Week 1.

Definition 1 *(Fields). A field $\mathbb{F}$ is a set on which two operations*

- *addition, $\mathbb{F} \times \mathbb{F} \to \mathbb{F}$, $(a, b) \mapsto a + b$ (the sum of a and b)*
- *multiplication, $\mathbb{F} \times \mathbb{F} \to \mathbb{F}$, $(a, b) \mapsto a \cdot b$ (the product of a and b)*

are defined, and such that the following conditions hold for all elements $a, b, c \in \mathbb{F}$.

(F 1) $a + b = b + a$ *and* $a \cdot b = b \cdot a$ *(Commutativity of addition and multiplication).*

(F 2) $(a+b)+c = a+(b+c)$ *and* $(a\cdot b)\cdot = a\cdot(b\cdot c)$ *(Associativity of addition and multiplication).*

(F 3) There exists distinct elements 0 and 1 in $\mathbb{F}$ such that $0 + a = a$ and $1 \cdot a = a$ (Existence of identity elements for addition and multiplication).

(F 4) For each element $a \in \mathbb{F}$, there exists an element $c \in \mathbb{F}$, called and additive inverse for a such that $a + c = 0$; and for each nonzero element $b \in \mathbb{F}\backslash\{0\}$, there exists and element $d \in \mathbb{F}$, called a multiplicative inverse for b, such that $b \cdot d = 1$ (Existence of inverse for addition and multiplication).

(F 5) $a \cdot (b + c) = a \cdot b + a \cdot c$ *(Distributivity of multiplication over addition).*

Remark 1. *Roughly speaking, a field is a set*

1. *Containing distinct elements $0, 1$ and possibly others.*
2. *Having four operations (addition, multiplication, subtraction, and division) so that, with the exception of division by zero, the sum, product, difference, and quotient of any two elements in the set is an element of the set.*
3. *Satisfying the "obvious" algebraic laws (commutativity, associativity, distributivity, existence of identities and inverses elements for addition and multiplication)*

Definition 2 *(Vector Spaces). A vector space over a field $\mathbb{F}$ is a set V on which two operations*

- *addition, $V \times V \to V$, $(x, y) \mapsto x + y$ (the sum of x and y),*
- *scalar multiplication, $\mathbb{F} \times V \to V$, $(a, x) \mapsto ax$ (the product of a and x),*

are defined, and such that the following conditions hold for all elements $x, y, z \in V$ and $a, b \in \mathbb{F}$:

(VS 1) $x+y=y+x$

(VS 2) $(x+y)+z=x+(y+z)$

(VS 3) *There exists a zero vector,* 0*, in* V *such that* $x+0=x$

(VS 4) *For each element* $x \in V$*, there exists an element* $y \in V$ *called an additive inverse for* x*, such that* $x+y=0$.

(VS 5) $1x = x$

(VS 6) $(ab)x=a(bx)$

(VS 7) $a(x+y)=ax+ay$

(VS 8) $(a+b)x=ax+bx$

The elements of the field $\mathbb{F}$ *are called scalars and the elements of the vector space* V *are called vectors.*

Remark 2*. Thanks to conditions (VS 1) and (VS 2), vector addition is equal irrespective of parentheses around terms, so simply write* $x+y+z+w+\cdots$*, omitting all parentheses.*

Remark 3*.* $\mathbb{Q}^n$ *is a vector space over* $\mathbb{Q}$*,* $\mathbb{R}^n$ *is a vector space over* $\mathbb{Q}$ *and* $\mathbb{R}$*,* $\mathbb{C}^n$ *is a vector space over* $\mathbb{Q}$*,* $\mathbb{R}$*, and* $\mathbb{C}$*.*

Definition 3 (Matrices)*. Let* $\mathbb{F}$ *be a field. Let* $m, n \geq 1$ *be fixed integers. An* $m \times n$ *matrix with entries from the field* $\mathbb{F}$ *is a rectangular array of the form*

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1,n} \\ a_{21} & a_{22} & \cdots & a_{2,n} \\ \vdots & \vdots & & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{pmatrix}$$

where $a_{ij} \in \mathbb{F}$ *for* $1 \leq i \leq m, 1 \leq j \leq n$*. We abbreviate the notation for this matrix by writing* $(a_{ij}), i = 1, \ldots, m$ *and* $j = 1, \ldots, n$*. We call* a_{ij} *the* ij*-entry of the matrix, i.e. the entry at the* i*th row and the* j*th column. The* $m \times n$ *matrix whose entries are all zero is called the zero matrix, denoted* O*. Two* $m \times n$ *matrices* $A = (a_{ij})$ *and* $B = (b_{ij})$ *are said to be equal if all their corresponding entries are equal, that is if* $a_{ij} = b_{ij}$ *for all* $1 \leq i, \leq m$*,* $1 \leq j \leq n$*.*

Matrix addition: Let $A = (a_{ij})$ *and* $B = (b_{ij})$ *be two* $m \times n$ *matrices with entries from* $\mathbb{F}$*. We defined* $A+B$ *to be an* $m \times n$ *matrix whose entries are the sum of corresponding entries of* A *and* B*. That is,* $(A+B)_{ij} \stackrel{\text{def}}{=} a_{ij} + b_{ij}$ *for all* $1 \leq i \leq m, 1 \leq j \leq n$*.*

Scalar multiplication: Let $A = (a_{ij})$ *be an* $m \times n$ *matrix with entries from* $\mathbb{F}$ *and* $c \in \mathbb{F}$*. We define* cA *to be an* $m \times n$ *matrix whose entries are the corresponding ones of* A*, multiplied by* c*. That is,* $(cA)_{ij} = c(A_{ij})$ *for* $1 \leq i \leq m, 1 \leq j \leq n$*.*

Example 3 *(The Space of Matrices). Denote* $\mathsf{M}_{m \times n}(\mathbb{F})$ *the set of all* $m \times n$ *matrices with entries from the field* $\mathbb{F}$*. Then the set* $\mathsf{M}_{m \times n}$ *with matrix addition and scalar multiplication is a vector space over* $\mathbb{F}$*.*

Example 4 *(Function Spaces). Let D be any nonempty set and $\mathbb{F}^D$ be the set of all functions from D to $\mathbb{F}$. The set $\mathbb{F}^D$ is a vector space with the following operations*

$$(f+g)(x) \overset{\text{def}}{=} f(x)+g(x) \qquad \textit{and} \qquad (cf)(x) \overset{\text{def}}{=} cf(x), x \in D,$$

for all $f, g \in \mathbb{F}^D$ and $c \in \mathbb{F}$.

Definition 4 *(Polynomials). A polynomial with coefficients from a field $\mathbb{F}$ is an expression of the form*

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$$

where n is a non-negative integer, $a_k \in \mathbb{F}$ for $0 \le k \le n$ (called the coefficient of x^k) and x is variable. The polynomial in which all coefficients are zero is called the zero polynomial $f(x) = 0$. Two polynomials from a field $\mathbb{F}$,

$$f(x) = a_n x^n + \cdots + a_1 x + a_0 \qquad \textit{and} \qquad g(x) = b_m x^m + \cdots + b_1 x + b_0$$

are said to be equal if $m = n$ and $a_k = b_k$ for $0 \le k \le n$.

Denote $\mathsf{P}_n(\mathbb{F})$ the set of all polynomials of degree at most n.

$$\mathsf{P}_n(\mathbb{F}) = \{a_n x^n + \cdots + a_1 x + a_0 : a_0, \ldots, a_n \in \mathbb{F}\}$$

Denote $\mathbb{F}[x]$ the set of all polynomials (of all degrees) with coefficients from $\mathbb{F}$.

Polynomial addition: Let $f, g \in \mathbb{F}[x]$

$$f(x) = a_n x^n + \cdots + a_1 x + a_0 \qquad \textit{and} \qquad g(x) = b_m x^m + \cdots + b_1 x + b_0$$

and assume without loss of generality $n \ge m$. Then let $b_j = 0$ for $j > m$, such that

$$g(x) = b_n x^n + \cdots + b_m x^m + \cdots + b_1 x + b_0$$

We define the sum $f + g$ as

$$(f+g)(x) \overset{\text{def}}{=} (a_n + b_n)x^n + \cdots + (a_1 + b_1)x + (a_0 + b_0)$$

Scalar Multiplication: Let $f \in \mathbb{F}[x]$, $f(x) = a_n x^n + \cdots + a_1 x + a_0$ and $c \in \mathbb{F}$. We define $cf \in \mathbb{F}[x]$ as

$$(cf)(x) \overset{\text{def}}{=} ca_n x^n + \cdots + ca_1 x + ca_0$$

Example 5 *(The Space of Polynomials). The set $\mathbb{F}[x]$ with the above operations of addition and scalar multiplication is a vector space over $\mathbb{F}$.*

Theorem 1.1 *(Cancellation Law for Vector Addition). Let V be a vector space. If $x, y, z \in V$ such that $x + z = y + z$, then $x = y$.*

Corollary 1.1.1*. Let V be a vector space and $x \in V$.*

1. *There is exactly one vector in V that can be the zero vector.*

2. *There is exactly one vector $y \in V$ such that $x + y = 0$, y is called the additive inverse.*

Definition 5. *Let V be a vector space and $x, z \in V$.*

- *Denote $-x$ the unique vector $y \in V$ such that $x + y = 0$.*
- *Denote $x - z$ the sum of $x + (-z)$.*

Theorem 1.2. *Let V be a vector space over $\mathbb{F}$, $x \in V$, and $a \in \mathbb{F}$. Then we have the following ($\underline{0}$ denotes the zero vector)*

1. $0x = \underline{0}$.
2. $a\underline{0} = \underline{0}$.
3. $(-a)x = -(ax) = a(-x)$. *In particular* $(-1)x = -x$

Definition 6. *A subset W of a vector space V over a field $\mathbb{F}$ is called a subspace of V if the following three conditions hold for the operations defined in V:*

(S 1) $W \neq \emptyset$, *W is nonempty.*

(S 2) If $x \in W$ and $y \in W$, then $x + y \in W$, W is closed under vector addition.

(S 3) If $c \in \mathbb{F}$ and $x \in W$, then $cx \in W$, W is closed under scalar multiplication.

Theorem 1.3. *If W is a subspace of a vector space V over a field $\mathbb{F}$, then W is a vector space over $\mathbb{F}$ under the operations of V restricted to W.*

Remark 4. *To check condition $(S1)$ in the definition of subspaces, we normally check whether $0 \in W$ or not.*

Week 2.

Definition 7. *Let V be a vector space over $\mathbb{F}$ and S a nonempty subset of V. A vector $x \in V$ is called a linear combination of vectors of S if there exists a finite number of vectors $u_1, \ldots, u_n \in S$ and scalars $a_1, \ldots, a_n \in \mathbb{F}$ such that*

$$x = a_1 u_1 + \cdots + a_n u_n$$

Note $n \geq 1$. We also say that x is linear combination of $u_1, \ldots, u_n$ and call $a_1, \ldots, a_n$ the coefficients of the linear combination.

Define the span of S, $\operatorname{span}(S)$, to be the set of all linear combinations of vectors in S.

For convenience, we define the span of the empty set to be $\operatorname{span}(\emptyset) = \{\underline{0}\}$.

Solving systems of linear equations by elimination. *The method of elimination can be used to solve systems of linear equations over any field. In general, the "allowed simplifications" of this method are:*

- *Add a scalar multiple of one equation to another equation.*
- *Multiply an equation by a nonzero scalar.*
- *Swap two equations.*

Only eliminate all occurrences (but one) of a variable once.

If an equation of the form $0 = c$ *is obtained where* $c \neq 0$*, then no solutions.*

(If an equation of the form $0 = 0$ *is found, you can delete it.)*

If no contradiction is found, then replace the variables not eliminated by parameters, move the parameters to the right-hand side, and add dummy equations ai = (parameter for a_i*) to them. The resulting equations should describe all solutions to the original system.*

***Theorem 1.4**. Let* S *be a subset of a vector space* V*. Then* $\operatorname{span}(S)$ *is a subspace of* V*. Moreover,* $\operatorname{span}(S)$ *is the smallest subspace of* V *which contains* S*, in the sense that*

1. $\operatorname{span}(S)$ *is a subspace of* V *containing* S*, and*
2. *If* W *is any other subspace of* V *containing* S*, then* $\operatorname{span}(S) \subseteq W$

***Definition 8**. Let* V *be a vector space and* S *be a subset of* V*. We say that* S *generates (or spans)* V *if* $\operatorname{span}(S) = V$*.*

***Remark 6**. Let* V *be a vector space and* S *a subset of* V*. Since* $\operatorname{span}(S)$ *is a subset of* V*, to prove* $\operatorname{span}(S) = V$*, it is sufficient to prove that every vector in* V *can be written as a linear combination of vectors in* S*.*

***Definition 9**. Let* V *be a vector space and* S *be a subset of* V*.*

- *The set* S *is called linearly dependent if there exist a finite number of distinct vectors* $u_1, \ldots, u_n$ *in* S *and scalars* $c_1, \ldots, c_n$*, <u>not all zero</u>, such that*

  $$c_1 u_1 + \cdots + c_n u_n = 0$$

  In this case, we also say that the vectors of S *are linearly dependent. Note that* $n \geq 1$*.*

- *The set* S *is called linearly independent if* S *is not linearly dependent. That is, for every choice of distinct* $u_1, \ldots, u_n \in S$ *with* $n \geq 1$*, whenever* $c_1, \ldots, c_n \in \mathbb{F}$ *are scalars satisfying*

  $$c_1 u_1 + \cdots + c_n u_n = 0$$

  then $c_i = 0$*, for all* $i = 1, \ldots, n$*. In this case we also say that the vectors of* S *are linearly independent.*

Remark 7

1. *For any vectors $u_1, \ldots, u_n$ in V, we always have the following representation of $0 \in V$ as a linear combination of $u_1, \ldots, u_n$:*

$$0u_1 + \cdots + 0u_n = 0$$

(all coefficients are 0). This is called the trivial representation of 0 as a linear combination of $u_1, \ldots, u_n$.

2. *The empty set is linearly independent since linearly dependent sets must be nonempty by definition.*

3. *The set $S = \{0\}$ is linearly dependent.*

4. *When S is a finite nonempty set, $S = \{u_1, \ldots, u_n\}$, where care has been taken to list the elements of S without repeats, then the definitions of linear independence and linear dependence can be simplified as follows:*

   - *$\{u_1, \ldots, u_n\}$ is linearly dependent if and only if there exist $(c_1, \ldots, c_n) \in \mathbb{F}^n$, $(c_1, \ldots, c_n) \neq (0, \ldots, 0)$ with*

   $$c_1u_1 + \cdots + c_nu_n = 0$$

   In other words, the equation witnessing linear dependence can be assumed to mention all of the vectors in S.

   - *$\{u_1, \ldots, u_n\}$ is linear independent if and only if the following condition is satisfied: Whenever $c_1, \ldots, c_n \in \mathbb{F}$ are such that*

   $$c_1u_1 + \cdots + c_nu_n = 0$$

   then $c_i = 0$ for all $i = 1, \ldots, n$. In other words, the definition of linear independence only needs to be checked for linear combinations involving all of the vectors in S.

5. *Any subset of a vector space that contains the zero vector is linearly dependent.*

Theorem 1.5 *Let S be a subset of a vector space V. Then S is linearly dependent if and only if either $S = \{0\}$ or some vector in S is a linear combination of other vectors in S.*

Week 3.

Definition 10 *(Basis). Let V be a vector space. A subset S of V is called a basis for V if it satisfies the following two conditions:*

1. *S is linearly independent.*

2. *S spans V.*

If S is a basis for V, we also say that the vectors of S form a basis for V.

Example 17 *(Standard Bases).*

1. *The empty set is the standard (and only) basis for the zero vector space for the zero vector space.*
2. *In $\mathbb{F}^n$, the set $\{e_1, \ldots, e_n\}$, where $e_j \in \mathbb{F}^n$ is the vector whose jth coordinate is 1 and other coordinates are 0 is the standard basis of $\mathbb{F}^n$*
3. *In $\mathsf{M}_{m\times n}(\mathbb{F})$, the set $\{E_{ij} \in \mathsf{M}_{m\times n}(\mathbb{F}) : 1 \leq i \leq m, 1 \leq j \leq n\}$ where E_{ij} is an $m \times n$ where the ijth entry is 1, and all other entries are 0 is the standard basis of $\mathsf{M}_{m\times n}(\mathbb{F})$.*
4. *In $\mathsf{P}_n(\mathbb{F})$, the set $\{1, x, \ldots, x^n\}$ is the standard basis of $\mathsf{P}_n(\mathbb{F})$.*
5. *In $\mathbb{F}[x]$, the set $\{1, x, x^2, \ldots\}$ is the standard basis of $\mathbb{F}[x]$.*

Theorem 1.6. *Let $\{v_1, \ldots v_n\}$ be a basis for a vector space V. Then every $x \in V$ can be uniquely expressed as a linear combination of $v_1, \ldots, v_n$. That is, there is a unique n-tuple $(a_1, \ldots, a_n) \in \mathbb{F}^n$ such that $x = a_1v_1 + \cdots + a_nv_n$.*

Definition 11

- *A set is countably infinite if there is a 1-1 mapping between the set and $\mathbb{N}$. For example, $\mathbb{Z}$ and $\mathbb{Q}$ are countably infinite.*
- *A set is countable if it is a finite set or it is countably infinite.*
- *A set is uncountable if it is not countable. For example, $\mathbb{R}$, $\mathbb{C}$, and $(0, 1)$ are uncountable.*

Theorem 1.7. *If a vector space V is generated by a countable set S, then some subset of S is a basis for V.*

Theorem 1.8 *(Existence Theorem). Every vector space has a basis.*

Theorem 1.9 *(Replacement Theorem). Suppose V is a vector space with a finite spanning set S. Let T be a linearly independent subset in V. Then*

1. *$|T| \leq |S|$.*
2. *There exists a set $H \subseteq S$ containing exactly $(|S|-|T|)$ vectors such that $T \cup H$ generates V.*

Corollary 1.9.1 *Suppose V is a finitely spanned vector space. Then all bases of V are finite and have the same number of elements.*

Definition 12.

- *A vector space is called finite-dimensional if it has a basis consisting of a finite number of vectors.*
- *Let V be a finite-dimensional vector space. The unique number of vectors in each basis for V is called the dimension of V and is denoted by* $\dim V$.
- *Convention:* $\dim\{0\} = 0$.
- *A vector space that is not finite-dimensional is called infinite-dimensional*

To find the dimension of a vector space, one can find a basis for that vector space and count the number of elements in that basis.

Corollary 1.9.2 *Let V be a vector space with dimension n.*

1. *Any finite spanning set for V contains at least n vectors.*
2. *A generating set for V that contains exactly n vectors is a basis for V.*
3. *Any linearly independent subset of V has at most n vectors.*
4. *Any linearly independent subset of V that contains exactly n vectors is a basis for V.*
5. *Every linearly independent subset of V can be extended to a basis for V.*
6. *Let W be a subspace of V. Then $\dim W \leq \dim V$. The equality happens if and only if $W = V$.*
7. *Let W be a subspace of V. Then any basis for W can be extended to a basis for V.*

Week 4.

Definition 13

- *For each $x \in V$, we define $x + W$ the following subset of V:*

$$x + W = \{x + w : w \in W\}$$

  The set $x + W$ is called a coset of W in V and x is called a representative of the coset $x + W$.

- *For $x, y \in V$, if $x - y \in W$, we write $x \equiv y \pmod{W}$.*
- *Denote V/W (pronounced "V mod W") the collection of cosets of W in V:*

$$V/W = \{x + W : x \in V\}$$

Proposition 1. *Let W be a subspace of a vector space V and $x, y \in V$.*

1. $x \in x + W$.

2. $x + W = y + W$ *if and only if* $x - y \in W$. *In particular,* $x + W = W$ *if and only if* $x \in W$.

Remark 8.

1. *The relation* $\equiv \pmod W$ *is an equivalence relation on* V. *That is,* $\equiv \pmod W$ *is reflexive, symmetric, and transitive.*

2. *It's easiest to think of* V/W *as the collection of cosets of* W *in* V, *it can be difficulty to visualize.*

Original Domain	$\mathbb{Z}$	*Vector space* V
Modding by	m	*subspace* W
Equivalence relation	$\equiv \pmod m$	$\equiv \pmod W$
Equivalence class	$[k]$	$x + W$
Set of equivalence class	$\mathbb{Z}_m$	V/W

3. *The construction of* V/W *is analogous to the construction of* $\mathbb{Z}_m$

Definition 14. *Let* V *be a vector space over* $\mathbb{F}$ *and* W *be a subspace of* V. *Operations of addition and scalar multiplication by* $\mathbb{F}$ *are defined naturally on* V/W *by representatives:*

$$(x + W) + (y + W) \stackrel{\text{def}}{=} (x + y) + W$$
$$a(x + W) \stackrel{\text{def}}{=} (ax) + W$$

for any $a \in \mathbb{F}$ *and* $x, y \in V$.

Lemma 1. *Under the assumptions in definition 14, the two operations are well-defined. It means for elements in* V/W,

1. *If* $x_1 + W = x_2 + W$ *and* $y_1 + W = y_2 + W$ *then* $(x_1 + W) + (y_1 + W) = (x_2 + W) + (y_2 + W)$

2. *If* $x_1 + W = x_2 + W$, *then* $a(x_1 + W) = a(x_2 + W)$ *for any* $a \in \mathbb{F}$.

Theorem 1.10 *(Quotient Space). The set* V/W *with the two operations defined in definition 14 is a vector space over* $\mathbb{F}$. *The vector space* V/W *is called the quotient space of* V *by* W.

Theorem 1.11. *Let* V *be a finite dimensional vector space and* W *be a subspace of* V. *Let* $\{v_1, \dots, v_n\}$ *be a basis for* V *such that* $\{v_1, \dots, v_k\}$ *is a basis for* W $(k \le n)$. *Then*

1. *The set* $\{v_{k+1} + W, \dots, v_n + W\}$ *is a basis for* V/W

2. $\dim(V/W) = \dim V - \dim W$

Remark 9. *There are also cases where both* V *and* W *are infinite dimensional, but* $\dim(V/W)$ *is finite. For example,* $V = \mathbb{F}^\infty$ *and* $W = \{(0, x_2, x_3, \dots) : x_k \in \mathbb{F}\}$. *Note that any element of* V/W *is just determined by the value of the first coordinate* x_1, *hence* $\dim(V/W) = 1$.

Definition 15. *Let* V *be a vector space over* $\mathbb{F}$ *and* W_1, W_2 *be subspaces of* V.

1. *Define the sum of subspaces* W_1 *and* W_2 *as follows:*

$$W_1 + W_2 \stackrel{\text{def}}{=} \{v_1 + v_2 : v_1 \in W_1, v_2 \in W_2\}$$

2. *If in addition,* $W_1 \cap W_2 = \{0\}$ *we say* W_1 *and* W_2 *are independent, or disjoint, and we write* $W_1 \oplus W_2$ *for* $W_1 + W_2$*. The set* $W_1 \oplus W_2$ *is also called the (internal) direct sum of the subspaces* W_1 *and* W_2*.*

3. *If* $W_1 \oplus W_2 = V$ *(i.e.* $W_1 + W_2 = V$ *and* $W_1 \cap W_2 = \{0\}$*), then* W_2 *is called a complementary subspace of* W_1*.*

Remark 10*. Let* V *be a vector space over* $\mathbb{F}$ *and* W_1, W_2 *be subspaces of* V*. Then the direct sum of* W_1 *and* W_2*,* $W_1 \oplus W_2$ *is defined whenever* $W_1 \cap W_2 = \{0\}$*.*

Lemma 2*. Let* V *be a vector space over* $\mathbb{F}$ *and* W_1, W_2 *be subspaces of* V*. Then*

1. $W_1 \cap W_2$ *is a subspace of* W_1*,* W_2*, and* V*.*

2. $W_1 + W_2$ *is the smallest subspace of* V *containing* W_1 *and* W_2*.*

3. $V = W_1 \oplus W_2$ *if and only if for every vector* v *in* V*, there exists unique elements* $w_1 \in W_1$ *and* $w_2 \in W_2$ *so that* $v = w_1 + w_2$*.*

Theorem 1.12*. Let* V *be a vector space over* $\mathbb{F}$ *and* W_1, W_2 *be two finite dimensional subspaces of* V*. Then*

1. $W_1 + W_2$ *is finite dimensional and*

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

2. *If* V *is finite dimensional and* $W_1 \oplus W_2 = V$*, then*

$$\dim W_1 + \dim W_2 = \dim V$$

Remark 11*.*

1. *(Existence of Complementary Subspaces) Every linearly independent subset of a vector space* V *can be extended to a basis for* V *(regardless of if the set countable or not or of the dimension of* V*). Therefore, every subspace of a vector space* V *has a complementary subspace.*

2. *This complementary subspace is not necessarily unique.*

Week 5.

Definition 16. *Let V and W be vector spaces over $\mathbb{F}$. A function $T : V \to W$ is called a linear transformation from V to W, or is said to be linear if, for all $x, y \in V$, we have*

(L 1) $T(x+y) = T(x) + T(y)$ *and*

(L 2) $T(cx) = cT(x)$

Proposition 2. *Let $T : V \to W$ be a mapping. T is linear if and only if $T(cx+y) = cT(x) + T(y)$ for all $x, y \in V$ and $c \in \mathbb{F}$.*

Example 23 *Let V and W be vector spaces. The following function are linear:*

(a) $T_0 : V \to W, T_0(x) = 0$ *for all $x \in V$. The function T_0 is called the zero transformation.*

(b) $I_V : V \to V, I_V(x) = x$ *for all $x \in V$. The function I_V is called the identity transformation.*

Proposition 3. *Let $T : V \to W$ be linear. Then*

1. $T(0) = 0$.
2. $T(x-y) = T(x) - T(y)$ *for all $x, y \in V$.*
3. $T(a_1x_1 + \cdots + a_nx_n) = a_1T(x_1) + \cdots + a_nT(x_n)$.

Theorem 2.1. *Let $\{v_1, \ldots, v_n\}$ be a basis for a vector space V, and let $\{w_1, \ldots, w_n\}$ be arbitrary elements of a vector space W. Then there exists a unique linear mapping $T : V \to W$ such that*

$$T(v_1) = w_1, \ldots, T(v_n) = w_n$$

Corollary 2.1.1. *Let $\{v_1, \ldots, v_n\}$ be a basis for a vector space V and $\{w_1, \ldots, w_n\}$ be arbitrary elements of a vector space W. Let $T : V \to W$ be the unique linear mapping such that*

$$T(v_1) = w_1, \ldots, T(v_n) = w_n.$$

Then for all $a_1 \in \mathbb{F}$, we have

$$T(a_1v_1 + \cdots + a_nv_n) = a_1w_1 + \cdots + a_nw_n$$

Definition 17. *Let V and W be vector spaces, and let $T : V \to W$ be linear. Define the following sets:*

- *Null space (or kernel) of T: $\mathcal{N}(T) \stackrel{\text{def}}{=} \{x \in V : T(x) = 0_W\}$.*

- *Range (or image) of T:* $\mathcal{R}(T) \stackrel{\text{def}}{=} \{T(x) : x \in V\}$

Theorem 2.2. *Let $T : V \to W$ be linear. Then $\mathcal{N}(T)$ is a subspace of V and $\mathcal{R}(T)$ is a subspace of W.*

Example 24. *The differential operator is defined as $D_n : \mathsf{P}_n(\mathbb{R}) \to \mathsf{P}_n(\mathbb{R}), D_n(p(x)) = p'(x) = \frac{d}{dx}p(x)$.*

Theorem 2.3. *Let V and W be vector spaces, and let $T : V \to W$ be linear. If $\{v_1, \dots, v_n\}$ is basis for V, then $\{T(v_1), \dots, T(v_n)\}$ generates $\mathcal{R}(T)$.*

Definition 18. *Let $T : V \to W$ be linear. If $\dim(\mathcal{N}(T)) < \infty$, define $\text{nullity}(T) \stackrel{\text{def}}{=} \dim(\mathcal{N}(T))$. If $\dim(\mathcal{R}(T)) < \infty$, define $\text{rank}(T) \stackrel{\text{def}}{=} \dim(\mathcal{R}(T))$.*

Theorem 2.4 *(Rank-Nullity Theorem). Let V and W be vector spaces and $T : V \to W$ be linear. If $\dim(V) < \infty$, then*

$$\text{rank}(T) + \text{nullity}(T) = \dim(V)$$

Definition 19. *Let $T : V \to W$ be linear. Then*

- *T is called one-to-one (or injective) if $T(x) = T(y)$ implies $x = y$, or equivalently $x \neq y$ implies $T(x) \neq T(y)$.*
- *T is called onto (or surjective) if $\mathcal{R}(T) = W$.*
- *T is called an isomorphism (or a bijection) if T is one-to-one and onto.*

Lemma 3. *Let $T : V \to W$ be linear. Then T is one-to-one if and only if $\mathcal{N}(T) = \{0\}$.*

Theorem 2.5. *Let W be a vector space over a field $\mathbb{F}$ and let V be a finite-dimensional vector space over $\mathbb{F}$ with a basis $\{v_1, \dots, v_n\}$. Consider a linear transformation $T : V \to W$. Then T is an isomorphism if and only if $\{T(v_1), \dots, T(v_n)\}$ is a basis for W.*

Remark. *From theorem 2.5, to construct an isomorphism (if there exists one) between two finite-dimensional vector spaces, we choose a basis $\{v_1, \dots, v_n\}$ for V and a basis $\{w_1, \dots, w_n\}$ for W. Then define a linear transformation $T : V \to W$ such that $T(v_k) = w_k$ for $k = 1, \dots, n$. By theorem 2.1, such a linear transformation exists. By theorem 2.5, T is an isomorphism.*

Definition 20. *Let V and W be two vector spaces over a field $\mathbb{F}$. The vector space V is said to be isomorphic to the vector space W if there is an isomorphism $T : V \to W$. We write $V \cong W$.*

Theorem 2.6. *Let V and W be two finite-dimensional vector spaces over a field $\mathbb{F}$. Then V is isomorphic to W if and only if $\dim V = \dim W$.*

Theorem 2.7. *Let V and W be two vector spaces over a field $\mathbb{F}$ of equal finite dimension. Let $T : V \to W$ be linear. Then the following statements are equivalent*

1. *T is one-to-one.*
2. *T is onto.*
3. $\text{rank}T = \dim V$.

Week 6.

Definition 21. *Let V and W be vector spaces over $\mathbb{F}$. We let $\mathcal{L}(V, W)$ denote the set of all linear transformation $T : V \to W$.*

Theorem 2.8. *Let V and W be vector spaces over $\mathbb{F}$. Then $\mathcal{L}(V, W)$ is a subspace of W^V.*

Definition 22 *(Matrix-Vector Multiplication). Let $A \in \mathsf{M}_{m\times n}(\mathbb{F})$ and $x \in \mathbb{F}^n$. We consider x as an $n \times 1$ column vector*

$$x = (x_1, \dots, x_n)^T = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}.$$

Then Ax is the $m \times 1$ column vector in $\mathbb{F}^m$ defined as:

$$Ax = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1,n} \\ a_{21} & a_{22} & \cdots & a_{2,n} \\ \vdots & \vdots & & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} \sum_{k=1}^{n} a_{1,k}x_k \\ \sum_{k=1}^{n} a_{2,k}x_k \\ \vdots \\ \sum_{k=1}^{n} a_{m,k}x_k \end{bmatrix}$$

That is the ith entry of Ax is the entries in the ith row of A each multiplies by the corresponding entries in x, then summed up.

Definition 23. *If $\mathbb{F}$ is a field and $A \in \mathsf{M}_{m\times n}(\mathbb{F})$, then L_A denotes the function $\mathbb{F}^n \to F^m$ given by $L_A(x) = Ax$.*

Matrix notation. *Let $A \in \mathsf{M}_{m\times n}(\mathbb{F})$*

- *Recall a_{ij} denotes the entry of A in the ith row and jth column. Thus $a_{ij} \in \mathbb{F}$.*
- *We often write a_j to denote the jth column of A. Thus $a_j \in \mathbb{F}^m$ and*

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1,j} & \cdots & a_{1,n} \\ a_{21} & a_{22} & \cdots & a_{2,j} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots & & \vdots \\ a_{i,1} & a_{i,2} & \cdots & a_{i,j} & \cdots & a_{i,n} \\ \vdots & \vdots & & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,j} & \cdots & a_{m,n} \end{bmatrix} \quad \text{and} \quad a_j = \begin{bmatrix} a_{1,j} \\ a_{2,j} \\ \vdots \\ a_{i,j} \\ \vdots \\ a_{m,j} \end{bmatrix}$$

- *We also write $a = [a_1\ a_2\ \cdots\ a_n]$ when $a_1, \ldots, a_n \in \mathbb{F}^m$ are the columns of A.*

Lemma 4. *Suppose $A \in \mathsf{M}_{m\times n}(\mathbb{F})$ and write $A = [a_1\ a_2\ \cdots\ a_n]$ where $a_1, \ldots, a_n \in \mathbb{F}^m$ are the columns of A.*

1. *For any $x = (x_1, \ldots, x_n)^T \in \mathbb{F}^n$ we have*

$$Ax = x_1a_1 + x_2a_2 + \cdots + x_na_n$$

   That is, Ax is the linear combination of the columns of A whose coefficients are the entries in x.

2. *If $e_1, \ldots, e_n$ are the standard basis vector for $\mathbb{F}^n$, then $Ae_j = a_j$ equals the jth column of A.*

Corollary 2.8.1 *(Matrix Equality Theorem). Let $A, B \in \mathsf{M}_{m\times n}(\mathbb{F})$. Then $A = b$ if and only if $Ax = Bx$ for all $x \in \mathbb{F}^n$.*

Theorem 2.9. *Let $A \in \mathsf{M}_{m\times n}(\mathbb{F})$. Then the function $L_A : \mathbb{F}^n \to \mathbb{F}^m$ is a linear transformation.*

Proposition 4. *In the above situation, $\mathsf{L} : \mathsf{M}_{m\times n}(\mathbb{F}) \to \mathcal{L}(F^n, F^m)$ is a one-to-one linear transformation.*

Definition 24. *Let V be a finite dimensional vector space. An ordered basis for V is a basis $\{v_1, \ldots, v_n\}$ endowed with a specific order.*

Definition 25. *Let $\beta = \{u_1, \ldots, u_n\}$ be an ordered basis for a finite dimensional vector space V. For $x \in V$, let $a_1, \ldots, a_n$ be the unique scalars such that $x = a_1u_1 + \cdots + a_nu_n$. We define the co-ordinate vector of x relative to β to be*

$$[x]_\beta \stackrel{\text{def}}{=} \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} \in \mathbb{F}^n$$

Theorem 2.10. *Let V be an n-dimensional vector space over $\mathbb{F}$ and β be an ordered basis for V. The map $[\ \ ]_\beta : V \to \mathbb{F}^n$ is an isomorphism.*

Definition 26. *Let V and W be finite dimensional vector spaces over $\mathbb{F}$ and let $T : V \to W$ be a linear transformation. Let $\beta = \{v_1, \ldots, v_n\}$ and $\gamma = \{w_1, \ldots, w_n\}$ be ordered bases for V and W respectively. The matrix representation of T in the ordered bases β and γ is the matrix $[T]_\beta^\gamma$ with entries from $\mathbb{F}$ defined as*

$$[T]_\beta^\gamma \stackrel{\text{def}}{=} \Big[[T(v_1)]_\gamma\ [T(v_2)]_\gamma\ \cdots\ [T(v_n)]_\gamma\Big].$$

When $T : V \to V$ is linear and β is an ordered basis of the finite dimensional vector space V, denote $[T]_\beta = [T]_\beta^\beta$.

Remark 12 *Under the assumptions of definition 26 for $T : V \to W$, if we denote $A = [T]_\beta^\gamma$, then*

1. *$A \in \mathsf{M}_{m\times n}(\mathbb{F})$, where (# of rows of A) $= m = \dim W$ and (# of columns of A) $= n = \dim V$.*

2. *For all $j = 1, \dots, n$, the jth column of A is $[T(v_j)]_\gamma$. If we write $A = (a_{ij})$ as usual, then the jth column of A is $(a_{1j}, a_{2j}, \dots, a_{mj})^T$, so by the definition of $[T(v_j)]_\gamma$ we have*

$$T(v_j) = \sum_{k=1}^{m} a_{kj} w_k$$

Example 33 *Let $A \in \mathsf{M}_{m\times n}(\mathbb{F})$ and consider $L_A : \mathbb{F}^n \to \mathbb{F}^m$. Let β and γ be the standard ordered bases of $\mathbb{F}^n$ and $\mathbb{F}^m$ respectively. Then $[L_A]_\beta^\gamma = A$*

Theorem 2.11*. Let $T : V \to W$ be linear and $\beta = \{v_1, \dots, v_n\}$ and $\gamma = \{w_1, \dots, w_m\}$ be ordered bases of V and W respectively. Then*

$$[T(x)]_\gamma = [T]_\beta^\gamma \cdot [x]_\beta \qquad \forall x \in V$$

Proposition 5*. Let V and W be finite dimensional vector spaces over $\mathbb{F}$ and let β and γ be ordered bases of V and W respectively.*

1. *For $T, U \in \mathcal{L}(V, W)$ and $c \in \mathbb{F}$, we have*

$$[T + U]_\beta^\gamma = [T]_\beta^\gamma + [U]_\beta^\gamma, \qquad [cT]_\beta^\gamma = c[T]_\beta^\gamma$$

2. *For every $C \in \mathsf{M}_{m\times n}(\mathbb{F})$ there exists a unique $T \in \mathcal{L}(V, W)$ such that $[T]_\beta^\gamma = C$.*

In other words, the map $[\quad]_\beta^\gamma : \mathcal{L}(V, W) \to \mathsf{M}_{m\times n}(\mathbb{F})$ is an isomorphism, where $m = \dim(W)$ and $n = \dim(V)$.

Corollary 2.11.1*. The map $\mathsf{L} : \mathsf{M}_{m\times n}(\mathbb{F}) \to \mathcal{L}(\mathbb{F}^n, \mathbb{F}^m)$ is an isomorphism.*

Definition 27*. Let $\mathbb{F}$ be a field. Suppose $A \in \mathsf{M}_{m\times n}(\mathbb{F})$ and $B \in \mathsf{M}(n \times p)(\mathbb{F})$. (Note that the number of columns in A equals the number of rows in B; this is required.) The matrix product AB is the $m \times p$ matrix $C \in \mathsf{M}_{m\times n}(\mathbb{F})$ whose row-i, column-j entry is the sum of products formed multiplying the entries in the ith row of A with the entries of jth column of B. That is*

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1,k} & \cdots & a_{1,n} \\ a_{21} & a_{22} & \cdots & a_{2,k} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots & & \vdots \\ \boxed{a_{i,1}} & a_{i2} & \cdots & a_{i,k} & \cdots & a_{i,n} \\ \vdots & \vdots & & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,k} & \cdots & a_{m,n} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1,j} & \cdots & b_{1,p} \\ b_{21} & b_{22} & \cdots & b_{2,j} & \cdots & b_{2,p} \\ \vdots & \vdots & \ddots & \vdots & & \vdots \\ b_{k,1} & b_{k,2} & \cdots & b_{k,j} & \cdots & b_{k,p} \\ \vdots & \vdots & & \vdots & \ddots & \vdots \\ b_{n,1} & b_{n,2} & \cdots & b_{n,j} & \cdots & b_{n,p} \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1,j} & \cdots & c_{1,p} \\ c_{21} & c_{22} & \cdots & c_{2,j} & \cdots & c_{2,p} \\ \vdots & \vdots & \ddots & \vdots & & \vdots \\ c_{i,1} & c_{i,2} & \cdots & \boxed{c_{i,j}} & \cdots & c_{i,p} \\ \vdots & \vdots & & \vdots & \ddots & \vdots \\ c_{m,1} & c_{m,2} & \cdots & c_{m,j} & \cdots & c_{m,p} \end{bmatrix}$$

where each entry $c_{i,j}$ of the product is given by $c_{i,j} = a_{i,1}b_{1,j} + a_{i,2}b_{2,j} + \cdots + a_{i,n}b_{n,j}$, or in summation notation,

$$c_{i,j} = \sum_{k=1}^{n} a_{i,k} b_{k,j}$$

If A and B are matrices and the number of columns of A does not equal the number of columns of B, then AB is not defined.

Remark 13.

1. *If $p = 1$ so B and AB column vectors, then the definition above is the matrix-vector product defined earlier.*
2. *In general (i.e. when B has several columns), B and AB have the same number of columns and the jth column of AB is obtained by multiplying A by the jth columns of B. That is if $B = [b_1\ b_2\ \cdots\ b_p]$ and $AB = C = [c_1\ c_2\ \cdots\ c_p]$, then $c_j = Ab_j$ for $j = 1, \dots, p$.*
3. *Combining the previous item with lemma 4(1), we see that the jth columns of AB is the linear combination of columns of A formed using the entries in the jth columns of B as coefficients.*
4. *This gives us an efficient algorithm to find the matrix product. Multiply the columns of A by the rows of B to form n $m \times p$ matrices which you then sum. (This isn't something they actually taught us, be wary of using it.)*

Matrix notation:

- *We usually use O to denote a zero matrix, i.e. a matrix in which every entry is 0. If we need to specify its number number of rows and columns, we may write $O_{m\times n}$.*
- *For each $n \geq 1$ we let I_n denote the $n \times n$ identity matrix. This is the matrix whose i, j entry is given by the Kronecker delta $\delta_{ij} = \begin{cases} 1, & \text{if } i = j \\ 0, & \text{if } i \neq j \end{cases}$ for example $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$.*

Lemma 5

1. *$A(B + C) = AB + AC$, where $A \in \mathsf{M}_{m\times n}(\mathbb{F})$ and $B, C \in \mathsf{M}_{n\times p}(\mathbb{F})$.*
2. *$(D + E)A = DA + EA$, where $A \in \mathsf{M}_{m\times n}(\mathbb{F})$ and $D, E \in \mathsf{M}_{q\times m}(\mathbb{F})$.*
3. *$\alpha(AB) = (\alpha A)B = A(\alpha B)$, $\forall \alpha \in \mathbb{F}$, where $A \in \mathsf{M}_{m\times n}(\mathbb{F})$ and $B \in \mathsf{M}_{n\times p}(\mathbb{F})$.*
4. *$(AB)^T = B^T A^T$, where $A \in \mathsf{M}_{m\times n}(\mathbb{F})$ and $B \in \mathsf{M}_{n\times p}(\mathbb{F})$.*
5. *$I_m A = AI_n = A$, where $A \in \mathsf{M}_{m\times n}(\mathbb{F})$.*
6. *$AO_{n\times p} = O_{m\times p}$ and $O_{q\times m}A = O_{q\times n}$, where $A \in \mathsf{M}_{m\times n}(\mathbb{F})$.*

Theorem 2.12. *Let $T : V \to W$ and $U : W \to Z$ be linear transformations between vector spaces. Then the composition function $U \circ T : V \to Z$ given by $(U \circ T)(x) = U(T(x))$ for $x \in V$ is also linear.*

Remark: *We usually denote $U \circ T$ by UT.*

Theorem 2.13 *(Matrix of Composition of Linear Transformations). Let V, W, and Z be finite dimensional vector spaces having ordered bases $\alpha = \{v_1, \ldots, v_p\}$, $\beta = \{w_1, \ldots, w_n\}$, and $\gamma = \{z_1, \ldots, z_m\}$, respectively. Let $T : V \to W$ and $U : W \to Z$ be linear transformations. Denote $A = [U]_\beta^\gamma \in \mathsf{M}_{m\times n}(\mathbb{F})$, $B = [T]_\alpha^\beta \in \mathsf{M}_{n\times p}(\mathbb{F})$ and $C = [UT]_\alpha^\gamma \in \mathsf{M}_{m\times p}(\mathbb{F})$. Then $C = AB$. That is, $[UT]_\alpha^\gamma = [U]_\beta^\gamma \cdot [T]_\alpha^\beta$.*

Corollary 2.13.1.

1. *$L_{AB} = L_A L_B$, whenever $A \in \mathsf{M}_{m\times n}(\mathbb{F})$ and $B \in \mathsf{M}_{n\times p}(\mathbb{F})$.*
2. *$A(BC) = (AB)C$, whenever the sizes of A, B, C make all the matrix products defined.*

Week 7.

Definition 28. *A square matrix $A \in \mathsf{M}_{n\times n}(\mathbb{F})$ is invertible if there exists a matrix $B \in \mathsf{M}_{n\times n}(\mathbb{F})$ such that $AB = BA = I_n$. We the matrix B is called the inverse of A, denoted by A^{-1}.*

Definition 29. *Let $T : V \to W$ be a linear mapping between vector spaces V and W. If there exists a function $U : W \to V$ such that $UT = I_V$ and $TU = I_W$, then T is said to be invertible and U is said to be an inverse of T.*

Lemma 6. *Suppose $T : V \to W$ is linear and invertible. Then the inverse of T is unique.*

Theorem 2.14. *Let $T : V \to W$ be linear. Then T is invertible if and only if T is an isomorphism.*

Lemma 7. *Suppose $T : V \to W$ is an isomorphism. Then T^{-1} is also linear.*

Theorem 2.15. *Let V and W be finite dimensional vector spaces, and α and β be ordered bases of V and W respectively. Let $T : V \to W$ be linear. Then*

1. *T is an isomorphism if and only if $[T]_\alpha^\beta$ is an invertible matrix.*
2. *In particular, if $A \in \mathsf{M}_{n\times n}(\mathbb{F})$, then L_A isomorphism if and only if A is invertible.*

Lemma 8.

1. *If a matrix A in invertible, then A^{-1} is also invertible and $(A^{-1})^{-1} = A$.*
2. *If A is invertible and $c \in \mathbb{F}$ with $c \neq 0$, then $(cA)^{-1}$ is also invertible and $\frac{1}{c} \cdot A^{-1}$.*
3. *If A is invertible $(A^T)^{-1} = (A^{-1})^T$*
4. *$A, B \in \mathsf{M}_{n\times n}(\mathbb{F})$ are invertible, then AB is also invertible and $(AB)^{-1} = B^{-1}A^{-1}$.*
5. *Conversely, if $A, B \in \mathsf{M}_{n\times n}(\mathbb{F})$ is invertible, then A and B are invertible matrices.*

Theorem 2.16 *(Invertible Matrix Theorem part 1). Let $A \in \mathsf{M}_{n\times n}(\mathbb{F})$. The following statement are equivalent:*

- *A is invertible.*
- *There exists a matrix $C \in \mathsf{M}_{n\times n}(\mathbb{F})$ such that $AC = I_n$.*
- *There exists a matrix $B \in \mathsf{M}_{n\times n}(\mathbb{F})$ such that $BA = I_n$.*

Theorem 2.17*. Let α and β be two ordered bases for a finite dimensional vector space V and let $Q = [I_V]_\alpha^\beta$. Then*

1. *Q is invertible, called the change of co-ordinate matrix from α to β.*
2. *For any $x \in V$, we have $[x]_\beta = Q[x]_\alpha$.*

Remark 15*. Suppose V is a finite dimensional vector space over $\mathbb{F}$. Let $\alpha = \{v_1, \dots, v_n\}$ and $\beta = \{w_1, \dots, w_n\}$ be ordered bases for V and $x \in V$. Then the change of co-ordinate matrix from α to β is*

$$[I_V]_\alpha^\beta = \Big[[v_1]_\beta \cdots [v_n]_\beta\Big]$$

Theorem 2.18*. Let $T : V \to W$ be linear and V be a finite dimensional vector space. Let α and β two ordered bases of V and Q be the change of co-ordinate matrix from α to β. Then*

$$[T]_\alpha = Q^{-1}[T]_\beta Q$$

Definition 30*. Let A and B be the matrices in $\mathsf{M}_{n\times n}(\mathbb{F})$. We say B is similar to A if there exists an invertible matrix Q such that $B = Q^{-1}AQ$*

Definition 31*. Let A be an $m\times n$ matrix. The following operations of the rows and columns of A are called elementary row / column operations:*

1. *Type 1: interchanging any two rows or columns of A: $R_i \leftrightarrow R_j$ or $C_i \leftrightarrow C_j$.*
2. *Type 2: multiplying any row or column of A by a non-zero scalar: $R_i \leftarrow aR_i$ or $C_i \leftarrow aC_i$.*
3. *Type 3: adding any scalar multiple of a row or column of A to another: $R_i \leftarrow R_i + aR_j$ or $C_i \leftarrow C_i + aC_j$.*

Definition 32*. An $n \times n$ elementary matrix is a matrix obtained by performing a single elementary matrix operation on I_n.*

Theorem 3.1*. Let $A \in \mathsf{M}_{m\times n}(\mathbb{F})$ and suppose B is obtained by performing an elementary row operation on A. Then there exists and $m \times m$ elementary matrix E such that $B = EA$. In fact, E is obtained by performing the elementary row operation I_m as was performed on A.*

Conversely, if E is an $m \times m$ elementary matrix, then EA is the matrix obtained from A by performing the same elementary row operation as that which obtained E from I_n.

Theorem 3.2. *Let $A \in \mathsf{M}_{m\times n}(\mathbb{F})$ and suppose B is obtained by performing an elementary column operation on A. Then there exists and $n \times n$ elementary matrix E such that $B = AE$. In fact, E is obtained by performing the elementary column operation I_m as was performed on A.*

Conversely, if E is an $m \times m$ elementary matrix, then EA is the matrix obtained from A by performing the same elementary column operation as that which obtained E from I_n.

Theorem 3.3. *Elementary matrices are invertible and the inverse of an elementary matrix is an elementary matrix of the same type.*

Week 8.

Definition 33. *Let $A \in \mathsf{M}_{m\times n}(\mathbb{F})$. We define the rank of the matrix A, denoted $\operatorname{rank}(A)$, to be the rank of the linear transformation $L_A : \mathbb{F}^n \to \mathbb{F}^m$ for $x \in \mathbb{F}^n$. That is $\operatorname{rank}(A) = \dim \mathcal{R}(L_A) = \dim L_A(\mathbb{F}^n)$.*

Remark 16.

1. *Let $A \in \mathsf{M}_{m\times n}(\mathbb{F})$. We have*

$$\mathcal{R}(L_A) = \operatorname{span}\{L_A(e_1), \dots, L_A(e_n)\} = \operatorname{span}\{Ae_1, \dots, Ae_n\} = \operatorname{span}\{a_1, \dots, a_n\}$$

   Therefore $\operatorname{rank}(A) = \dim \mathcal{R}(L_A) = \dim \operatorname{span}\{a_1, \dots, a_n\}$. That is the rank of a matrix is the dimension of the subspace generated by its columns.

2. *Since $\{a_1, \dots, a_n\}$ generates $\mathcal{R}(L_A)$ and any finite spanning set for $\mathcal{R}(L_A)$ contains at least $\dim \mathcal{R}(L_A) = \operatorname{rank}(A)$ vectors, we have $n \geq \operatorname{rank}(A)$.*

   Since $\mathcal{R}(L_A)$ is a subspace of $\mathbb{F}^m$, $\dim \mathcal{R}(L_A) \leq \dim(\mathbb{F}^m) = m$. Hence $\operatorname{rank}(A) \leq m$. Therefore $\operatorname{rank}(A) \leq \min(m, n)$.

Lemma 9. *Let $T : V \to W$ be a linear and one-to-one mapping from a vector space V to a vector space W. Let V_0 be a subspace of V. Then*

1. *$T(V_0) = \{T(x) | x \in V_0\}$ is a subspace of W.*
2. *If $\dim(V_0) < \infty$, then $\dim(V_0) = \dim(T(V_0))$.*

Theorem 3.4. *Let A be an $m \times n$ matrix. If P and Q are invertible $m \times m$ and $n \times n$ matrices respectively, then*

$$\operatorname{rank}(AQ) = \operatorname{rank}(PA) = \operatorname{rank}(PAQ) = \operatorname{rank}(A)$$

Corollary 3.4.1 *(Invertible Matrix Theorem Part 2). Let* $A \in \mathsf{M}_{n\times n}(\mathbb{F})$*. Then* A *is invertible if and only if* $\operatorname{rank}(A) = n$*.*

Corollary 3.4.2 *Elementary row and column operations are rank-preserving.*

Theorem 3.5*. Let* $A \in \mathsf{M}_{m\times n}(\mathbb{F})$*. Then by means of a finite number of elementary row and column operations,* A *can be transformed into the matrix*

$$D = \begin{bmatrix} I_r & O_1 \\ O_2 & O_3 \end{bmatrix}$$

where O_1, O_2, O_3 *are zero matrices. Moreover,* $r = \operatorname{rank}(A)$*.*

Corollary 3.5.1*. Let* A *be an* $m\times n$ *with* $\operatorname{rank}(A) = r$*. Then there exist invertible matrices* B *and* C *of sizes* $m\times m$ *and* $n\times n$ *respectively such that* $D = BAC$*, where* $D = \begin{bmatrix} I_r & O_1 \\ O_2 & O_3 \end{bmatrix}$ *is the* $m\times n$ *matrix in which* O_1, O_2, O_3 *are zero matrices.*

Theorem 3.6*. Let* $A \in \mathsf{M}_{m\times n}(\mathbb{F})$ *be of rank* r*. Then by means of a finite number of elementary row and column operations* A *can be transformed into the matrix*

$$D_{upper} = \begin{bmatrix} 1 & d_{12} & d_{13} & \cdots & d_{1,r} & d_{1,r+1} & \cdots & d_{1,n} \\ 0 & 1 & d_{23} & \cdots & d_{2,r} & d_{2,r+1} & \cdots & d_{2,n} \\ \vdots & \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 & d_{r,r+1} & \cdots & d_{r,n} \\ 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{bmatrix}$$

Remark 17*. Theorem 3.6 suggest a system way to transform a matrix* A *to the matrix* D_{upper}*:*

Step 1 *Find a non-zero entry of* A*.*

Step 2 *Apply at most one type-1 row operation and at most one type-1 column operation to move that entry to the* $(1,1)$ *position.*

Step 3 *Apply at most one type-2 row (or column) operation so that the entry at the (1,1) position is* $1_{\mathbb{F}}$*.*

Step 4 *Apply at most (m-1) type-3 elementary row operations so that all the remaining entries in the first are* 0*. The updated matrix is now of the form*

$$\left[\begin{array}{c|ccc} 1 & d_{12} & \cdots & d_{1n} \\ \hline 0 & & & \\ \vdots & & B & \\ 0 & & & \end{array}\right]$$

Step 5 *Repeat steps 1-4 on the matrix* B*. Continue this process until you get a matrix of the form of* D_{upper}*.*

Step 6 Then $\text{rank}(A) = r =$ *number of non-zero rows of* D_{upper}

Corollary 3.6.1 *Let A be an $m \times n$. Then*

1. $\text{rank}(A^T) = \text{rank}(A)$.
2. $\text{rank}(A) =$ *the dimension of the subspace generated by the columns of* $A =$ *the dimension of the subspace generated by the rows of* A.

Theorem 3.7 *Let A and B be matrices such that the product AB is defined. Then*

$$\text{rank}(AB) \leq \min\{\text{rank}(A), \text{rank}(B)\}$$

Definition 34. *Let $A \in \mathsf{M}_{m\times n}(\mathbb{F})$. Define*

$$\begin{aligned}\text{Col}(A) &\stackrel{\text{def}}{=} \{Ax | x \in \mathbb{F}^n\} \\ &= \{\textit{all linear combinations of columns of } A\} \\ &= \text{span}\{\textit{columns of } A\}, \ \textit{called the column space of } A\end{aligned}$$

$$\begin{aligned}\text{Row}(A) &\stackrel{\text{def}}{=} \text{Col}(A^T) = \{A^T y | y \in \mathbb{F}^m\} \\ &= \{\textit{all linear combinations of rows of } A\} \\ &= \text{span}\{\textit{rows of } A\}, \ \textit{called the row space of } A\end{aligned}$$

$$\mathcal{N}(A) \stackrel{\text{def}}{=} \{x \in \mathbb{F}^n | Ax = 0\}, \ \textit{called the null space of } A$$

$$\mathcal{N}(A^T) \stackrel{\text{def}}{=} \{y \in \mathbb{F}^m | A^T y = 0\}, \ \textit{called the left null space of } A$$

Denote $\text{nullity}(A) \stackrel{\text{def}}{=} \dim \mathcal{N}(A)$

Theorem 3.8. *Let $A \in \mathsf{M}_{m\times n}(\mathbb{F})$. Then*

1. $\text{Col}(A)$ *and* $\mathcal{N}(A^T)$ *are subspaces of* $\mathbb{F}^m$*;* $\text{Row}(A)$ *and* $\mathcal{N}(A)$ *are subspaces of* $\mathbb{F}^n$.
2. $\text{rank}(A) = \dim \text{Col}(A) = \dim \text{Row}(A)$.
3. $\text{nullity}(A^T) = m - \text{rank}(A)$ *and* $\text{nullity}(A) = n - \text{rank}(A)$.
4. *If* $\mathbb{F} = \mathbb{R}$*, then* $\mathbb{R}^m = \text{Col}(A) \oplus \mathcal{N}(A^T)$ *and* $\mathbb{R}^n = \text{Row}(A) \oplus \mathcal{N}(A)$.

Theorem 3.9 *(Invertible Matrix Theorem Part 3). Let $A \in \mathsf{M}_{n\times n}(\mathbb{F})$. Then the following statements are equivalent.*

1. *A is invertible.*

2. *The columns of A form a basis for $\mathbb{F}^n$.*

3. *The rows of A form a basis for $\mathbb{F}^n$.*

4. *A is a product of elementary matrices.*

Theorem 3.10.

1. *If A is an invertible $n \times n$ matrix, it is possible to transform $(A|I_n)$ into the matrix $(I_n|A^{-1})$ by means of a finite number of elementary row operations.*

2. *Conversely, suppose A is an $n \times n$ matrix and there exists an $n \times n$ matrix B such that $(A|I_n) \rightsquigarrow (I_n|B)$ via a finite number of elementary row operations, then A is invertible and $B = A^{-1}$.*

Remark 18. *Theorem 3.10 suggest an algorithm suggests an algorithm to check whether a square matrix is invertible or not how to find A^{-1}. It is called the Gauss-Jordan method to find the inverse of a square matrix.*

Step 1 *If the first column of A is a zero vector, A is not invertible. Otherwise the first column of $(A|I_n)$ has a non-zero entry. From now on we consider the matrix $(A|I_n)$.*

Step 2 *By means of at most one type-1 and one type-2 elementary row operation we can move that non-zero entry to the (1,1) position and its new value is 1.*

Step 3 *By means of at most $(n-1)$ type-3 row operations, we change all the remaining entries in the first row to be 0. Thus, we have transformed $(A|I_n)$ to a matrix of the form*

$$\left[\begin{array}{c|ccc} 1 & d_{12} & \cdots & d_{1,2n} \\ \hline 0 & & & \\ \vdots & & Q & \\ 0 & & & \end{array}\right]$$

Step 4 *Repeat steps 1-3 on the matrix Q until you get the matrix*

$$C' = \begin{bmatrix} 1 & d_{12} & \cdots & d_{1,n} & d_{1,n+1} & \cdots & d_{1,2n} \\ 0 & 1 & \cdots & d_{2,n} & d_{2,n+1} & \cdots & d_{2,2n} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 & d_{n,n+1} & \cdots & d_{n,2n} \end{bmatrix}$$

From steps 1-4 we have been going forward to transform $(A|I_n)$ to a matrix whose main diagonal has 1 and whose entries below the main diagonal are zero. For the remaining steps we move backwards to produce zeroes above the main diagonal.

Step 5 *By means of at most $(n-1)$ type-3 row operations we can transform all entries at of the n-th column of C' to zeroes, except for the last entry. E.g. we get $C' \rightsquigarrow C_n$*

$$C_n = \begin{bmatrix} 1 & d_{12} & \cdots & 0 & d_{1,n+1} & \cdots & d_{1,2n} \\ 0 & 1 & \cdots & 0 & d_{2,n+1} & \cdots & d_{2,2n} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 & d_{n,n+1} & \cdots & d_{n,2n} \end{bmatrix}$$

Step 6 *By means of at most $(n-2)$ type-3 row operations, we can transform all entries of the $(n-1)$-th column of C_n to zeroes, except for the entry at the $(n-1, n-1)$ position. That is $C_n \rightsquigarrow C_{n-1}$.*

Step 7 *We can continues this process until we get the matrix of the form $(I_n|B)$. B is the inverse of A.*

Week 9.

Definition 35

- *The system of equations*

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1,n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2,n}x_n &= b_2 \\ &\vdots \\ a_{m,1}x_1 + a_{m,2}x_2 + \cdots + a_{m,n}x_n &= b_m \end{aligned}$$

  where $a_{ij}, b_i \in \mathbb{F}$ for all $i = 1, \ldots, m$ and $j = 1, \ldots, n$ and $x_1, \ldots, x_n$ are variables taking values in $\mathbb{F}$, is called a system of m linear equation in n unknowns over the field $\mathbb{F}$. Note that this system can be written as the matrix product $Ax = b$, where

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1,n} \\ a_{21} & a_{22} & \cdots & a_{2,n} \\ \vdots & \vdots & & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{bmatrix}, \qquad b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}, \qquad \text{and} \qquad x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

- *The matrix A is called the coefficient matrix of the system.*

- *The $m \times (n+1)$ matrix $(A|b)$ is called the augmented matrix of the system $Ax = b$.*

- *A solution to the system is an n-tuple $c = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} \in \mathbb{F}^n$ such that $Ac = b$.*

- *The set of all solutions to the system is called the solution set of the system.*

- *The system is said to be consistent if its solution set is nonempty. Otherwise, the system is said to be inconsistent.*

- *The system is said to be homogeneous if the $b = 0$. Otherwise, the system is said to be inhomogeneous.*

- *The system $Ax = 0$ is said to be the homogeneous system corresponding to the system $Ax = b$.*

- *We often denote the solution set to* $Ax = b$ *by* K *and the solution set to* $Ax = 0$ *by* K_H.

Theorem 3.11. *Let* $A \in \mathsf{M}_{m\times n}(\mathbb{F})$ *and consider the homogeneous system* $Ax = 0$. *The solution set* K_H *to* $Ax = 0$ *is a subspace of* $\mathbb{F}^n$ *and* $\dim K_H = n - \text{rank}(A)$.

Remark 19. *Let* K_H *be the solution set to* $Ax = 0$. *Then*

1. $K_H \neq \emptyset$. *Indeed,* $0 \in K_H$ *is the trivial solution of* $Ax = 0$.
2. $K_H = \{0\}$ *if and only if* $\text{rank}(A) = n$. *in this case, we say the matrix* A *is of full column rank.*
3. *If* $m < n$, *then* $\text{rank}(A) \leq m < n$ *by theorem 3.7, so the system* $Ax = 0$ *has a non-zero solution. In other words a homogeneous system of linear equations with more unknowns than equations has a non-zero solution.*

Theorem 3.12 *Given* $A \in \mathsf{M}_{m\times n}(\mathbb{F})$ *and* $b \in \mathbb{F}^m$, *let*

$$K = \{X \in \mathbb{F}^n | Ax = b\} \qquad \textit{and} \qquad K_H = \{x \in \mathbb{F}^n | Ax = 0\}.$$

Then for any solution c *to* $Ax = b$ *(e.g.* $c \in K$*) we have*

$$K = c + K_H = \{c + k | k \in K_H\}$$

Hence if $Ax = b$ *is consistent, then its solution set* K *is a coset of the solution set of its corresponding homogeneous system* K_H.

Theorem 3.13 *(Invertible Matrix Theorem Part 4). Let* A *be an* $n \times n$ *matrix. Then the following are equivalent:*

1. A *is invertible.*
2. *For some* $b \in \mathbb{F}^m$, *the equation* $Ax = b$ *has a unique solution.*
3. *For all* $b \in \mathbb{F}^m$, *the equation* $Ax = b$ *has a unique solution.*

Theorem 3.14. *Let* $Ax = b$ *be a system of linear equations. Then the system is consistent if and only if* $\text{rank}(A) = \text{rank}(A|b)$.

Definition 36. *Two systems of linear equation are said to be equivalent if they have the same solution set.*

Theorem 3.15. *Let* $Ax = b$ *be a system of* m *linear equations in* n *unknowns and let* C *be a* $m \times m$ *invertible matrix. Then the system* $(CA)x = Cb$ *is equivalent to* $Ax = b$.

Corollary 3.15.1. *Let* $Ax = b$ *be a system of* m *linear equations in* n *unknowns. If* $(A|b) \rightsquigarrow (A'|b')$ *via a sequence of a finite number of elementary row operations, then the system* $A'x = b'$ *is equivalent to the system* $Ax = b$.

Definition 37. *A matrix is said to be in reduced row echelon form (RREF) if the following four conditions are met:*

1. *Non-zero rows (if any) are at the top of the matrix and zero rows (if any) are at the bottom.*

2. *The first non-zero entry in each non-zero row is* 1*, called a leading one.*

3. *The leading one in a non-zero row is the only non-zero entry in its column.*

4. *The leading one in each non-zero row is to the right of the leading one in any row above it.*

Every matrix can be transformed into a matrix in RREF via a finite sequence of elementary row operations. We call such a transformation a row reduction. Gaussian Elimination is an efficient algorithm to row reduce any matrix.

Gaussian Elimination to Row Reduce a Non-zero Matrix into RREF

Suppose we wish to row reduce the matrix $A \in \mathsf{M}_{m\times n}$.

Step 1 *In the leftmost non-zero column, use at most one type-1 and at most one type-2 elementary row operation to get a* 1 *in the first row. (This will be a leading one.)*

Step 2 *By means of at most* $(m-1)$ *type-3 elementary row operations using the first row, create zeroes in all the remaining entries of the leftmost non-zero column, e.g. below the leading one created in step 1.*

Step 3 *Consider the submatrix one column to the right and one row below the leading one we just obtained. Use at most one type-1 and at most one type-2 elementary row operation to get a* 1 *at the top of the first non-zero columns of this submatrix. (This will be a leading one.)*

Step 4 *Use elementary type-3 row operations to obtain zeroes below the* 1 *created in step 3. (Do not create zeroes above the leading one now; we do this later.)*

Step 5 *Repeat steps 3-4 until no non-zero rows remain. This completes the forward phase.*

Step 6 *Now we will create zeroes above the leading ones. Working backwards, beginning with the last non-zero row, use type-3 row operations to create zeroes above the leading one.*

Step 7 *Repeat step 6 with the previous (second-to-last, then third-to-last, etc.) leading one until it has been performed on every non-zero row except the first. This completes the backward phase, and at this point the matrix should be in RREF.*

Theorem 3.16 *Gaussian elimination transforms any matrix into RREF.*

Definition 38*. Let* B *be the RREF of the coefficient matrix* A *of the system of linear equations* $Ax = b$*. If the* j*-th column of* B *does not contain a leading one, then we call* x_j *a free variable.*

Remark 20*. Let* B *be the RREF of the coefficient matrix* A *of the system of linear equations* $Ax = b$*. Then*

1. $\text{rank}(A) = \text{rank}(B) =$ *number of leading ones in* B *=number of non-zero rows in* B.
2. *Number of free variable=* $n-$*number of leading ones=* $n - \text{rank}(A)$

Algorithm for Solving a System of Linear Equations. *To solve the system* $Ax = b$ *for* $A \in \mathsf{M}_{m\times n}(\mathbb{F})$ *and* $b \in \mathbb{F}^m$*, follow these steps:*

Step 1 *Write the augmented matrix of the system* $(A|b)$

Step 2 *Use elementary row operations to row reduce the augmented matrix into RREF* $(A'|b')$. *E.g. use Gaussian elimination.*

Step 3 *Write the system of linear equations corresponding to the RREF.*

Step 4 *If the system contains an equation of the form* $0 = 1$*, then stop as the system is inconsistent.*

Step 5 *Otherwise assign parametric values* $t_1, \dots, t_n - r$ *to the free variables in the system, then solve the remaining variables in terms of the parameters. Here* r *denotes the number of non-zero rows in* A' *or the number of leading zeroes.*

Step 6 *Reorganize the equations from the previous step as a vector equation in the form* $x = x_0 + t_1u_1 + \cdots + t_{n-r}u_{n-r}$.

Step 7 *The solution set to* $Ax = b$ *is the set*

$$K = \{x_0 + t_1u_1 + \cdots + t_{n-r}u_{n0r} | t_1, \dots, t_{n-r} \in \mathbb{F}\} = x_0 + \text{span}\{u_1, \dots, u_{n-r}\}$$

Notice if the solution set to K *is the coset* $x_0 + \text{span}\{u_1, \dots, u_{n-r}\}$*, then the solution set to* K_H *is* $\text{span}\{u_1, \dots, u_{n-r}\}$.

Theorem 3.17 *Let* $(A|b)$ *be the augmented matrix of a consistent system of* m *linear equations in* n *variables. Suppose the RREF of* $(A|b)$ *has* r *non-zero rows. If the general solution to* $Ax = b$ *obtained by the algorithm outlined above is of the form*

$$x = x_0 + t_1u_1 + \cdots + t_{n-r}u_{n-r} \qquad t_1, \dots, t_{n-r} \in \mathbb{F}$$

then $x_0 \in \mathbb{F}^n$ *is a solution to the* $Ax = b$ *and* $\{u_1, \dots, u_{n-r}\}$ *is a basis for the solution set of the corresponding homogeneous system.*

Theorem 3.18 *The RREF of a matrix is unique.*

Week 10.

Definition 39. *Let* $A \in \mathsf{M}_{n\times n}(\mathbb{F})$. *We define the determinant of* A, $\det(A)$ *or* $|A|$*, as follows:*

- *For* $n = 1$, $\det(A) \overset{\text{def}}{=} A_{11}$

- *For* $n \geq 2$, $\det(A) = \sum_{i=1}^{n} (-1)^{i+1} A_{i1} \cdot \det(\widetilde{A}_{i1})$, *where* $\widetilde{A}_{i1}$ *denotes the* $(n-1) \times (n-1)$ *matrix obtained from* A *by deleting row* i *and column* 1.

Remarks.

- *The above definition of determinants is called cofactor expansion along the first column.*
- *More generally, if* A *is an* $n \times n$ *matrix with* $n > 1$, *then* $\widetilde{A}_{ij}$ *denotes the* $(n-1) \times (n-1)$ *matrix obtained from* A *by deleting row* i *and column* j.
- *The scalar* $(-1)^{i+j} \det(\widetilde{A}_{ij})$ *is called the cofactor of entry* A *in row* i, *column* j.
- *The recursive definition of determinants expresses* A *as the sum of entries in the first column multiplied by their corresponding cofactors.*

Theorem 4.1 *Let* $A \in \mathsf{M}_{2\times 2}$. *Then* A *is invertible if and only if* $\det(A) \neq 0$. *Moreover, if* A *is invertibl, then*

$$A^{-1} = \frac{1}{\det(A)} \begin{bmatrix} A_{22} & -A_{12} \\ -A_{21} & A_{11} \end{bmatrix}$$

Example 48 *For all* $n \geq 1$, $\det(I_n) = 1$.

Lemma 10. *If* $A \in \mathsf{M}_{n\times n}(\mathbb{F})$ *is upper-triangular, then* $\det(A)$ *is equal to the product of its entries along the main diagonal. That is,*

$$\text{if} \qquad A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ 0 & a_{22} & a_{23} & \cdots & a_{2n} \\ 0 & 0 & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & a_{nn} \end{bmatrix} \qquad \text{then} \qquad \det(A) = \prod_{i=1}^{n} a_{ii}$$

Lemma 11. *If* $A \in \mathsf{M}_{n\times n}$ *and* A *has a row of zeros, then* $\det(A) = 0$.

Theorem 4.2. *The* det *mapping is "linear in each row." That is, if we fix* $n \geq 1$, $1 \leq i \leq n$ *and* $a_1, \ldots, a_{i-1}, a_{i+1}, \ldots, a_n \in \mathbb{F}^n$, *then for all* $b, c \in \mathbb{F}^n$ *and* $\alpha \in \mathbb{F}$,

$$\begin{vmatrix} — & a_1 & — \\ & \vdots & \\ — & b + \alpha c & — \\ & \vdots & \\ — & a_n & — \end{vmatrix} = \begin{vmatrix} — & a_1 & — \\ & \vdots & \\ — & b & — \\ & \vdots & \\ — & a_n & — \end{vmatrix} + \alpha \begin{vmatrix} — & a_1 & — \\ & \vdots & \\ — & c & — \\ & \vdots & \\ — & a_n & — \end{vmatrix}$$

where $b + \alpha c$, b, *and* c *are inserted in row* i.

Theorem 4.3 *(Determinant after a type 2 elementary row operation). Let* $A \in \mathsf{M}_{n\times n}(\mathbb{F})$ *and* B *be the matrix obtained from* A *by multiplying a row of* A *by a scalar* c. *Then* $\det(B) = c\det(A)$.

Lemma 12. *If a square matrix A has two identical rows, then* $\det(A) = 0$.

Theorem 4.4 *(Determinant after a type 1 elementary row operation). Let $A \in \mathsf{M}_{n\times n}(\mathbb{F})$ and suppose $A \xrightarrow{R_i \leftrightarrow R_j} B$. Then* $\det(B) = -\det(A)$.

Theorem 4.5 *(Determinant after a type 3 elementary row operation). Let $A \in \mathsf{M}_{n\times n}(\mathbb{F})$ and suppose $A \xrightarrow{R_i \leftarrow R_i + cR_j} B$. Then* $\det(B) = \det(A)$.

Corollary 4.5.1. *Let E be an elementary matrix obtained from I_n by an elementary row operation. Then*

1. *If E is obtained by a type 1 row operation* $\det(E) = -1$.
2. *If E is obtained by a type 2 row operation with scalar $c \neq 0$,* $\det(E) = c$.
3. *If E is obtained by a type 3 row operation,* $\det(E) = 1$.

In all cases, $\det(E) \neq 0$.

Corollary 4.5.2. *Let E be an elementary matrix obtained from I_n by an elementary row operation. Then*

1. $\det(E^T) = \det(E)$.
2. $\det(E^{-1}) = \frac{1}{\det(E)}$

Theorem 4.6. *Let E be an $n \times n$ elementary matrix and $A \in \mathsf{M}_{n\times n}(\mathbb{F})$. Then* $\det(EA) = \det(E)\det(A)$.

Corollary 4.6.1. *Let $A \in \mathsf{M}_{n\times n}$ and $E_1, \dots, E_k$ be elementary matrices. Then*

1. $\det(E_1E_2\cdots E_kA) = \det(E_1)\det(E_2)\cdots\det(E_k)\det(A)$.
2. $\det(E_1E_2\cdots E_k) = \det(E_1)\det(E_2)\cdots\det(E_k)$

Theorem 4.7 *(Invertible Matrix Theorem part 5). Let $A \in \mathsf{M}_{n\times n}(\mathbb{F})$. Then A is invertible if and only if* $\det(A) \neq 0$.

Corollary 4.7.1. *Let $A \in \mathsf{M}_{n\times n}(\mathbb{F})$. If* $\operatorname{rank}(A) < n$, *then* $\det(A) = 0$.

Theorem 4.8. *Let $A, B \in \mathsf{M}_{n\times n}(\mathbb{F})$. Then* $\det(AB) = \det(A)\det(B) = \det(BA)$.

Theorem 4.9. *Let $A \in \mathsf{M}_{n\times n}(\mathbb{F})$. Then* $\det(A) = \det(A^T)$.

Corollary 4.9.1. *Suppose $A \in \mathsf{M}_{n\times n}(\mathbb{F})$. If B is obtained from A by swapping to columns, then* $\det(B) = -\det(A)$.

Theorem 4.10. *The determinant of A can be evaluated by cofactor expansion along any column. That, for any fixed $1 \leq i \leq n$, we have*

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+j} A_{ij} \cdot \det(\widetilde{A}_{ij}) = \sum_{i=1}^{n} A_{ij} \cdot \left(\underbrace{(-1)^{i+j}\det(\widetilde{A}_{ij})}_{\text{cofactor } A \text{ at } i,j} \right)$$

Corollary 4.10.1. *The determinant of A can be evaluated by cofactor expansion along any row. That is, for any fixed $1 \le i \le n$, we have*

$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} A_{ij} \cdot \det(\widetilde{A}_{ij}) = \sum_{j=1}^{n} A_{ij} \cdot \left(\underbrace{(-1)^{i+j} \det(\widetilde{A}_{ij})}_{\text{cofactor } A \text{ at } i,j} \right)$$

The following is optional, up until the start of week 11

Lemma 13. *Suppose $A \in \mathsf{M}_{n\times n}(\mathbb{F})$ and let A's columns be $a_1, \dots, a_n$. Fix $1 \le i, j \le n$ and let B be the matrix obtained from A by replacing column i with e_j. That is, $A = \begin{bmatrix} a_1 & \cdots & a_{i-1} & a_i & a_{i+1} & \cdots & a_n \end{bmatrix}$ and $B = \begin{bmatrix} a_1 & \cdots & a_{i-1} & e_j & a_{i+1} & \cdots & a_n \end{bmatrix}$. Then $\det(B) = (-1)^{i+j} \det(\widetilde{A}_{ji})$*

Lemma 14. *Suppose $C \in \mathsf{M}_{n\times n}(\mathbb{F})$, and let its entry at row i column j be C_{ij}. Fix $1 \le i, j \le n$ and let X_{ij} be the $n \times n$ matrix obtained from I_n by replacing column i with $\mathrm{Col}_j(C)$. That is, $X_{ij} = \begin{bmatrix} e_1 & \cdots & e_{i-1} & \mathrm{Col}_j(C) & e_{i+1} & \cdots & e_n \end{bmatrix}$. Then $\det(X_{ij}) = C_{ij}$.*

Theorem 4.11 *(Formula for A^{-1}). Suppose $A \in \mathsf{M}_{n\times n}(\mathbb{F})$ and A is invertible. Then $A^{-1} = \frac{1}{\det(A)} Q$ where Q is the $n \times n$ matrix whose row i, column j entry (i.e. Q_{ij}) is the row j, column i cofactor of A (for all i and j). That is,*

$$Q_{ij} = (-1)^{i+j} \det(\widetilde{A}_{ji})$$

Corollary 4.11.1. *Suppose $A \in \mathsf{M}_{n\times n}(\mathbb{Q})$ and suppose that every entry of A is an integer. If $|\det(A)| = 1$, then (A is invertible and) every entry of A^{-1} is also an integer.*

Theorem 4.12 *(Leibniz Expansion). If $A \in \mathsf{M}_{n\times n}(\mathbb{F})$ then*

$$\det(A) = \sum_{\sigma} \mathrm{sgn}(\sigma) A_{1\sigma(1)} A_{2\sigma(2)} \cdots A_{n\sigma(n)}$$

where the sum is over all n element permutations σ of $\{1, \dots, n\}$, $\sigma(i)$ denotes the ith element of the permutation σ (e.g. if $\sigma = \{2, 1, 3\}$, then $\sigma(2) = 1$), and $\mathrm{sgn}(\sigma) = +1$ if the parity of σ is even, i.e. σ is obtained from the set $\{1, \dots, n\}$ via an even number of swaps, and $\mathrm{sgn}(\sigma) = -1$ if the parity of σ is odd, i.e. σ is obtained from the set $\{1, \dots, n\}$ via an odd number of swaps.

Week 11.

Definition 40. *Let $A \in \mathsf{M}_{n\times n}(\mathbb{F})$.*

- *A nonzero vector $v \in \mathbb{F}^n$ is called an eigenvector of A if there exists a scalar $\lambda \in \mathbb{F}$ such that $Av = \lambda v$. Such a λ is called the eigenvalue of A corresponding to the eigenvector v and (λ, v) is called an eigenpair of the matrix A.*

- *If $\lambda \in \mathbb{F}$ is an eigenvalue of A, the set*

$$\begin{aligned} E_\lambda &= \{\textit{eigenvectors of } A \textit{ corresponding to } \lambda\} \cup \{0\} \\ &= \{v \in \mathbb{F}^n | Av = \lambda v\} \\ &= \{v \in \mathbb{F}^n | (A - \lambda I_n)v = 0\} = \mathcal{N}(A - \lambda I_n) \end{aligned}$$

  is called the eigenspace of A corresponding to λ.

Remark 22. *Let $A \in \mathsf{M}_{n\times n}$ and $\lambda \in \mathbb{F}$ be an eigenvalue of A. From the above definitions,*

1. *A vector $v \in \mathbb{F}^n$ is an eigenvector of A corresponding to the eigenvalue λ if and only if v is a non-zero solution to the linear system $(A - \lambda I_n)v = 0$.*
2. *Since $E_\lambda = \mathcal{N}(A - \lambda I_n)$, E_λ is a subspace of $\mathbb{F}^n$. So $\dim(E_\lambda) \leq n$. Also, since E_λ contains at least one eigenvector, $E_\lambda \neq \{0\}$. Hence, $1 \leq \dim(E_\lambda) \leq n$.*

Theorem 5.1. *Let $A \in \mathsf{M}_{n\times n}(\mathbb{F})$. Then a scalar λ is an eigenvalue of A if and only if* $\det(A - \lambda I_n) = 0$

Definition 41. *Let $A \in \mathsf{M}_{n\times n}(\mathbb{F})$. The polynomial of variable t, $p_A(t) \stackrel{\text{def}}{=} \det(A - tI_n)$ is called the characteristic polynomial of A. That is*

$$p_A(t) = \det(A - tI - n) = \begin{vmatrix} a_{11} - t & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - t & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} - t \end{vmatrix}$$

Theorem 5.2 *(Properties of Characteristic Polynomials). Let $A \in \mathsf{M}_{n\times n}(\mathbb{F})$. Denote* $\operatorname{tr}(A) = \sum_{i=1}^{n} a_{ii}$*, called the trace of A.*

1. *Then*

$$p_A(t) = (-1)^n t^n + (-1)^{n-1}\operatorname{tr}(A)t^{n-1} + c_{n-2}t^{n-2} + \cdots + c_1 t + \det(A)$$

   That is, $p_A(t)$ is a polynomial of degree n with leading coefficient $(-1)^n$ and constant coefficient $\det(A)$. Also, the coefficient of t^{n-1} in $p_A(t)$ is $(-1)^{n-1}\operatorname{tr}(A)$. In addition, the matrix A has at most n distinct eigenvalues.

2. *If $B \in \mathsf{M}_{n\times n}(\mathbb{F})$ is similar to A, then $p_B(t) = p_A(t)$. (Recall: Let $A, B \in \mathsf{M}_{n\times n}(\mathbb{F})$. The matrix B is said to be similar to A if there exists an invertible matrix P such that $B = P^{-1}AP$.)*

Nomenclature. *We call a linear transformation $T : V \to V$ from a vector space V to itself a linear operator.*

Definition 42.

- *Let $T : V \to V$ be a linear operator on a vector space V. A scalar $\lambda \in \mathbb{F}$ if called an eigenvalue of the linear operator T if there exists a non-zero vector $v \in V$ such that $T(v) = \lambda v$. Such a vector v is called an eigenvector of T corresponding to the eigenvalue λ and (λ, v) is called an eigenpair of the linear operator T.*
- *Let $T : V \to V$ be a linear operator on an n-dimensional vector space V with ordered basis β. We defined the characteristic polynomial of T to be the characteristic polynomial of $A = [T]_\beta$*

Theorem 5.3. *Let $T : V \to V$ be a linear operator on a vector space V. Then*

1. *A scalar $\lambda \in \mathbb{F}$ is an eigenvalue of T if and only if $(T - \lambda \mathrm{Id}_V)$ is not invertible.*
2. *Let λ be an eigenvalue of T. A vector $v \in V$ is an eigenvector of T corresponding to λ if and only if $v \neq 0$ and $v \in \mathcal{N}(T - \lambda \mathrm{Id}_V)$.*

Theorem 5.4. *Let V be an n-dimensional vector space with ordered basis β. Then the characteristic polynomial of the linear operator T does not depend on the chosen basis. That is, if α is another ordered basis for V, the characteristic polynomial of T (and so of $[T]_\beta$) is also the characteristic polynomial of $[T]_\alpha$.*

Definition 43.

- *A linear operator T on a finite-dimensional vector space V is said to be diagonalizable if there is an ordered basis β for V such that $[T]_\beta$ is a diagonal matrix.*
- *A square matrix A is said to be diagonalizable if L_A is diagonalizable.*

Theorem 5.5. *Let $T : V \to V$ be a linear operator on an n-dimensional vector space V. Then T is diagonalizable if and only if there is an ordered basis β for V consisting of eigenvectors of T. Moreover, if T is diagonalizable and $\beta = \{v_1, \ldots, v_n\}$ is an ordered basis for V consisting of eigenvectors of T, then the diagonal entries of $[T]_\beta$ are eigenvalues of T corresponding to the eigenvectors $v_1, \ldots, v_n$.*

Theorem 5.6. *Let $A \in \mathsf{M}_{n\times n}(\mathbb{F})$. Then A is diagonalizable if and only if there is an ordered basis β for $\mathbb{F}^n$ consisting of eigenvectors of A. Moreover, if A is diagonalizable and β is an ordered basis for $\mathbb{F}^n$ consisting of eigenvectors of A, then $[L_A]_\beta$ is a diagonal matrix whose diagonal entries are eigenvalues of A corresponding to the vectors in β.*

Theorem 5.7 *(Diagonalizable Matrices). Let $A \in \mathsf{M}_{n\times n}(\mathbb{F})$. Then A is diagonalizable if and only if there exists an invertible matrix P and a diagonal matrix D such that $P^{-1}AP = D$.*

Remark 26.

1. *From the proof of theorem 5.7, if there exists and invertible matrix P and a diagonal matrix D such that $P^{-1}AP = D$, then the columns of P are eigenvectors of A and the diagonal entries of D are the eigenvalues of A corresponding to the columns of P. Note that this factorization is not unique. Even if we sort the entries of D in a given order to ensure D is unique, P is still not.*

2. *If $A \in \mathsf{M}_{n\times n}(\mathbb{F})$ is diagonalizable with invertible matrix P and diagonal matrix D such that $P^{-1}AP = D \iff A = PDP^{-1}$, then by induction we can prove*

$$A^m = PD^mP^{-1}$$

for all $m \geq 1$. Note further that if $D = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$, then $D^m = \operatorname{diag}(\lambda_1^m, \dots, \lambda_n^m)$.

Week 12.

Theorem 5.8. *Let V be a vector space and $T : V \to V$ be linear. Let $v_1, \dots, v_k$ be eigenvectors of T, with eigenvalues $\lambda_1, \dots, \lambda_k$ respectively. Assume that these eigenvalues are distinct, i.e. $\forall i \neq j, \lambda_i \neq \lambda_j$. Then $E_{\lambda_i} \cap E_{\lambda_j} = \{0\}$ for all $1 \leq i \neq j \leq k$, and $\{v_1, \dots, v_k\}$ is linearly independent.*

Corollary 5.8.1

1. *Let $T : V \to V$ be linear on an n-dimensional vector space V. If T has n distinct eigenvalues, then T is diagonalizable.*

2. *Let $A \in \mathsf{M}_{n\times n}(\mathbb{F})$. If A has n distinct eigenvalues, A is diagonalizable.*

Remark 27 *The converse of corollary 5.8.1 is not true, for instance take $A = I_n$ which has only one eigenvalue $\lambda = 1$.*

Theorem 5.9. *Let $T : V \to V$ be linear and let $\lambda_1, \dots, \lambda_k$ be distinct eigenvalues of T. For each $i = 1, \dots, k$, Let S_i be a finite, linearly independent subset of the eigenspace E_{λ_i}. Then*

1. *$S_i \cap S_j = \emptyset, \forall 1 \leq i \neq j \leq k$.*

2. *The set $S = \bigcup_{i=1}^{k} = S_1 \cup S_2 \cup \cdots \cup S_k$ is a linearly independent subset of V and $|S| = \sum_{i=1}^{k} |S_i|$.*

Definition 44. *A polynomial $f(t) \in \mathbb{F}[t]$ splits over $\mathbb{F}$ is there are scalars $c, a_1, \dots, a_n$ (not necessarily distinct) in such that $f(t) = c(t - a_1)(t - a_2)\cdots(t - a_n)$.*

Theorem 5.10. *Let V be a finite-dimensional vector space. The characteristic polynomial of any diagonalizable linear operator $T : V \to V$ splits.*

Remark. *The converse of theorem 5.10 is not true.*

Definition 45. *Let V be a finite dimensional vector space and $T : V \to V$ be linear. Let λ be an eigenvalue of T and $p(t)$ be the characteristic polynomial of T.*

- *The algebraic multiplicity of λ is the largest positive integer k for which $(t - \lambda)^k$ is a factor of $p(t)$.*

- *The geometric multiplicity of λ is the dimension of the eigenspace E_λ.*

Theorem 5.11. *Let $T : V \to V$ be a linear operator on a finite dimensional vector space V, and let λ be an eigenvalue of T having algebraic multiplicity m_λ. Then $1 \leq \dim(E_\lambda) \leq m_\lambda$.*

Theorem 5.12. *Let $T : V \to V$ be a linear operator on a finite dimensional vector space V. Let $\lambda_1, \ldots, \lambda_k$ be all distinct eigenvalues of T and let $m_1, \ldots, m_k$ be their multiplicities. Then T is diagonalizable if and only if*

1. *$p_T(t)$ splits, i.e. $p_T(t) = (-1)^n(t-\lambda_1)^{m_1}(t-\lambda_2)^{m_2}\cdots(t-\lambda_k)^{m_k}$ and*
2. *For each $i = 1, \ldots, k$, $\dim(E_{\lambda_i}) = m_i$.*

Remark 28. *The proof of theorem 5.12 provides a procedure to find the diagonalized factorization of A if it exists.*

Step 1 *Computer the characteristic polynomial $p_A(t)$. If $p_A(t)$ does not split, A is not diagonalizable.*

Step 2 *Find all eigenvalues of A (i.e. roots of $p_A(t)$). Suppose $\lambda_1, \ldots, \lambda_k$ are all distinct eigenvalues of A and m_i is the algebraic multiplicity of λ_i for all $1 \leq i \leq k$.*

Step 3 *Find a basis for each eigenspace E_{λ_i} for all $1 \leq i \leq k$. If there is a $1 \leq j \leq k$ such that $\dim(E_j) \neq m_j$, then A is not diagonalizable. Otherwise, A is diagonalizable. In particular, $\beta = \bigcup_{i=1}^k \beta_i = \beta_1 \cup \cdots \cup \beta_k$ is an ordered basis for V, where for all $1 \leq j \leq k$ β_j is an ordered basis for E_{λ_j}. Hence, if we let P be the square matrix whose columns are vectors from β and let D be the diagonal matrix whose entries are the eigenvalues of A corresponding to the columns of P, then $A = PDP^{-1}$, as desired.*

Definition 46. *Let $A \in \mathsf{M}_{n\times n}(\mathbb{F})$ and $f(t) = a_N t^N + \cdots + a_1 t + a_0 \in \mathbb{F}[t]$. Define*

$$f(A) \overset{\text{def}}{=} a_N A^N + \cdots + a_1 A + a_0 I_n \in \mathsf{M}_{n\times n}(\mathbb{F})$$

Lemma 15. *Let $f, g \in \mathbb{F}[t]$ and $A \in \mathsf{M}_{n\times n}(\mathbb{F})$. Recall $(fg)(t) \overset{\text{def}}{=} f(t)g(t) \in \mathbb{F}[t]$. Then*

- *$(f+g)(A) = f(A) + g(A)$.*
- *$(cf)(A) = cf(A)$ for $c \in \mathbb{F}$.*
- *$(fg)(A) = f(A)g(A)$.*
- *$f(A)g(A) = g(A)f(A)$.*

Lemma 16. *Suppose $A \in \mathsf{M}_{n\times n}(\mathbb{F})$. Then there exists a non-zero polynomial $f \in \mathbb{F}[t]$ such that $f(A) = 0$.*

Definition 47. *A field $\mathbb{F}$ is called algebraically closed if every polynomial in $\mathbb{F}[t]$ of degree at least 1 has a root in $\mathbb{F}$.*

Lemma *(unnamed). Let $A \in \mathsf{M}_{n\times n}(\mathbb{F})$ be an upper-triangular matrix. Then $p_A(A) = 0$ where $p_A(t)$ is the characteristic polynomial of A.*

Theorem 5.13 *(Cayley-Hamilton Theorem). Let $\mathbb{F}$ be algebraically closed. For every $A \in \mathsf{M}_{n\times n}(\mathbb{F})$, we have $p_A(A) = 0$, where $p_A(t)$ is the characteristic polynomial of A.*

Theorem 5.14 *Let $\mathbb{F}$ be algebraically closed. Every $A \in \mathsf{M}_{n\times n}(\mathbb{F})$ is similar to an upper-triangular matrix.*