomorphic to $W_1 \oplus \cdots \oplus W_n$.*

Proof. Define the mapping $\mathcal{A} : V \to W_1 \oplus \cdots \oplus W_n$ by $\mathcal{A}(x) = (x_1, \dots, x_n)$ where $x = x_1 + \cdots + x_n$ and each $x_i \in W_i$. It is easy to see that $\mathcal{A}$ is bijective, $\mathcal{A}(x+y) = \mathcal{A}(x) + \mathcal{A}(y)$ and $\mathcal{A}(\alpha x) = \alpha\mathcal{A}(x)$. Hence, $\mathcal{A}$ is an isomorphism. $\square$

Exercise 4.7. Let V be a vector space over a field F and let $W_1, \dots, W_n$ be subspaces of V. Prove that the following statements are equivalent:

(i) every $x \in V$ can be uniquely written in the form $x = x_1 + \cdots + x_n$ where each $x_i \in W_i$, and
(ii) $V = \sum_{i=1}^n W_i$ and $W_j \cap \sum_{i \neq j} W_i = 0$ for each $j = 1, \dots, n$.

5. Rank of a matrix

Definition 5.1. Let A be an $m \times n$ matrix over F, let $a_1, \dots, a_m$ denote rows of A, and let $b_1, \dots, b_n$ denote columns of A. The *row (column) space* of A is the subspace $\langle a_1, \dots, a_m \rangle \subseteq F^n$ ($\langle b_1, \dots, b_n \rangle \subseteq F^m$). The *row (column) rank* of A is the dimension of row (column) space of A.

Equivalently, the row (column) rank of A is the number of vectors in a maximal linearly independent subsystem of $\{a_1, \dots, a_m\}$ ($\{b_1, \dots, b_n\}$).

It can be easily checked that

(i) elementary row operations do not change the row rank of A,
(ii) the row-echelon form of A produced by the Gauss elimination method has the same row rank that A, and
(iii) the row rank of A is the number of nonzero rows in the row-echelon form of A.

Theorem 5.2. *The row rank and the column rank are equal.*

Proof. It suffices to show that the row rank of A does not exceed the column rank of A. Let k be the row rank of A. Pick a linearly independent system of k rows of A, say $\{a_{i_1}, \dots, a_{i_k}\}$, and let C denote the $n \times k$ matrix with columns $a_{i_1}, \dots, a_{i_k}$. Since the system $\{a_{i_1}, \dots, a_{i_k}\}$ is linearly independent, the homogeneous system of linear equations $Cx = 0$ has only trivial solution. It follows that the row echelon form of C has k nonzero rows, so the row rank of C is k. Pick a linearly independent system of k rows of C, say $\{c_{j_1}, \dots, c_{j_k}\}$. Then the corresponding system $\{b_{j_1}, \dots, b_{j_k}\}$ of columns of A is also linearly independent. Hence the column rank of A is $\geq k$. $\square$

Since the row rank and the column rank of A are equal, they are simply called the *rank* of A.

Exercise 5.3. Prove that a system of linear equations $Ax = b$ has a solution iff $\operatorname{rank}(A|b) = \operatorname{rank}(A)$.

Proof. The system $Ax = b$ has a solution iff b is a linear combination of the columns of A iff $\operatorname{rank}(A|b) = \operatorname{rank}(A)$. □

Exercise 5.4. Given a homogeneous system of linear equations $Ax = 0$, find a basis of the solution space.

Exercise 5.5. Given a system of linear equations $Ax = b$, find its general solution in the form $x = c_0 + \alpha_1 c_1 + \cdots + \alpha_k c_k$, where c_0 is its particular solution and $c_1, \ldots, c_k$ is a basis of the solution space of the corresponding homogeneous system $Ax = 0$.