MAD 4471: Introduction to Cryptography and Coding Theory **Fall 2022**

Lecture 13: Linear Algebra Review

Lecturer: Jean-François Biasse *TA: William Youmans*

In this lecture, we recall some elementary facts on linear algebra that are necessary to understand lattices. For a thorough introduction, we refer to the linear algebra pre-requisites of this class.

13.1 Fields

Fields are elements of abstract algebra that play a role in linear algebra. However, vector spaces are sometimes introduced without refering to fields by restricting the scope to "real vector spaces", i.e. vector spaces that just involve real numbers. While this is convenient to get most of the points across, we need other kinds of fields to introduce lattice based cryptosystems. In particular, we need "finite fields". Therefore, we must go through the formal definition of a field.

Definition 13.1 (Field) *We say that a set K is a field if there are two binary operations $+, \cdot : K \times K \to K$ such that*

1. $\forall a, b, c$, $a + (b + c) = (a + b) + c$ *and* $a \cdot (b \cdot c) = (a \cdot b) \cdot c$.
2. $\forall a, b$, $a + b = b + a$ *and* $a \cdot b = b \cdot a$.
3. *There are* $0 \neq 1 \in F$ *such that* $\forall a \in F$, $a + 0 = a$ *and* $a \cdot 1 = a$.
4. $\forall a \in F$, $\exists -a \in F$ *such that* $a + (-a) = 0$.
5. $\forall a \neq 0$, $\exists a^{-1}$ *such that* $a \cdot a^{-1} = 1$.
6. $a \cdot (b + c) = (a \cdot b) + (a \cdot c)$.

Example 1 (Real numbers) *The most obvious example of a field is $(\mathbb{R}, +, \cdot)$ where $+$ and $\cdot$ are the usual addition and multiplication between real numbers. All the properties of a field can be immediately checked.*

As mentioned above, real numbers are not enough. In particular, we need the so-called "finite fields" to define the lattice-based cryptosystems that are instroduced in this class.

Proposition 13.2 (Prime fields) *For all prime p, the field $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$ is a field of size p.*

Proof: We know that the cardinality of $\mathbb{Z}/p\mathbb{Z}$ is p. Moreover, the addition and multiplication of residue classes satisfy properties 1, 2, 3, and 6 for neutral elements [0] and [1]. Additionally, all residue classes have an inverse for the addition: $[a] + [-a] = [0]$. Finally, we know that all a coprime to p is invertible modulo p. This means that all classes $[a] \neq [0]$ are invertible, thus satisfying Property 5. ■

Proposition 13.3 (Finite fields) *For all prime powers $q = p^n$, there exists a field $\mathbb{F}_q$ of cardinality q.*

Proof: We will prove this statement in the lecture on modular polynomials. In a nutshell, the field $\mathbb{F}_q$ is made of polynomials over $\mathbb{F}_p$ that are reduced modulo an irreducible polynomial of degree n. ■

13.2 Vector spaces

Vector spaces can have many shapes, but their behavior is modeled after that of vectors with entries in a field K. In this type of space, the natural way to define the addition of two vectors (of the same length) $\mathbf{u}$ and $\mathbf{v}$ is coordinate-wise. Similarly, the multiplication of a vector $\mathbf{u}$ by $\lambda \in K$ is most naturally defined as the multiplication of every entry of $\mathbf{u}$ by λ. Such multiplication has distributivity over the addition of vectors.

Definition 13.4 (Vector space) *A vector space V over a field K is a set equipped with binary operations $+ : V \times V \to V$ and $\cdot : K \times V \to V$ such that*

1. $\forall \mathbf{u}, \mathbf{v}, \mathbf{w} \in V$: $\mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w}$.
2. $\forall \mathbf{u}, \mathbf{v} \in V$, $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$.
3. $\exists \mathbf{0} \in V$ *such that* $\forall \mathbf{v}$, $\mathbf{v} + \mathbf{0} = \mathbf{v}$.
4. $\forall \lambda, \mu \in K$, *and* $\mathbf{v} \in V$, $\lambda \cdot (\mu \cdot \mathbf{v}) = (\lambda\mu) \cdot \mathbf{v}$.
5. $\forall \mathbf{v} \in V$, $1 \cdot \mathbf{v} = \mathbf{v}$.
6. $\forall \lambda \in K$, *and* $\mathbf{u}, \mathbf{v} \in V$, $\lambda \cdot (\mathbf{u} + \mathbf{v}) = \lambda \cdot \mathbf{u} + \lambda \cdot \mathbf{v}$.
7. $\forall \lambda, \mu \in K$, *and* $\mathbf{v} \in V$, $(\lambda + \mu) \cdot \mathbf{v} = \lambda \cdot \mathbf{v} + \mu \cdot \mathbf{v}$.

Example 2 (Vectors in $\mathbb{R}^n$) *As mentioned before, the prototypical example of a vector space is the vectors of a given length n with entries in $\mathbb{R}$ (which we denote $\mathbb{R}^n$). With this space, operations are defined as*

$$(u_1, \ldots, u_n) + (v_1, \ldots, v_n) = (u_1 + v_1, \ldots, u_n + v_n)$$
$$\lambda \cdot (v_1, \ldots, v_n) = (\lambda v_1, \ldots, \lambda v_n).$$

The verification of the properties making $\mathbb{R}^n$ a vector space is immediate.

There can be much more "exotic" vector spaces. For example the set of continuous functions $\mathbb{R} \to \mathbb{R}$ is a vector space over $\mathbb{R}$ for the usual addition of functions $(f + g)(x) = f(x) + g(x)$ and the intuitive scalar multiplication $(\lambda \cdot f)(x) = \lambda f(x)$. The most important example for us is the vector spaces made of vectors with entries in a finite field $\mathbb{F}_q$.

Example 3 (Vectors of $\mathbb{F}_q^n$) *Vectors of $\mathbb{F}_q^n$ have the structure of a vector space similar to that of $\mathbb{R}^n$, with addition and salar multiplication defined as*

$$(u_1, \ldots, u_n) + (v_1, \ldots, v_n) = (u_1 + v_1, \ldots, u_n + v_n)$$
$$\lambda \cdot (v_1, \ldots, v_n) = (\lambda v_1, \ldots, \lambda v_n).$$

We are interested in sets of elements that generate a vector space. In general, any set of elements in V generates its own vector space $V' \subseteq V$.

Definition 13.5 (Span of a set of elements) *Let V be a K-vector space, and $S = \{\mathbf{v}_1, \mathbf{v}_2, \dots$ a set (finite or infinite) of elements of V. We denote by $Span(S)$ (or sometimes $Span_K(S)$ to emphasize the field of definition) the vector space generated by all linear combinations of elements of S:*

$$Span(S) = \{\mathbf{v} = \lambda_1\mathbf{v}_1 + \lambda_2\mathbf{v}_2 + \dots \mid \text{for } \lambda_i \in K,\ \mathbf{v}_i \in V\}$$

Example 4 *Let $K = \mathbb{R}$, $V = \mathbb{R}^3$, and $S = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ for $\mathbf{v}_1 = (1, 0, 0)$, $\mathbf{v}_2 = (0, 1, 0)$ and $\mathbf{v}_3 = (0, 0, 1)$. We have that $Span(S) = \mathbb{R}^3$ because*

$$\forall \mathbf{x} = (x_1, x_2, x_3) \in \mathbb{R}^3,\ \mathbf{x} = x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + x_3\mathbf{v}_3 \in Span(S).$$

If a K-vector space V is of the form $\text{Span}(S)$ for a finite set S, then we say that V is *finitely generated*.

Definition 13.6 (Basis of a vector space) *A basis of a K-vector space V is a minimal generating set S such that $V = Span(S)$.*

Example 5 *For our previous example $V = \mathbb{R}^3$, the generating set $S = \mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$, is also a basis. Indeed, S generates V, but no subset of S generates V. On the other hand, we can see that the set*

$$S' = \{(1, 0, 0), (0, 1, 0), (0, 0, 1), (1, 1, 0)\}$$

is not a basis of V. Indeed, the element $(1, 1, 0) = (1, 0, 0) + (0, 1, 0)$ is redundant, and can be removed.

Note that a basis can be infinite. For example, the space of polynomial functions $\mathbb{R} \to \mathbb{R}$ is an $\mathbb{R}$-vector space, and has a basis $(f_i)_{i\geq 0}$ where $f_i : x \mapsto x^i$.

Definition 13.7 (Dimension of a vector space) *The cardinality of a basis of a vector space V is called the dimension of V. If the basis is inifinite, then V is said to have infinite dimension.*

Example 6 *The dimension of $\mathbb{R}^3$ is 3, and more generally, the dimension of $\mathbb{R}^n$ is n.*

13.3 Matrices

Matrices over a field K are another example of a K-vector space. Matrices are objects that are ubiquitous in mathematics. Their "bare-bone" definition is very simple though: they are two-dimensional arrays of numbers in K.

Definition 13.8 (Matrices over K) *Let K be a field, and $m, n > 0$ be integers. The space of n by m matrices over K is the set of two-dimensional arrays of the form $(a_{i,j})_{i\leq n, j\leq m}$ where $a_{i,j} \in K$. This space is denoted by $K^{n\times m}$.*

Matrices have a typical representation of the following form:

$$(a_{i,j})_{i\leq n, j\leq m} = \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,m} \\ a_{2,1} & a_{2,2} & \dots & a_{2,m} \\ \vdots & \vdots & & \vdots \\ a_{n,1} & a_{n,2} & \dots & a_{n,m} \end{pmatrix}$$

Definition 13.9 (Addition of matrices) *There is a natural addition* $+ : K^{n\times m} \times K^{n\times m} \to K^{n\times m}$ *defined as*

$$\begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,m} \\ a_{2,1} & a_{2,2} & \dots & a_{2,m} \\ \vdots & \vdots & & \vdots \\ a_{n,1} & a_{n,2} & \dots & a_{n,m} \end{pmatrix} + \begin{pmatrix} b_{1,1} & b_{1,2} & \dots & b_{1,m} \\ b_{2,1} & b_{2,2} & \dots & b_{2,m} \\ \vdots & \vdots & & \vdots \\ b_{n,1} & b_{n,2} & \dots & b_{n,m} \end{pmatrix} = \begin{pmatrix} a_{1,1}+b_{1,1} & a_{1,2}+b_{1,2} & \dots & a_{1,m}+b_{1,m} \\ a_{2,1}+b_{2,1} & a_{2,2}+b_{2,2} & \dots & a_{2,m}+b_{2,m} \\ \vdots & \vdots & & \vdots \\ a_{n,1}+b_{n,1} & a_{n,2}+b_{n,2} & \dots & a_{n,m}+b_{n,m} \end{pmatrix}$$

Proposition 13.10 *The space of* n *by* m *matrices over* K *is a vector space over* K *of dimension* nm*.*

Proof: The definition of matrix addition, together with the scalar multiplication defined by

$$\lambda \cdot \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,m} \\ a_{2,1} & a_{2,2} & \dots & a_{2,m} \\ \vdots & \vdots & & \vdots \\ a_{n,1} & a_{n,2} & \dots & a_{n,m} \end{pmatrix} = \begin{pmatrix} \lambda a_{1,1} & \lambda a_{1,2} & \dots & \lambda a_{1,m} \\ \lambda a_{2,1} & \lambda a_{2,2} & \dots & \lambda a_{2,m} \\ \vdots & \vdots & & \vdots \\ \lambda a_{n,1} & \lambda a_{n,2} & \dots & \lambda a_{n,m} \end{pmatrix}$$

satisfy all the properties of the vector space definition. Here, we are essentially treating matrices as vectors of K^{nm}. Likewise, matrices $M_{i,j}$ defined by $a_{i,j} = 1$ and $a_{i',j'} = 0$ for $i' \neq i$ or $j' \neq j$ form a basis of $K^{n\times m}$, thus showing that the dimension of $K^{n\times m}$ is nm. ■

The vector space structure of $K^{n\times m}$ does not really distinguish it from K^{nm}. Indeed, as far as addition and scalar multiplication go, we treat matrices as vectors of length nm. However, $K^{n\times m}$ can be endowed with a multiplication operation that sets them appart from vectors of K^{nm}. This multiplication is not coordinate-wise. Instead, it has a very specific definition as follows:

Definition 13.11 (Matrix multiplication) *Let* $A = (a_{i,j})_{i\leq n, j\leq l} \in K^{n\times l}$ *and* $B = (b_{j,k})_{j\leq l, k\leq m} \in K^{l\times m}$*. We define* $C = A \times B \in K^{n\times m}$ *by*

$$C = (c_{i,j})_{i\leq n, j\leq m} \quad \text{for } c_{i,j} = \sum_{k=1}^{l} a_{i,k} b_{k,j}.$$

The coefficient $c_{i,j}$ is the dot-product of the i-th row of A and the j-th column of B. As a reminder, a dot product between two vectors $\mathbf{x} = (x_1, \dots, x_n)$ and $\mathbf{y} = (y_1, \dots, y_n)$ is $\langle \mathbf{x}, \mathbf{y} \rangle = \sum_i x_i y_i$. Note that for this operation to be possible, the number of columns of A needs to match the number of rows of B (i.e. l). We illustrate the computation of a coefficient of the matrix product in Figure 13.3.

Example 7

$$\begin{pmatrix} 3 & 2 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 2 & 2 & 2 \\ 3 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 10 & 6 & 8 \\ 4 & 2 & 2 \\ 2 & 2 & 2 \end{pmatrix}$$

There is a matrix which plays an important role regarding the matrix product, namely the *identity matrix*.

Definition 13.12 (Identity matrix) *Let* K *be a field. The identity matrix* $I_n \in K^{n\times n}$ *is defined as*

$$I_n = \begin{pmatrix} 1 & (0) & \\ & \ddots & \\ (0) & & 1 \end{pmatrix}.$$

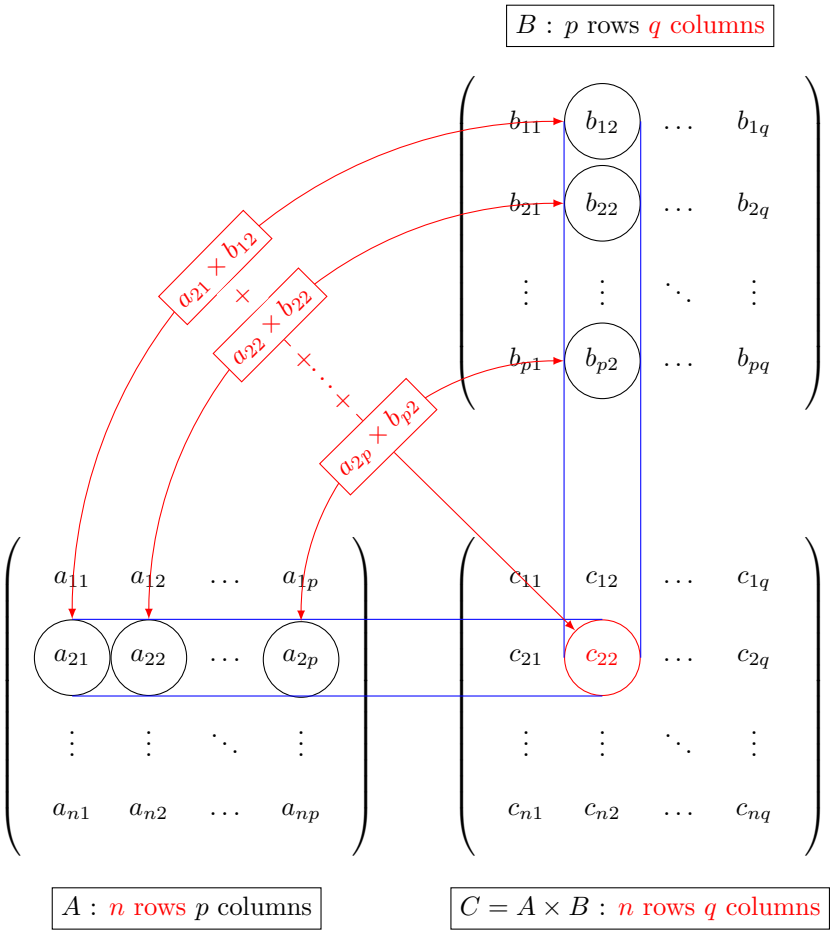


Figure 13.1: Computation of a coefficient of the matrix product

Proposition 13.13 *Let K be a field, $\forall A \in K^{n\times n}$, $A \times I_n = I_n \times A = A$.*

The identity matrix is a neutral element for the matrix multiplication operation. However, $n \times n$ matrices are not a group for the multiplication. Indeed, not every matrix has an inverse. However, some of them are, which is a special case of high importance in many linear algebra problems.

Definition 13.14 (Invertible matrix) *Let K be a field and $A \in K^{n\times n}$. We say that A is invertible if there is $B \in K^{n\times n}$ such that*

$$AB = BA = I_n.$$

In this case, we call B the inverse of A and we denote it $B = A^{-1}$.

13.4 Determinants

A determinant of a matrix in $K^{n\times n}$ is a scalar (i.e. an element of K). The generic formula for computing the determinant of a matrix is quite involved, and determinants of large matrices can be difficult to compute

by hand (unless some "tricks" can be used to simplify the calculation). The case of a 2×2 matrix is well known:

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - cb.$$

To give a general definition of a determinant, we need to discuss permutations of $\{1, \dots, n\}$. A permutation σ of $\{1, \dots, n\}$ is a bijection map $\{1, \dots, n\} \to \{1, \dots, n\}$.

Definition 13.15 (Signature of a permutation) *Let σ be a permutation of $\{1, \dots, n\}$. We define the signature $\operatorname{sgn}(\sigma)$ as*

- $\operatorname{sgn}(\sigma) = 1$ *if the reordering given by σ can be achieved by successively interchanging two entries in an even number of times.*
- $\operatorname{sgn}(\sigma) = -1$ *otherwise.*

Example 8 (Permutations of $\{1, 2, 3\}$) *Let us consider elements of S_3 (i.e. permutations of $\{1, 2, 3\}$). We start with:*

$$\sigma(1) = 1,\ \sigma(2) = 3,\ \sigma(3) = 2.$$

The reordering of $\{1, 2, 3\}$ given by sigma is $\{1, 3, 2\}$. It is achieved by exchanging 2 and 3. So $\operatorname{sgn}(\sigma) = -1$. Now let us consider

$$\sigma(1) = 2,\ \sigma(2) = 3,\ \sigma(3) = 1.$$

The reordering of $\{1, 2, 3\}$ given by sigma is $\{2, 3, 1\}$. It is achieved by first exchanging 1 and 2 (thus giving $\{2, 1, 3\}$, and then exhanging 2 and 3. Therefore $\operatorname{sgn}(\sigma) = 1$.

Definition 13.16 (Determinant of a matrix) *Let $A \in K^{n \times n}$ be a square matrix over a field K. The determinant of A is an element of K defined as*

$$\det(A) = \sum_{\sigma \in S_n} \left(\operatorname{sgn}(\sigma) \prod_{i=1}^{n} a_{i,\sigma(i)} \right).$$

This complicated formula does not seem very practical. Indeed, for large n, it is difficult to evaluate it by hand. But first, we can easily check that it is exactly the formula for 2×2 matrices that we already know!

Example 9 (Determinant formula on 2×2 matrices) *When $n = 2$, we have only two permutations: the identity defined by $\sigma(1) = 1$ and $\sigma(2) = 2$, which satisfies $\operatorname{sgn}(\sigma) = 1$, and σ' defined by $\sigma'(1) = 2$ and $\sigma'(2) = 1$. Since the reordering $\{2, 1\}$ is obtained by exchanging 1 and 2, we have $\operatorname{sgn}(\sigma') = -1$. Let*

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix}$$

The determinant of A is given by

$$\begin{aligned} \det(A) &= \operatorname{sgn}(\sigma) a_{1,\sigma(1)} a_{2,\sigma(2)} + \operatorname{sgn}(\sigma') a_{1,\sigma'(1)} a_{2,\sigma'(2)} \\ &= a_{1,1} a_{2,2} - a_{1,2} a_{2,1} \end{aligned}$$

Proposition 13.17 (Determinant of a 3×3 matrix) *Let $A = (a_{i,j})_{i,j \leq 3} \in K^{3 \times 3}$. Then the determinant of A is given by*

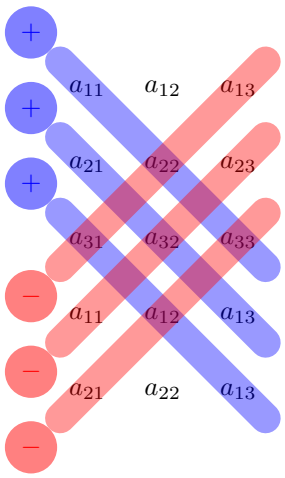
$$\det(A) = a_{1,1}a_{2,2}a_{3,3} - a_{1,1}a_{2,3}a_{3,2} + a_{1,3}a_{2,1}a_{3,2} - a_{1,3}a_{2,2}a_{3,1} + a_{1,2}a_{2,3}a_{3,1} - a_{1,2}a_{2,1}a_{3,2}.$$

Proof: There are $3! = 6$ permutations $\sigma \in S_3$ defined by their reordering $\{\sigma(1), \sigma(2), \sigma(3)\}$ of $\{1, 2, 3\}$:

- $\{1, 2, 3\}$: $\text{sgn}(\sigma) = 1$, $\prod_{i=1}^{3} a_{i,\sigma(i)} = a_{1,1}a_{2,2}a_{3,3}$.
- $\{1, 3, 2\}$: $\text{sgn}(\sigma) = -1$, $\prod_{i=1}^{3} a_{i,\sigma(i)} = a_{1,1}a_{2,3}a_{3,2}$.
- $\{3, 1, 2\}$: $\text{sgn}(\sigma) = 1$, $\prod_{i=1}^{3} a_{i,\sigma(i)} = a_{1,3}a_{2,1}a_{3,2}$.
- $\{3, 2, 1\}$: $\text{sgn}(\sigma) = -1$, $\prod_{i=1}^{3} a_{i,\sigma(i)} = a_{1,3}a_{2,2}a_{3,1}$.
- $\{2, 3, 1\}$: $\text{sgn}(\sigma) = 1$, $\prod_{i=1}^{3} a_{i,\sigma(i)} = a_{1,2}a_{2,3}a_{3,1}$.
- $\{2, 1, 3\}$: $\text{sgn}(\sigma) = -1$, $\prod_{i=1}^{3} a_{i,\sigma(i)} = a_{1,2}a_{2,1}a_{3,2}$.

Putting all the terms together in a sum according to the definition of the determinant gives us the desired formula. ■

Visually, this can be represented as follows (also known as Rule of Sarrus):



13.5 Inverting matrices

Determinants play a crucial role in matrix inversion. In this section, we recall (without proof) a few facts leading to the matrix inversion formula.

Definition 13.18 (Cofactor) *Let $A = (a_{i,j}) \in K^{n \times n}$ be a square matrix over a field K, and $i, j \leq n$. The cofactor of $a_{i,j}$ in A is $(-1)^{i+j} A_{i,j}$ where $A_{i,j}$ is the $(n-1) \times (n-1)$ determinant obtained by removing the i-th row and j-th column of A.*

Example 10 *Cofactor of $a_{1,1} = 0$ in A:*

$$\begin{bmatrix} 0 & 2 & 1 \\ 3 & -1 & 2 \\ 4 & 0 & 1 \end{bmatrix}$$

Proposition 13.19 (Laplace expansion) *For any $i \leq n$, the determinant of $A \in K^{n\times n}$ can be expressed as a sum of cofactors:*

$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{i,j} A_{i,j},$$

where $(-1)^{i+j} A_{i,j}$ is the cofactor of $a_{i,j}$.

Example 11 *Let*

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

Then, by choosing $i = 1$, we have

$$\begin{aligned} \det(A) &= 0 \begin{vmatrix} -1 & 2 \\ 0 & 1 \end{vmatrix} - 2 \begin{vmatrix} 3 & 2 \\ 4 & 1 \end{vmatrix} + 1 \begin{vmatrix} 3 & -1 \\ 4 & 0 \end{vmatrix} \\ &= 0 \times (-1) - 2 \times (-5) + 1 \times 4 = 14 \end{aligned}$$

Proposition 13.20 (Inverse of a matrix) *Let $A \in K^{n\times n}$ such that $\det(A) \neq 0$. Then A is invertible with*

$$A^{-1} = \frac{1}{\det(A)} C^T$$

where $C_{i,j}$ is the cofactor of $a_{i,j}$ and C^T denotes the transpose of C (i.e. the matrix whose entry at i, j is $C_{j,i}$).

Example 12 *Let*

$$A = \begin{pmatrix} 0 & 2 & 1 \\ 3 & -1 & 2 \\ 4 & 0 & 1 \end{pmatrix}.$$

We already established that $\det(A) = 14$. Now the cofactors are

$$C_{1,1} = \begin{vmatrix} -1 & 2 \\ 0 & 1 \end{vmatrix} = -1,\ C_{1,2} = -\begin{vmatrix} 3 & 2 \\ 4 & 1 \end{vmatrix} = 5,\ C_{1,3} = \begin{vmatrix} 3 & -1 \\ 4 & 0 \end{vmatrix} = 4$$

$$C_{2,1} = -\begin{vmatrix} 2 & 1 \\ 0 & 1 \end{vmatrix} = -2,\ C_{2,2} = \begin{vmatrix} 0 & 1 \\ 4 & 1 \end{vmatrix} = -4,\ C_{2,3} = -\begin{vmatrix} 0 & 2 \\ 4 & 0 \end{vmatrix} = 8$$

$$C_{3,1} = \begin{vmatrix} 2 & 1 \\ -1 & 2 \end{vmatrix} = 5,\ C_{3,2} = -\begin{vmatrix} 0 & 1 \\ 3 & 2 \end{vmatrix} = 3,\ C_{3,3} = \begin{vmatrix} 0 & 2 \\ 3 & -1 \end{vmatrix} = -6$$

Hence

$$C = \begin{pmatrix} -1 & 5 & 4 \\ -2 & -4 & 8 \\ 5 & 3 & -6 \end{pmatrix},$$

and

$$A^{-1} = \frac{1}{\det(A)} C^T = \frac{1}{14} \begin{pmatrix} -1 & -2 & 5 \\ 5 & -4 & 3 \\ 4 & 8 & -6 \end{pmatrix}$$

13.6 Change of basis

A matrix can represent a change of basis from a basis $\mathcal{B}_1$ to a matrix $\mathcal{B}_2$ of the same K-vector space V. Assume that

$$\mathcal{B}_1 = \mathbf{v}_1, \ldots, \mathbf{v}_n$$
$$\mathcal{B}_2 = \mathbf{w}_1, \ldots, \mathbf{w}_n,$$

and that there are $A = (a_{i,j})_{i,j \leq n}$ and $B = (b_{i,j})_{i,j \leq n}$ such that

$$\forall j \leq n,\ \mathbf{v}_j = \sum_i a_{i,j} \mathbf{w}_i$$
$$\forall j \leq n,\ \mathbf{w}_j = \sum_i b_{i,j} \mathbf{v}_i.$$

Then, first of all, A and B are inverse of each other (i.e. $AB = BA = I_n$). Additionally, assume a (column) vector $\mathbf{v}$ is in the span of the $\mathbf{v}_i$:

$$\mathbf{v} = x_1 \mathbf{v}_1 + \ldots + x_n \mathbf{v}_n,$$

then the vector $\mathbf{y} = (y_1, \ldots, y_n)$ such that

$$\mathbf{v} = y_1 \mathbf{w}_1 + \ldots + y_n \mathbf{w}_n$$

is given by $\mathbf{y} = A\mathbf{x}$.

Example 13 *Suppose that $\mathcal{B}_1 = \{(1,0),(1,1)\}$ and $\mathcal{B}_2 = \{(0,1),(1,0)\}$ are two basese of $V = \mathbb{R}^2$. Then with the previous definitions,*

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$$

Assume that $\mathbf{v}$ is represented by the vector $\mathbf{x} = (1,1)$ in $\mathcal{B}_1$, then in $\mathcal{B}_2$, it is represented by

$$\mathbf{y} = A\mathbf{x} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

A basis of interest is always the *canonical basis.*

Definition 13.21 (Canonical basis) *Let $V = K^n$ for a field K. The canonical basis of V is $\mathbf{v}_1, \ldots, \mathbf{v}_n$ where*

$$\mathbf{v}_i = (0, \ldots, 0, 1, 0, \ldots, 0)$$

with the coefficient 1 is in i-th position.

13.7 Linear maps

A matrix can also represent the evaluations of a linear map f at the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_m$ of a basis of a K-vector space V of dimension m.

Definition 13.22 (Linear maps) *Let V, W be a K-vector spaces. We say that $f : V \to W$ is linear if*

$$\forall \mathbf{v}_1, \mathbf{v}_2 \in V \text{ and } \lambda \in K,\ f(\lambda \mathbf{v}_1 + \mathbf{v}_2) = \lambda f(\mathbf{v}_1) + f(\mathbf{v}_2).$$

Example 14 *Let $V = \mathbb{R}^2$, $W = \mathbb{R}$, and f defined by*

$$f(x_1, x_2) = x_1 + x_2.$$

Then let $\mathbf{v_1} = (x_1, x_2)$, $\mathbf{v}_2 = (y_1, y_2)$, and $\lambda \in \mathbb{R}$. We have

$$\begin{aligned} f(\lambda \mathbf{v}_1 + \mathbf{v}_2) &= f(\lambda x_1 + y_1, \lambda x_2 + y_2) \\ &= \lambda x_1 + y_1 + \lambda x_2 + y_2 \\ &= \lambda(x_1 + x_2) + (y_1 + y_2) \\ &= \lambda f(\mathbf{v}_1) + f(\mathbf{v}_2). \end{aligned}$$

The matrix of a linear map $f : V \to W$ with respect to a basis $\mathcal{B}_1 = \mathbf{v}_1, \ldots, \mathbf{v}_m$ of V and $\mathcal{B}_2 = \mathbf{w}_1, \ldots, \mathbf{w}_n$ of W is the matrix A whose j-th column satisfies

$$f(\mathbf{v}_j) = a_{1,m}\mathbf{w}_1 + a_{2,m}\mathbf{w}_2 + \ldots + a_{n,m}\mathbf{w}_n.$$

With that in mind, if an input $\mathbf{v} \in V$ decomposes as $\mathbf{v} = \sum_j x_j \mathbf{v}_j$, then the image $f(\mathbf{w})$ satisfies $f(\mathbf{v}) = \sum_i y_i \mathbf{w}_i$ for

$$\mathbf{y} = A\mathbf{x}.$$