Math 250- Introductory Linear Algebra
Class notes

Tom Benhamou

Solutions of Linear Systems and Echlon Form

0.1 Systems of Linear Equations

Definition.
- A *Linear Equation* in the variable $x_1, ..., x_n$ is an equation that can be written in the form:

$$a_1x_1 + a_2x_2 + ... + a_nx_n = b$$

  where b, and the *coefficients* $a_1, ..., a_n, b$ are either real or complex numbers. n may be any positive integer.

- A *Linear System* in the variables $x_1, ..., x_n$ is a list of one or more linear equations in the same variables $x_1, ..., x_n$.

- A *solution* to a linear system in the variables $x_1, ..., x_n$ is a list $(s_1, ..., s_n)$ of numbers, such that when the values $s_1, .., s_n$ are substituted for $x_1, ..., x_n$, then equality holds in *all* the equations.

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

- Two linear systems are called *equivalent* if they have the same set of solutions.

Remark. Given any linear equations, we can always add variables or changes the variable names so that the equations are in the same variables.

Fact 1. A linear system of equations either have $0, 1$ of ∞ many solutions.

A system with at least one solution is called *consistent*, and if it has no solutions then it is called *inconsistent*. To determined how many solutions does a linear system has, we will usually ask the two following questions regarding existence and uniqueness:

Question 1. (1) Is the consistent? namely, does at least one solution *exists*? (either 1 or ∞ many solutions)

(2) If a solution exists, is it the only one? namely is the solution unique? (either 0 or 1 solutions)

Definition. An $m \times n$ *matrix* is a rectangle of numbers, with m rows and n columns. (The number of rows always comes first)

Given a linear system

$$a_{1,1}x_1 + \ldots.a_{1,n}x_n = b_1$$

$$a_{2,1}x_1 + \ldots.a_{2,n}x_n = b_2$$

$$\ldots$$

$$a_{m,1}x_1 + \ldots.a_{m,n}x_n = b_m$$

The *coefficient matrix* of the sustem is defined as:

$$\begin{pmatrix} a_{1,1} & \ldots & a_{1,n} \\ a_{2,1} & \ldots & a_{2,n} \\ \vdots & \ddots & \vdots \\ a_{m,1} & \ldots & a_{m,n} \end{pmatrix}$$

The *augmented matrix* of the system is defined as:

$$\begin{pmatrix} a_{1,1} & \ldots & a_{1,n} & b_1 \\ a_{2,1} & \ldots & a_{2,n} & b_2 \\ \vdots & \ddots & \vdots * \vdots & \\ a_{m,1} & \ldots & a_{m,n} & b_m \end{pmatrix}$$

Remark. Note that the rows always correspond to equations while the columns correspond to variables.

Definition. The three basic Row/Equation operations are:

(1) (Replacement) Replace one row by the sum of that of itself and a multiple of another row. ("Add to row i the j row multiplied by α")

(2) (Interchange) Interchange two rows.

(3) (Scaling) Multiply all entries in a row by a nonzero constant.

Two matrices are called *row equivalent* if there is a (finite) sequence of elementary operations which leads from one matrix to the other.

Remark. Row operation are reversible in the sense that every row operation can be canceled by a (possibly) other operation:

(1) Interchange back the same rows.

(2) If we multiplies by the (nonzero) constant c, then we can multiply by $\frac{1}{c}$.

(3) If we added the j^{th} row multiplied by α to the i^{th} row, then we can add the j^{th} row multiplied by $-\alpha$ to the i^{th} row.

Theorem. If two augmented matrices of two linear systems are row equivalent, then the two systems are equivalent, that is, they have the same solution set.

Proof. Checking one by one each operation, we see that if $(s_1, ...s_n)$ is a solution to the first system, then any of the basic operations will not change this fact. For example, if $a_{i,1}s_1 + ...a_{i,n}s_n = b_i$ and $a_{j,1}s_1 + ... + a_{j,n}s_n = b_j$, and we add the j-th row multiplied by α to row i, then $s_1, ..., s_n$ is still a solution since

$$(a_{i,1} + \alpha a_{j,1})s_1 + ...(a_{i,n} + \alpha a_{j,n})s_n = b_i + \alpha b_j.$$

Since every operation is reversible, then the same argument shows that if $(s_1, ..., s_n)$ is a solution to the second system, it is a solution to the first.

□

0.2 Row Reduction and Echelon Form

Definition. A leading entry in a row (equation) is the first nonzero entry (the first variable that appears).

Definition (echalon form)**.** A rectangular matrix is in *echalon form* if it has the following three properties:

(1) All zero rows are at the bottom.

(2) Each leading entry in a nonzero row is in a column strictly to the right to the column of the leading entry in the row above it. In particular, all the entries in a column below a leading entry are 0.

It is moreover called *reduced echalon form* if:

(3) all the leading entries are 1.

(4) each leading 1 is the only nonzero entry in its column.

Theorem. Each matrix is row equivalent to one and only one reduced echalon matrix

Remark. Teh echalon form is convenient to deduce how many solution are there and the reduced echalon form is good to represent solutions

Solutions of linear system

(1) A pivot column is a column in the echlon form where there is a leading entry.

(2) Each variable corresponding to a pivot column is called a *basic variable*.

(3) The other variables are called *free variables*.

That means that choosing any value for the free variables determines uniquely the basic variables and yield a solution.

This gives a description of the solution set in parametric representation.

Theorem. A linear system is consistent if and only if the right-most columns of any echlon form of the augmented matrix is not a pivot row. Namely, there is not a "contradictory line".

If a linear system is consistent, then there is a unique solution if and only if there are no free variables.

1 Vectore in $\mathbb{R}^n$

1.1 The palne $\mathbb{R}^2$

A verctor in $\mathbb{R}^2$ is a pair $\bar{u} = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$ where u_1, u_2 are real numbers. $\mathbb{R}^2 = \{\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \mid u_1, u_2 \in \mathbb{R}\}$

Remark. $\mathbb{R}^2$ can be identified with the plane.

Remark. We add the bar in '$\bar{u}$' to remind ourselves that the variable represents a vector (rather than a number)

Vector Equality: $\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} c \\ d \end{pmatrix}$ iff $a = c$ and $b = d$.

Vector addition: $\begin{pmatrix} a \\ b \end{pmatrix} + \begin{pmatrix} c \\ d \end{pmatrix} = \begin{pmatrix} a + c \\ b + d \end{pmatrix}$

Remark. Vector addition corresponds geometrically to the parallelogram rule.

Multiplication by scalar: A *scalar* is just a number in $\mathbb{R}$. For a scalar α we define: $\alpha \cdot \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \alpha \cdot a \\ \alpha \cdot b \end{pmatrix}$.

Remark. scalar multiplication corresponds geometrically to:

(1) If $\alpha > 1$- stratch.

(2) If $0 < \alpha < 1$- shrink.

(3) If $\alpha < 0$- reverse direction.

We identify $\begin{pmatrix} a \\ b \end{pmatrix}$ with a 2×1-matrix and $\begin{pmatrix} a & b \end{pmatrix}$ with 1×2 matrix.

1.2 Vectors in $\mathbb{R}^n$

Similarily we can idemtify $\mathbb{R}^3$ with the space.

Definition. n-tuples are $n \times 1$-matrices $\bar{u} = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix}$ $\mathbb{R}^n = \{\begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} \mid u_1, \ldots, u_n \in \mathbb{R}\}$=the set of **all** possible n-tuples.

Remark. For $n = 3$ $\mathbb{R}^3$ can be idetified eith the space. For $n > 3$ the is no (clear) geometry associated.

Definition. $\bar{0} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$

Addition: $\begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} + \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} u_1 + v_1 \\ u_2 \\ \vdots \\ u_n + v_n \end{pmatrix}$.

Multiplicationby scalars: $\alpha \cdot \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} = \begin{pmatrix} \alpha u_1 \\ \alpha u_2 \\ \vdots \\ \alpha u_n \end{pmatrix}$.

Claim. *For all $\bar{u}, \bar{v}, \bar{w} \in \mathbb{R}^n$ and all α, β scalars we have*

(i) $\bar{v} + \bar{u} = \bar{u} + \bar{v}$.

(ii) $\bar{v} + (\bar{u} + \bar{w}) = (\bar{v} + \bar{u}) + \bar{w}$.

(iii) $\bar{v} + \bar{0} = \bar{v}$.

(iv) $\bar{u} + (-\bar{u}) = \bar{0}$.

(v) $\alpha \cdot (\bar{u} + \bar{v}) = \alpha \cdot \bar{u} + \alpha \cdot \bar{v}$.

(vi) $(\alpha + \beta) \cdot \bar{u} = \alpha \cdot \bar{u} + \beta \cdot \bar{u}$.

(vii) $\alpha \cdot (\beta \cdot \bar{u}) = (\alpha\beta) \cdot \bar{u}$.

(viii) $1 \cdot \bar{u} = \bar{u}$.

(ix) $0 \cdot \bar{u} = \bar{0}$.

Definition. A *linear combination* of $\bar{v}_1, \ldots, \bar{v}_n$ is a vector $\bar{y} \in \mathbb{R}^n$ that can be written as

$$y = c_1\bar{v}_1 + \ldots + c_n\bar{v}_n = \sum_{i=1}^{n} c_i\bar{v}_i.$$

Vector equation A vector equation is an equation of the form

$$x_1\bar{v}_1 + \ldots x_n\bar{v}_n = \bar{b}.$$

Theorem. The vector equation $x_1\bar{v}_1 + \ldots x_n\bar{v}_n = \bar{b}$. has exactly the same solutions as the augmented matrix

$$\begin{pmatrix} | & & | & | \\ \bar{v}_1 & \ldots & \bar{v}_n & \bar{b} \\ | & & | & | \end{pmatrix}$$

Definition. Given vectors $\bar{v}_1, \ldots, \bar{v}_n \in \mathbb{R}^n$, define Span$\{\bar{v}_1, \ldots, \bar{v}_n\}$ as the set of **all** linear combinations of $\bar{v}_1, \ldots, \bar{v}_n$.

Remark. $\bar{b} \in$ Span$\{\bar{v}_1, \ldots, \bar{v}_n\}$ iff $x_1\bar{v}_1 + \ldots x_n\bar{v}_n = \bar{b}$ has a solution iff

$$\begin{pmatrix} | & & | & | \\ \bar{v}_1 & \ldots & \bar{v}_n & \bar{b} \\ | & & | & | \end{pmatrix}$$

has a solution.

2 The Matrix Equation $Ax = b$.

Compactifying our notations even more, we write the vector equation

$$x_1\bar{v}_1 + \ldots x_n\bar{v}_n = \bar{b}$$

in matrix form as

$$\begin{pmatrix} | & & | \\ \bar{v}_1 & \ldots & \bar{v}_n \\ | & & | \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \bar{b}$$

Definition. An equation of the form $A \cdot \bar{x}$ denoted the linear combinations of the columns of A with weights $\bar{x}$.

Theorem. Let A be an $M \times n$-matrix with columns $\bar{a}_1, ..., \bar{a}_n$ and $\bar{b} \in \mathbb{R}^m$. The following have the same solutions:

(1) The matrix equation: $A\bar{x} = \bar{b}$.

(2) the vector equation $x_1\bar{a}_1 + ... + x_n\bar{a}_n = \bar{b}$.

(3) The linear system whose augmented matrix is:

$$\begin{pmatrix} | & & | & | \\ \bar{a}_1 & \dots & \bar{a}_n & \bar{b} \\ | & & | & | \end{pmatrix}$$

2.1 Existence of solutions

Corollary. The following are eequivalent:

(1) $\bar{b} \in \text{Span}(\bar{a}_1, .., \bar{a}_n)$.

(2) $A\bar{x} = \bar{b}$ has a solution.

Corollary. The following are equivalent

(1) The columns of A span $\mathbb{R}^m$.

(2) for every b, $Ax = b$ is consistent.

(3) In the Echlon form of A, every row has a leading entry.

Corollary. If $n < m$, then no n-vectors in $\mathbb{R}^m$ can span $\mathbb{R}^m$.

Matric-vector product:

Definition. $(A \cdot \bar{x})_i = (A)_{i,1}x_1 + ...(A)_{i,n}x_n$.

Theorem. (1) $A(\bar{u} + \bar{v}) = A\bar{u} + A\bar{v}$.

(2) $A(\alpha\bar{u}) = \alpha(A\bar{u})$.

(3) $A(\sum_{i=1}^{n} \alpha_i\bar{v}_i) = \sum_{i=1}^{n} \alpha_i \cdot (A\bar{v}_i)$.

2.2 Homogeneous Linear Systems

Definition. A linear system is said to be homogeneous if it can be written in the form $A\bar{x} = \bar{0}$. In other words if it is equivalent to the form:

$$a_{1,1}x_1 + ...a_{1,n}x_n = 0$$

$$a_{2,1}x_1 + ...a_{2,n}x_n = 0$$

$$...$$

$$a_{m,1}x_1 + ...a_{m,n}x_n = 0$$

Proposition. (1) An homogeneous system always has a solution.

(2) The set of solutions of a linear system can always be written as a span.

If the set of solutions of $Ax = 0$ is Span$(\bar{v}_1, ..., \bar{v}_k)$, then the *general solution* or *parametric vector form of the solution* is

$$\bar{x} = t_1\bar{v}_1 + ... + t_k\bar{v}_k$$

2.3 Solution to non-homogeneous systems

Definition. If $Ax = b$ is a non homogeneous linear equation, the corresponding homogeneous system is $Ax = 0$.

Any specific solution to $Ax = b$ is called a private solution.

Theorem. Suppose that $AX = b$ is a consistent linear system. Then the general solution for $Ax = b$, has the form:

$$\bar{x} = \bar{p} + \bar{v}_h$$

where p is **any** fixed private solution for $Ax = b$ and $\bar{v}_h$ is the general solution for the corresponding homogeneous system $Ax = 0$.

Corollary. If $Ax = b$ is consistent then $Ax = b$ has the same number of solutions as $Ax = 0$.

3 Linear Independence

Definition. A sequence of vectors $\{\bar{v}_1, ...\bar{v}_n\} \subseteq \mathbb{R}^m$ is called *linearly independent*(LI) if the vector equation

$$(\star)\ \ x_1\bar{v}_1 + ...x_n\bar{v}_n = 0$$

has only the trivial solution. It is called linearly dependent (LD) otherwise, that is, if there are coeficients $\alpha_1, .., \alpha_n$, not all 0, such that

$$(*)\alpha_1\bar{v}_1 + ... + \alpha_n\bar{v}_n = 0$$

Any solution to $(\star)$ (including the trivial one) is called a linear dependence of $\{\bar{v}_1, ...\bar{v}_n\}$.

Remark. Note that $\bar{0}$ cannot be a part of a linearly independent sequence since we can always put a non-zero coefficient to it to produce a non-zero linear dependence.

Claim. *A single non-zero vector is linearly independence.*

Corollary. TFAE:

(1) the sequence $\{\bar{v}_1, ...\bar{v}_n\}$ is linearly independent.

(2) $x_1\bar{v}_1 + ...x_n\bar{v}_n = 0$ has a unique solution.

(3) In (any) reduced form of A, evey column has a leading entry, where

$$\begin{pmatrix} | & & | \\ \bar{v}_1 & \dots & \bar{v}_n \\ | & & | \end{pmatrix}$$

Corollary. more than m vectors in $\mathbb{R}^m$ are LD.

This is not very useful in abstract settings (but very useful for specific examples). To prove that a sequence of vector $\{\bar{v}_1, ...\bar{v}_n\}$, you can use the following: "Suppose that $\alpha_1\bar{v}_1 + ... + \alpha_n\bar{v}_n = 0$, let us prove that $\alpha_1 = \alpha_2 = ... - = \alpha_n = 0$.

Example 1. Show that if $\{\bar{v}_1, ...\bar{v}_n\}$ is linearly independent, then also

$$\{\bar{v}_1 + \bar{v}_1, ...\bar{v}_n + \bar{v}_1\}$$

is linearly independent.

Proof. Suppose that

$$(*)\ \ \alpha_1(\bar{v}_1 + \bar{v}_1) + ... + \alpha_n(\bar{v}_n + \bar{v}_1) = 0$$

we want to prove that $\alpha_1 = ... = \alpha_n = 0$. rearranging $(*)$ we have

$$(2\alpha_1 + \alpha_2 + ... + \alpha_n)\bar{v}_1 + \alpha_2\bar{v}_2 + ... + \alpha_n\bar{v}_n = 0$$

By our assumption, $\bar{v}_1, ...\bar{v}_n$ are linearly independent, and therefore

$$2\alpha_1 + \alpha_2 + ... + \alpha_n = 0$$

$$\alpha_2 = \alpha_3 = ... = \alpha_n = 0$$

Thenrefore $2\alpha_1 = 0$ and thus $\alpha_1 = 0$. □

Proposition. Let $\bar{v}_1, \bar{v}_2$ be any two vectors, then $\{\bar{v}_1, \bar{v}_2\}$ is linearly independent iff nither of them is a scalar multiplicity of the other. For example, in $\mathbb{R}_2$ and $\mathbb{R}^3$, this means that they do not lay on the same line.

Proof. If $\bar{v}_1 = \alpha \cdot \bar{v}_2$ for example, then $(1, -\alpha)$ is a non-zero linear dependence, sop the vectors are LD. If (α, β) is a non-zero lineare dependence, then $\alpha\bar{v}_1 + \beta\bar{v}_2 = 0$, suppose for example that $\alpha \neq 0$, then $\bar{v}_1 = \frac{\beta}{\alpha}\bar{v}_2$. □

In general we have the following (which is not very convenient in practice but good for the intuition):

Theorem. TFAE:

(1) A sequence $\{\bar{v}_1, ..., \bar{v}_n\}$ is LD.

(2) $\bar{v}_1 = \bar{0}$ or there is $1 < j \leq n$ such that $\bar{v}_j \in \text{Span}(\bar{v}_1, ..., \bar{v}_{j-1})$ (i.e. $\bar{v}_j$ is a linear combination of the previous ones.).

(3) There is $1 \leq j \leq n$ such that $\bar{v}_j \in \text{Span}(\bar{v}_1, ..., \bar{v}_{j-1}, \bar{v}_{j+1}, ..., \bar{v}_n)$.

Problem 1. Prove that $\bar{v}_1, ..., \bar{v}_n$ is LI if and only if every $\bar{b} \in \text{Span}(\bar{v}_1, ..., \bar{v}_n)$ can be represented uniquely as a linear combination of $\bar{v}_1, ..., \bar{v}_n$.

Definition. A sequenc of vector $\bar{v}_1, ..., \bar{v}_n$ is called a base for $\mathbb{R}^m$ if it is spanning and LI.

Corollary. $\bar{v}_1, ..., \bar{v}_n$ is a base for $\mathbb{R}^m$ iff every $\vec{b} \in \mathbb{R}^m$ can be written uniquely as a linear combination of $\bar{v}_1, ..., \bar{v}_n$.

Corollary. Any base of $\mathbb{R}^m$ consisted of exactly m vectors.

4 Linear Transformations

Definition. A transformation/function/mapping from $\mathbb{R}^n$ to $\mathbb{R}^m$ is an assigment f, such that:

1. f assigns to **every** $a \in A$, an element $b \in B$.
2. the element $b \in B$ which is assigned to $a \in A$, is uniquely determined by a and is denoted by $f(a)$.

We denote it we $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$. $\mathbb{R}^n$ is the *domain f the function* f which is denoted by $\text{dom}(f)$ and $\mathbb{R}^m$ is the *co-domain of the function* f which we denote by $\text{codom}(f)$.

Further notations: $f(a)$ is called the *image* if a under f and if $b = f(a)$ then a is call *a preimage of* b. The *image/range of* f is the set of all elements in B which are images of elements of A under f. Namely $\text{ran}(f) = \{f(a) \mid a \in A\}$.

Example 2. (1) Given an $m \times n$ matrix A, we define the *matrix tranformation*, $T_A : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ defined by $T_A(\bar{x}) = A \cdot \bar{x}$. For example if

$$B = \begin{pmatrix} 1 & 2 & 0 \\ 1 & 1 & -1 \end{pmatrix}$$

Then $T_B : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$.

$$T_B(\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}) = \begin{pmatrix} 1 & 2 & 0 \\ 1 & 1 & -1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

$\text{dom}(T_B) = \mathbb{R}^3$, $\text{codom}(f) = \mathbb{R}^2$. The vector $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ is the image of $\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$ under T_B. At this point it is not clear what is the range of T_B, but we will see later that it is $\mathbb{R}^2$.

Example 1 p.68

Recall that the identity matrix is defined by $I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$. Then T_{I_3} is what we call the identity map.

Definition. A transformation $f : \mathbb{R}^n \to \mathbb{R}^m$ is called *linear* if for every $\bar{v}, \bar{u} \in \mathbb{R}^n$ and $\alpha \in \mathbb{R}$:

(1) $f(\bar{v} + \bar{u}) = f(\bar{v}) + f(\bar{u})$.

(2) $f(\alpha \cdot \bar{v}) = \alpha \cdot f(\bar{v})$

Example 3. [Exmaple 4 p.71] Let $r \in \mathbb{R}$ be a scalar. Define $f_r : \mathbb{R}^n \to \mathbb{R}^n$ be the function $f_r(\bar{x}) = r \cdot \bar{x}$.

Example 5

Corollary. If T is a linear transformation then $T(\bar{0}) = \bar{0}$

Proof. $T(\bar{0}) = T(0 \cdot \bar{v}) = 0T(\bar{v}) = \bar{0}$. □

Example 4. The above corollary is a simple way to rule out certain function from being linear. For example $f(x) = x + 1$ is not linear since $f(0) \neq 0$.

Corollary. If $T : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation, then T moves a linear combination to a linear combination. Formally, for every $\bar{v}_1, ..., \bar{v}_k \in]\mathbb{R}^n$, and any $\alpha_1, ..., \alpha_k \in \mathbb{R}$ scalars, $f(\sum_{i=1}^{k} \alpha_i \bar{v}_i) = \sum_{i=1}^{k} \alpha_i f(\bar{v}_i)$; that is:

$$f(\alpha_1 \bar{v}_1 + ...\alpha_k \bar{v}_k) = \alpha_1 f(\bar{v}_1) + \alpha_2 f(\bar{v}_2) + ... + \alpha_k f(\bar{v}_k)$$

Proof. Do it for 2, and the general case is clear. □

4.1 The matrix of a linear transformation

We denote the unit vectors by $\bar{e}_i = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \\ 0 \end{pmatrix} \leftarrow i^{\text{th}}$ place. Note that the identity matrix $I_n = \begin{pmatrix} \bar{e}_1 & \dots & \bar{e}_n \end{pmatrix}$. Hence for every vector $\bar{v} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}$ we have $\bar{v} = I_n \bar{v} = v_1 \bar{e}_1 + \dots + v_n \bar{v}_n$ (by the definition of multiplication $A\bar{v}$ as the linear combinations of the columns).

Theorem. Every matrix transformation T_A is linear. Moreover $T_A(\bar{e}_i) = C_i(A)$, where $C_i(A)$ is the i^{th}-column of A.

Proof. For linearity, use the properties we proved regarding he product $A \cdot \bar{x}$. To see that $T_A(\bar{e}_i) = C_i(A)$, recall that by the definition of multiplication $A \cdot \bar{x}$ as a linear combination of the columns we have:

$$T_A(\bar{e}_i) = A \cdot \bar{e}_i = 0 \cdot C_1(A) + 0 \cdot C_2(A) + \dots + 1 \cdot C_i(A) + \dots + 0 \cdot C_n(A) = C_i(A)$$

□

Theorem. Every linear map $T : \mathbb{R}^n \to \mathbb{R}^m$ is a matrix transformation. Namely, there is a unique $m \times n$ matrix A such that $T = T_A$ (hence $C_i(A) = T(\bar{e}_i)$)

Proof. Let

$$A = \begin{pmatrix} | & \dots & | \\ T(\bar{e}_1) & \dots & T(\bar{e}_n) \\ | & \dots & | \end{pmatrix}$$

Note that since $\text{dom}(T) = \mathbb{R}^n$, there are n unit vectors $\bar{e}_1, .., \bar{e}_n$ to plug in T. Since codom(T) = $\mathbb{R}^m$, the matrix A is an $m \times n$ matrix. To see that $T = T_A$

we will show that for every vector $\bar{v}$, $T(\bar{v}) = A \cdot \bar{v} = T_A(\bar{v})$. To see this, note that

$$\bar{v} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = v_1 \cdot \bar{e}_1 + v_2\bar{e}_2 + \ldots + v_n\bar{e}_n$$

For example

$$\begin{pmatrix} 2 \\ 4 \\ -1 \end{pmatrix} = 2 \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + 4 \cdot \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + (-1) \cdot \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

Since T is linear it preserves linear combinations and therefore:

$$T(\bar{v}) = T(v_1 \cdot \bar{e}_1 + v_2\bar{e}_2 + \ldots + v_n\bar{e}_n) = v_1 \cdot T(\bar{e}_1) + v_2T(\bar{e}_2) + \ldots + v_nT(\bar{e}_n) =$$

$$= \begin{pmatrix} | & \ldots & | \\ T(\bar{e}_1) & \ldots & T(\bar{e}_n) \\ | & \ldots & | \end{pmatrix} \cdot \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = A \cdot \bar{v}$$

□

The matrix A is called *the standard matrix for the linear transformation T*.

Example 5. Define $T : \mathbb{R}^2 \to \mathbb{R}^3$ by $T(\binom{x}{y}) = \begin{pmatrix} x + 2y \\ x - 2y \\ 2y + 3x \end{pmatrix}$. Show that T is linear and find the standard matrix for the linear transformation T.

Example 6. Define $T_{\frac{\pi}{2}} : \mathbb{R}^2 \to \mathbb{R}^2$, the rotation matrix by $\frac{\pi}{2}$-radian (90°).

It is not hard to see geometrically that $T_{\frac{\pi}{2}}$ is linear. Hence by the above theorem it is suppose to be a matrix transformation T_A. To find the matrix A we need to compute $T_{\frac{\pi}{2}}(\binom{1}{0}) = \binom{0}{1}$ and $T_{\frac{\pi}{2}}(\binom{0}{1}) = \binom{-1}{0}$. Hence $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. The general formula for rotation matrix by θ-radians is:

$$\begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & cos(\theta) \end{pmatrix}$$

4.2 surjective and injective transformations

Definition. A transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ is called onto/surjective if any $\bar{b} \in \mathbb{R}^m$ is the image of some $\bar{a} \in \mathbb{R}^n$ under f. Namely, $f(\bar{a}) = \bar{b}$. Equivalently, if $\text{ran}(f) = \mathbb{R}^m$.

Problem 2. Formulate what does it mean that T is not onto.

Problem 3. Prove that in general $\text{ran}(T) = \text{Span}(T(\bar{e}_1), ..., T(\bar{e}_n))$.

Definition. A mapping $T : \mathbb{R}^n \to \mathbb{R}^m$ is called one-to-one/injective if for each $\bar{b} \in \mathbb{R}^m$ there is at most one $\bar{a} \in \mathbb{R}^n$ such that $T(\bar{a}) = \bar{b}$. In other words, if $\bar{a} \neq \bar{a}'$ then $T(\bar{a}) \neq T(\bar{a}')$.

Example 7 (Exmple 4 p. 81)**.**

Theorem. Let $F : \mathbb{R}^n \to \mathbb{R}^m$, and let A be the standard matrix of T. TFAE:

(1) T is onto.

(2) For every $\bar{b} \in \mathbb{R}^m$, the equation $A\bar{x} = \bar{b}$ has a solution.

(3) The columns of A span $\mathbb{R}^m$.

Proof. (2) and (3) are equivalent from previous theorems regarding spanning sequences (Theorem 4 p.39 of the textbook).

To that (1) is equivalent to (2), given any $\bar{b} \in \mathbb{R}^m$, since $T_A = T$, we have that, for any $\bar{a} \in \mathbb{R}^n$, $T(\bar{a}) = T_A(\bar{a}) = A \cdot \bar{a}$, hence $\bar{a}$ is a solution to $A\bar{x} = \bar{b}$ iff $T(\bar{a}) = \bar{b}$. That means that for every $\bar{b}$ $A\bar{x} = \bar{b}$ has a solution iff for every $\bar{b}$ there is $\bar{a}$ such that $T(\bar{a}) = \bar{b}$, namely, T is onto.

For (2) implies (1), □

Theorem. T is one-to-one the only $\bar{x} \in \mathbb{R}^n$ such that $T(\bar{x}) = \bar{0}$ is $\bar{x} = \bar{0}$ (i.e. $T(\bar{x}) = \bar{0}$ has only the trivial solution).

Proof. If T is one-to-one, and supposed that $T(\bar{x}) = \bar{0}$. Then $\bar{x} = \bar{0}$ since if $\bar{x} \neq \bar{0}$, then $T(\bar{x}) = \bar{0} = T(\bar{0})$, which is in contradiction to T being one-to-one. In the other direction, suppose that the only solution to $T(\bar{x}) = \bar{0}$ is $\bar{x} = \bar{0}$, and let us prove that T is one-to-one. Suppose that $T(\bar{v}) = T(\bar{u})$, and let us prove that $\bar{v} = \bar{u}$. Indeed, $T(\bar{v} - \bar{u}) = T(\bar{v}) - T(\bar{u}) = \bar{0}$. Hence $\bar{v} - \bar{u}$ is a solution to $T(\bar{x}) = \bar{0}$. Since the only solution to $T(\bar{x}) = \bar{0}$, it follows that $\bar{v} - \bar{u} = \bar{0}$ and therefore $\bar{v} = \bar{u}$. □

Corollary. Let $F : \mathbb{R}^n \to \mathbb{R}^m$, and let A be the standard matrix of T. TFAE:

(1) T is one-to-one.

(2) the equation $A\bar{x} = \bar{0}$ has a unique solution.

(3) the columns of A are linearly independent.

Proof. (2), (3) are equivalent as we have already seen in previous theorems (in the textbook this is green remark on p.61)

To see that (1), (2) are equivalent, by the previous theorem T is one-to-one iff $T(\bar{x}) = \bar{0}$ has a unique solution. Since $T(\bar{x}) = A \cdot \bar{x}$, this is the same equation as $A\bar{x} = \bar{0}$. Hence (1) and (2) are equivalent.

□

Example 8 (Example 5 p.82)**.**

5 Matrix Altgebra

Let A be the $m \times n$ matrix

$$A = \begin{bmatrix} a_{1,1} & \dots & a_{1,n} \\ a_{2,1} & \dots & a_{2,n} \\ \vdots & \ddots & \vdots \\ a_{m,1} & \dots & a_{m,n} \end{bmatrix}$$

To access the (i, j)-cell in A we denote $A_{i,j} = a_{i,j}$.

Definition. Let A, B be two matrices of the same $m \times n$ dimension and $r \in \mathbb{R}$ be a scalar. Define:

(1) $A + B$ is a matrix of dimension $m \times n$ defined by $(A + B)_{i,j} = A_{i,j} + B_{i,j}$.

(2) rA is a matrix of dimension $m \times n$ defined by $(rA)_{i,j} = r(A_{i,j})$

Theorem. (1) $A + B = B + A$.

(2) $(A + B) + C = A + (B + C)$.

(3) $A + 0 = A$, where 0 is the zero matrix defined by $0_{i,j} = 0$.

(4) $r(A + B) = rA + rB$.

(5) $(r + s)A = rA + sA$.

(6) $r(sA) = (rs)A$.

5.1 Matrix multiplication

Definition. Let $f : \mathbb{R}^n \to \mathbb{R}^m$ and $g : \mathbb{R}^m \to \mathbb{R}^k$ be function. Define the composition of g on f, denoted by $g \circ f : \mathbb{R}^n \to \mathbb{R}^k$ by

$$g \circ f(\bar{x}) = g(f(\bar{x}))$$

Claim. *If $T : \mathbb{R}^n \to \mathbb{R}^m$ and $S : \mathbb{R}^m \to \mathbb{R}^k$ are linear transformations then $S \circ T$ is a linear transformation.*

If $T = T_A$ and $S = T_B$, then

$$T \circ S(\bar{x}) = T(S(\bar{x})) = B(Ax) = B(x_1 C_1(A) + ... + x_n C_n(A)) =$$

$$= x_1 BC_1(A) + ... + x_n BC_n(A) = \begin{bmatrix} | & | & \cdots & | \\ BC_1(A) & BC_2(A) & \ldots & BC_n(A) \\ | & | & \cdots & | \end{bmatrix} \cdot \bar{x}$$

Definition. Let B be $k \times m$ and A be $m \times n$ matrices, define

$$B \cdot A = \begin{bmatrix} | & | & \cdots & | \\ BC_1(A) & BC_2(A) & \ldots & BC_n(A) \\ | & | & \cdots & | \end{bmatrix}$$

Remark. Matrix multiplication is only defined when the number of columns of the left metrix is the same as the number of rows of the right matrix.

Theorem. If $T : \mathbb{R}^n \to \mathbb{R}^m$ and $S : \mathbb{R}^m \to \mathbb{R}^k$ are linear transformations, $T = T_A$ and $S = T_B$, then $T \circ S = T_{B \cdot A}$.

Corollary. The columns of $B \cdot A$ are linear combinations of the columns of B.

Fast Computation: $(B \cdot A)_{i,j} = b_{i,1} a_{1,j} + ... + b_{i,m} a_{m,j}$.

Corollary. The rows of $B \cdot A$ are linear combinations of the rows of A.

Theorem. Whenever the products below are defined, we have the following:

(1) $(A \cdot B) \cdot C = A \cdot (B \cdot C)$.

(2) $A(B + C) = AB + AC$.

(3) $(B+C)A = BA + CA$.

(4) $r(AB) = (rA)B = A(rB)$.

(5) $I_m A = A = AI_n$ (given that A is $m \times n$)

Remark. (1) In general $AB \neq BA$.

(2) No cancellation law: $AB = AC \nRightarrow B = C$.

(3) $AB = 0 \nRightarrow A = 0$ or $B = 0$.

(4) $A \cdot \bar{x}$ is a spectial case of matrix multiplication when we think of $\bar{x}$ as a $n \times 1$ matrix.

Definition. Let A be a square matrix. Define $A^m = \underbrace{A \cdot A .... \cdot A}_{m\text{-times}}$. Also let $A^0 = I_n$ and $A^1 = A$.

Definition. Let A be an $m \times n$ matrix. The transpose of A is an $n \times m$ matrix defined by $(A^T)_{i,j} = A_{j,i}$.

Theorem. (1) $(A^T)^T = A$.

(2) $(A+B)^T = A^T + B^T$.

(3) $(rA)^T = rA^T$.

(4) $(AB)^T = B^T A^T$.

5.2 Inverse of Matrix

Definition. A square matrix A of dimension $n \times n$ is invertible if there is an $n \times n$ matrix C such that

$$AC = CA = I_n$$

Claim. *If A is invertible then C is unique. We denote $C = A^{-1}$*

Remark. We are only allowed to write A^{-1} after we proved (somehopw) that A is invertible.

non-invertible matrices are sometimes called singular matrices.

Theorem. Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. If $ad - bc \neq 0$, then A is invertible and $A^{-1} = \frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$. If $ad - bc = 0$, then A is singular.

Theorem. If A is invertible then for every $b \in \mathbb{R}^n$, $A\bar{x} = \bar{b}$ has a unique solution given by $\bar{x} = A^{-1}\bar{b}$.

Theorem. (1) If A is invertible then A^{-1} is invertible and $(A^{-1})^{-1} = A$.

(2) If A is invertible then A^T is invertible and $(A^T)^{-1} = (A^{-1})^T$.

(3) If A, B are invertible then AB is invertible and $(AB)^{-1} = B^{-1}A^{-1}$.

Definition. An elementary matrix is a matrix obtained by performing a basic row operation on I_n.

Theorem. Let E be an elementary matrix. EA is the matrix obtained by performing the basic operation of E on A.

Corollary. Each elementary matrix is invertible and the inverse matrix is the elementary matrix of the inverse operation.

Theorem. If A is reducible to I_n then A is invertible. Moreover, any sequence $E_1 .... E_n$ of basic operations that reduces A to I also reduces I to A^{-1}

Remark. We can view $[A|I]$ and trying to solve n-many equations simoultanuiously $Ax = e_i$. Then the columns of A^{-1} would exactly be those solutions.

5.3 Characterization of invertible matrix

Theorem. TFAE for any $n \times n$ matrix A:

(1) A is invertible.

(2) A is reducible to I.

(3) in any Eachelon form of A every column has a leading entry.

(4) $AX = 0$ has a unique solution.

(5) the columns of A are linearly independent.

(6) T_A is one-to-one.

(7) there is an $n \times n$ matrix C such that $CA = I$ (left invertible)

(8) in any Eachelon form of A there are no zero lines.

(9) for every b, $AX = b$ has at least one solution.

(10) the columns of A span $\mathbb{R}^n$.

(11) the map T_A is onto.

(12) there is an $n \times n$ matrix C such that $AC = I$ (right invertible)

(13) A^T is invertible.

Corollary. If $AB = I$ e then A, B have to be both invertible with $A = B^{-1}$ and $B = A^{-1}$.

Corollary. If AB is invertible then A, B are both invertible.

Proof. $ABC = I$ and $A(BC) = I$ and therefore A is invertible. $CAB = I$ and therefore B is invertible. □

5.4 invertible linear transformations

$T : \mathbb{R}^n \to \mathbb{R}^n$ is called invertible if there is $D : \mathbb{R}^n \to \mathbb{R}^n$ such that $T(S(x)) = x$ and $S(T(x)) = x$. In which case we denote $S = T^{-1}$.

Theorem. T is invertible iff T is one to one and onto.

Theorem. $T = T_A$ is invertible iff A is invertible in which case $T^{-1} = T_{A^{-1}}$.

Corollary. For a linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$, T is invertible iff T is one-to-one or onto.

6 Determinants

The following is a recursive definition:

Definition (/Theorem)**.** Let A be an $n \times n$ square matrix, then

$$\det(A) = \sum_{k=1}^{n} a_{ik} \det(A_{ik}) = \sum_{k=1}^{n} a_{kj} \det(A_{kj})$$

Where a_{ij} is the element in the i^{th}-row and the j^{th}-column of A, and A_{ij} is the $(n-1) \times (n-1)$-matrix obtained from A by erasing the i^{th}-row and the j^{th}-column.

If $A = [a_{11}]$ is a 1×1 matrix then $\det(A) = a_{11}$. If $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$, then $\det(A) = a_{11}a_{22} - a_{21}a_{12}$.

Theorem. If $A = \begin{bmatrix} a_{11} & * & \dots & * \\ 0 & a_{22} & \dots & * \\ \ddots & \vdots & \dots & * \\ 0 & 0 & \dots & a_{nn} \end{bmatrix}$ is a triangular matrix then $\det(A) = a_{11} \cdot a_{22} \cdot a_{nn}$.

Theorem. Let A be a square matrix.

(a) If B is the outcome of multiplying a row of A by $\alpha \neq 0$ then $\det(A) = \frac{1}{\alpha} \det(B)$.

(b) If B is the outcome of interchanging rows of A then $\det(A) = -\det(B)$.

(c) If B is the outcome of adding a row of A multiplied by a scalar c to another row then $\det(A) = \det(B)$.

Theorem. A is invertible iff $\det(A) \neq 0$

Theorem. $\det(A) = \det(A^T)$

Theorem. $\det(A \cdot B) = \det(A) \det(B)$

Theorem (Cramer's rule)**.** Let A be an invertable matrix then for every $\bar{b} \in \mathbb{R}^n$, $A\bar{x} = \bar{b}$ has a unique solution given by

$$x_i = \frac{\det(A_i(\bar{b}))}{\det(A)}$$

where $A_i(\bar{b})$ is the matrix obtained from A by replacing the i^{th} column by $\bar{b}$.

Definition. The adjugate matrix of A, is defined by

$$(\operatorname{adj}(A))_{ij} = (-1)^{i+1} \det(A_{ji})$$

Theorem. If A is invertible then $A^{-1} = \frac{1}{\det(A)} \cdot \operatorname{adj}(A)$

Appendix: Complex numbers

A *complex number* is a number of the form $a + bi$, where $a, b \in \mathbb{R}$ are real numbers, and i is an new number which satisfies $\sqrt{-1} = i$, that is, $i^2 = -1$. a is called the *real* part and b the *imaginary part*.

Example 9. $2 + 2i$, $i - 1$, $5i$, π are all complex numbers. The real part of $i - 1$ is -1 and the imaginary part if 1.

Complex numbers equality:

$$a + bi = c + di \text{ if and only if } a = c,\ b = d$$

We denote by $\mathbb{C} = \{a + bi \mid a, b \in \mathbb{R}\}$ the set of all complex numbers. So $\mathbb{C}$ can be identified with pairs (a, b) and therefore can be thought of as a plane- called the Complex/Gauss Plane.

(include graphics of Gauss plane)

We add and subtract complex numbers as follows:

$$(a + ib) + (c + id) = (a + c) + i(b + d)$$
$$(a + ib) - (c + id) = (a - c) + i(b - d)$$

We multiply them as follows:

$$(a + ib)(c + id) = (ac - bd) + i(ad + bc)$$

To divide them we need something called the complex conjugate:

Definition. Given a complex number $z = a + ib$, the complex conjugate of z, denoted by $\bar{z}$ is the complex number $\bar{z} = a - ib$.

Claim. $z \cdot \bar{z} = a^2 + b^2 \in \mathbb{R}$ *is a non-negative real number. Moreover, if* $z \neq 0$, *then* $a^2 + b^2 > 0$.

Proof. For the first part

$$z \cdot \bar{z} = (a + ib)(a - ib) = a^2 - (-b^2) + i(ab - ba) = a^2 + b^2$$

For the second part, if $z \neq 0$, then either $a \neq 0$ and then $a^2 > 0$ or $b \neq 0$ in which case $b^2 > 0$. In any case $a^2 + b^2 > 0$. □

So in order to divide $\frac{z}{w}$, where $w \neq 0$ we can do the following:

$$\frac{z}{w} = \frac{z\bar{w}}{w\bar{w}}$$

Example 10.

Together with those operations, $\mathbb{C}$ is what we call a field, which similar to $\mathbb{R}$, gives meaning to equations with complex variables.