II. LINEAR TRANSFORMATIONS

RODICA D. COSTIN

 CONTENTS

2. Linear Transformations 1
2.1. Definition and examples 1
2.2. The matrix of a linear transformation 2
2.3. Operations with linear transformations and with their associated matrices 5
2.4. Null space and range 5
2.5. Column space and rank of a matrix 6
2.6. Rank, minors, the dimension of the span of a finite set 7
2.7. Systems of linear equations, column space and null space 8
2.8. Dimensionality properties 8
2.9. Invertible transformations, isomorphisms 10
2.10. Change of basis, similar matrices 12
2.11. Examples. 15
2.12. Summary of some properties of \( n \) dimensional vector spaces 17
2.13. Left inverse, right inverse of functions and of matrices 17

2. Linear Transformations

2.1. Definition and examples. Let \( U, V \) be two vector spaces over the same field \( F \) (which for us is \( \mathbb{R} \) or \( \mathbb{C} \)).

Definition 1. A linear transformation \( T : U \to V \) is a function which is linear, in the sense that it satisfies

\[
(1) \quad T(x + y) = T(x) + T(y) \quad \text{for all } x, y \in U
\]

and

\[
(2) \quad T(cx) = cT(x) \quad \text{for all } x \in U, \ c \in F
\]

The properties (1), (2) which express linearity, are often written in one formula as

\[
T(cx + dy) = cT(x) + dT(y) \quad \text{for all } x, y \in U, \ c, d \in F
\]
Note that if \( T \) is linear, then also
\[
T \left( \sum_{k=1}^{r} c_k x_k \right) = \sum_{k=1}^{r} c_k T(x_k) \quad \text{for all } x_1, \ldots, x_r \in U, \; c_1, \ldots, c_r \in F
\]

**Notation:** for linear transformations it is customary to denote simply \( T x \) rather than \( T(x) \).

Note that a linear transformation takes the zero vector to the zero vector: \( T \mathbf{0} = \mathbf{0} \). (Indeed, take \( c = 0 \) in (2).)

**Definition 2.** A linear transformation \( \phi : V \to F \), from a vector space to its scalar field, is called a linear functional.

By contrast, linear transformations \( T : U \to V \) with target space \( V \) not the scalar field are also called linear maps, or linear operators.

**Examples.**
1. The transformation \( T : \mathbb{R}^n \to \mathbb{R}^n \) given by \( T(x) = 2x \) is linear.
   More generally, scaling transformation is linear: for some fixed scalar \( 2 \in \mathbb{R} \), let \( T x = \lambda x \). In particular, \( T x = 0 \) for all \( x \) is linear, and so is the identity transformation, \( I x = x \).
   And even more generally, dilations: \( T : \mathbb{R}^n \to \mathbb{R}^n \), \( T(x_1, x_2, \ldots, x_n) = (\lambda_1 x_1, \lambda_2 x_2, \ldots, \lambda_n x_n) \) are linear.
   The rotation of the plane by \( \theta \) (rad) counterclockwise: \( R : \mathbb{C} \to \mathbb{C} \), \( R(z) = e^{i \theta} z \) is linear (over \( F = \mathbb{C} \)).
2. The projection operator \( P : \mathbb{R}^3 \to \mathbb{R}^2 \) defined by \( P(x_1, x_2, x_3) = (x_1, x_2) \) is a linear operator.
   The projection \( P : \mathbb{R}^3 \to \mathbb{R} \), \( P(x_1, x_2, x_3) = x_2 \) is a linear functional.
3. The translation transformation \( T : \mathbb{R}^n \to \mathbb{R}^n \), given by \( T(x) = x + e_1 \) is not linear.
   The usually called ”linear function” \( f(x) = ax + b \) is not a linear transformation. It is more correctly called an affine transformation. An affine transformation is a linear transformation followed by a translation.
4. A few examples in infinite dimensions. Denote by \( P \) the vector space of polynomials with coefficients in \( F \).
   o) Evaluation at some \( t_0 \): the functional \( E_{t_0} : P \to F \) defined by \( E_{t_0}(p) = p(t_0) \) is linear.
   a) The differentiation operator: \( D : P \to P \), defined by \( (Dp)(t) = p'(t) \) is linear.
   b) The integration operator: \( J : P \to P \), defined by \( (Jp)(t) = \int_0^t p(s)ds \) is linear, and so is the functional \( I : P \to F \), by \( Ip = \int_0^1 p(s)ds \). Note that \( I \) is the composition \( I = E_1 \circ J \).

2.2. The matrix of a linear transformation.
   I. A linear transformation is completely determined by its action on a basis.
Indeed, let \( T : U \to V \) be a linear transformation between two finite dimensional vector spaces \( U \) and \( V \). Let \( B_U = \{ u_1, \ldots, u_n \} \) be a basis of \( U \). Then any \( x \in U \) can be uniquely written as

\[
x = x_1 u_1 + \ldots + x_n u_n \quad \text{for some} \quad x_1, \ldots, x_n \in F
\]

and by the linearity of \( T \) we have

\[
T x = T(x_1 u_1 + \ldots + x_n u_n) = x_1 T u_1 + \ldots + x_n T u_n
\]

hence, once we give the values \( T u_1, \ldots, T u_n \) then the action of \( T \) is determined on all the vectors in \( U \).

II. To determine \( T u_1, \ldots, T u_n \) we need to specify the representation of these vectors in a given basis of \( V \). Let then \( B_V = \{ v_1, \ldots, v_m \} \) be a basis of \( V \), and denote by \( M_{ij} \) the coefficients of these representations:

\[
T u_j = \sum_{i=1}^{m} M_{ij} v_i, \quad \text{for some} \quad M_{ij} \in F, \quad \text{for all} \quad j = 1, \ldots, n
\]

The matrix \( M = M_T = [M_{ij}]_{i=1 \ldots m, j=1 \ldots n} \) is called the **matrix representation of the linear transformation \( T \) in the bases \( B_U \) and \( B_V \)**.

**Note** that the matrix representation of a linear transformation depends not only on the basis chosen for \( U \) and \( V \), but also on the order on which the vectors in each basis are enumerated. It would have been more precise to write the two basis \( B_U \), \( B_V \) as a multiplet, rather than a set.

**Note.** Let \( M_T \) is the matrix representation of the linear transformation \( T \) in the bases \( B_U, B_V \). Any \( x \in U \) can be represented as (3), and denote \( T x = y = y_1 v_1 + \ldots + y_m v_m \). Then, organizing the coordinates of \( x, y \) as columns, the action of \( T \) is matrix multiplication on the coordinate vector:

\[
M_T \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix}
\]

which can be seen after a direct verification: \( T x = \sum_j x_j T u_j = \sum_j x_j \sum_i M_{ij} v_i = \sum_i (\sum_j M_{ij} x_j) v_i = \sum_i y_i v_i \).

Note also the block representation of the matrix of \( T \): if we organize the coordinates \( M_{1j}, \ldots, M_{mj} \) of \( T u_j \) as columns,

\[
M_T = [T u_1 \mid T u_2 \mid \ldots \mid T u_n]
\]

From now on we have to write the coordinates of vectors as columns.

Conversely, **every matrix determines a linear transformation**. Let \( M \) be an \( m \times n \) matrix, and consider the transformation which multiplies vectors \( x \in \mathbb{R}^n \) by the matrix \( M \)

\[
T : \mathbb{R}^n \to \mathbb{R}^m, \quad T(x) = M x
\]
$T$ is a linear transformation (check!), whose matrix in the standard bases is exactly $M$. Indeed,

$$Tx = Mx = M \left( \sum_{j=1}^{n} x_j e_j \right) = \sum_{j=1}^{n} x_j M e_j$$

where $M e_j$ is the column $j$ of $M$ (convince yourselves!), so

(5)  

$$M e_j = \sum_{i=1}^{m} M_{ij} e_i$$

Note that (5) can be used to give a block representation of $M$ where each block is a column:

$$M = [M e_1 \mid M e_2 \ldots \mid M e_n]$$

Therefore any linear transformation is a matrix multiplication.

**Example 1.** Consider the standard basis $\{e_1, e_2\}$ for $\mathbb{C}^2$, where

(6)  

$$e_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad e_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

For any $z \in \mathbb{C}^2$ we have

$$z = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = z_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + z_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} = z_1 e_1 + z_2 e_2$$

$\mathbb{C}$ is a vector space over $F = \mathbb{C}$, of dimension 1. Denote its standard basis (consisting of $\{f_1\}$ for $\mathbb{C}$, where $f_1 = [1]$).

Define a linear transformation $T : \mathbb{C}^2 \to \mathbb{C}$ (note that here the scalars are $F = \mathbb{C}$, and $n = 2$, $m = 1$) be the linear transformation defined by its action on the basis of its domain: let $T(e_1) = f_1$, $T(e_2) = -f_1$. By linearity, its action on any element of $\mathbb{C}^2$ is completely determined:

$$T(z) = T(z_1 e_1 + z_2 e_2) = z_1 T(e_1) + z_2 T(e_2) = z_1 [1] + z_2 [-1] = (z_1 - z_2)[1]$$

The same calculation in matrix notation: $[M_{ij}]_{i=1 \ldots j=2} = [1, -1]$ is the matrix of $T$, and $T(z) = Mz$ (with $z$ written as a vertical string).

**Example 2.** Consider the standard basis $p_j(t) = t^j$, $j = 0, 1, \ldots, n$ of $\mathcal{P}_n$, and let us find the matrix representation of the differentiation operator $D : \mathcal{P}_n \to \mathcal{P}_n$, $Dp = p'$.

Noting that $Dp_k = kp_{k-1}$, the matrix of $D$ has the block form

$$M = M_D = [Dp_0 \mid Dp_1 \mid Dp_2 \ldots \mid Dp_n] = [0 \mid 1 \mid 2p_1 \ldots \mid np_{n-1}]$$
2.3. Operations with linear transformations and with their associated matrices. The sum of two linear transformations \( S, T : U \to V \) is defined like for any functions, as \( (S + T)(u) = Su + Tu \) and so is multiplication by scalars, as \( (cT)(u) = cTu \). It turns out that \( S + T \) and \( cT \) are also linear transformations, and so is the composition:

**Theorem 3.** Let \( U, V, W \) be vector spaces over \( F \), and fix some bases \( B_U, B_V, B_W \). Let \( T, S : U \to V \) be linear transformations, with matrix representations \( M_T, M_S \) in the bases \( B_U, B_V \).

(i) Any linear combination \( cS + dT : U \to V \) is a linear transformation, and its matrix representation is \( cM_S + dM_T \) (in the bases \( B_U, B_V \)).

(ii) Let \( R : V \to W \) be linear, with matrix \( M_R \) in the basis \( B_V, B_W \). Then the composition \( R \circ S : U \to W \), defined as (usually) by \( (R \circ S)(x) = R(S(x)) \) is a linear transformation, with matrix \( M_R M_S \) (in the bases \( B_U, B_W \)).

The proof of the Theorem relies on immediate calculations; the calculations needed in part (ii) will be detailed below in §2.10.1 (and it shows why we multiply matrices using that strange rule...). □

For linear transformations we simply denote \( R \circ S \equiv RS \).

It is easy to check that the operations with linear transformations satisfy the axioms of vector spaces, and the set of all linear transformations from \( U \) to \( V \), denoted \( \mathcal{L}(U, V) \) is a linear space. Moreover, for \( U = V \), there is an extra operation, the composition of linear transformations, which behaves very nicely with respect to addition and scalar multiplication, in the sense that usual algebra rules apply: \( (RS)T = R(ST) \), \( R(cS + dT) = cRS + dRT \), \( cRS + dTS = (cR + dT)S \) except that ”multiplication” is not commutative, since, in general, \( RS \) does not equal \( SR \) (recall that, in general, for two functions \( f, g \), \( f \circ g \neq g \circ f \)).

2.4. Null space and range. The following definitions and properties are valid in finite or infinite dimensions.

Let \( U, V \) be vector spaces over the same scalar field \( F \).

**Definition 4.** Let \( T : U \to V \) be a linear transformation.

The **null space** (or **kernel**) of \( T \) is

\[
\mathcal{N}(T) = \{ x \in U : Tx = 0 \}
\]

The **range** of \( T \) is

\[
\mathcal{R}(T) = \{ y \in V : y = Tx \text{ for some } x \in U \}
\]
Example. Let $T: \mathbb{R}^3 \to \mathbb{R}$ be the linear transformation defined by $T(x_1, x_2, x_3) = x_2 - 3x_1$ (this is a functional, to be more precise). Then $\mathcal{N}(T) = \{(x_1, x_2, x_3) \in \mathbb{R}^3 \mid x_2 - 3x_1 = 0\}$ (a plane) and $\mathcal{R}(T) = \mathbb{R}$.

**Theorem 5.** Let $T: U \to V$ be a linear transformation. Then:

(i) $\mathcal{N}(T)$ is a subspace of $U$.

(ii) $\mathcal{R}(T)$ is a subspace of $V$.

(iii) $T$ is one to one if and only if $\mathcal{N}(T) = \{0\}$.

Why: the proof of (i)-(ii) is immediate.

To show (iii), first assume that $T$ is one-to-one. If $x \in \mathcal{N}(T)$ then $Tx = 0 = T0$ hence $x = 0$ and therefore the only element of $\mathcal{N}(T)$ is $0$. Conversely, assume that $\mathcal{N}(T) = \{0\}$. If $x, y \in U$ are so that $Tx = Ty$ then $T(x - y) = 0$ (because $T$ is linear) which means that $x - y \in \mathcal{N}(T)$, hence $x - y = 0$ which implies that $T$ is one-to-one. \square

One of the key features of linear transformations is that if a (linear) property is valid at $0$ then it is valid at any point. Theorem 5 (iii) illustrates this principle: if $T$ takes the value $0$ only once, then, if $T$ takes some value, then it takes it only once.

Some buzz-words: the **rank** of $T$ is
\[ \text{rank}(T) \equiv \dim \mathcal{R}(T) \]
and the **nullity** of $T$ is
\[ \text{nullity}(T) \equiv \dim \mathcal{N}(T) \]

2.5. **Column space and rank of a matrix.** Recall that matrices are convenient ways to represent linear transformations (once bases are chosen). It is then useful to transcribe the notions of null space, nullity, range, and rank for matrices.

Let $M$ be an $n \times m$ matrix with entries in $F$, and $T$ its associated linear transformation $T: F^n \to F^m$, $T_Mx = Mx$. Recall that $M$ is the matrix of $T_M$ associated to the standard bases.

Since vectors in $\mathcal{R}(T_M)$ have the form $Mx = M(\sum_j x_je_j) = \sum_j x_jMe_j$ we see that
\[ \mathcal{R}(T_M) = \text{Sp}(Me_1, \ldots, Me_n) \]
which is the space spanned by the columns of $M$, called the **column space**, and which we will denote (by abuse of notation) by $\mathcal{R}(M)$.

The $\text{rank}(T_M)$, is, by definition, the dimension of its range. It is natural to define the rank of a matrix as
\[ \text{rank}(M) = \text{rank}(T_M) = \dim \text{of the column space of } M \]

**Proposition 6.** If $\dim \text{Sp}(x_1, \ldots, x_k) = r > 0$ then there are $r$ vectors independent among $x_1, \ldots, x_k$.

Here is an orderly procedure to choose them. Since $r \geq 1$, there is a nonzero vector. Let $x_j$ be the first nonzero vector in the list $x_1, \ldots, x_k$. \[ \text{Example.} \]
If $r = 1$ then $\{x_j\}$ is a basis, and we are done.

For $r \geq 2$: if $x_j, x_{j+1}$ are dependent, then we ignore $x_{j+1}$ (in this case $x_{j+1} \in Sp(x_1, \ldots, x_j)$) and ask if $x_j, x_{j+2}$ are dependent, etc. Eventually we must have for some $q$ that $x_j, x_q$ are independent, while $Sp(x_1, \ldots, x_j) = Sp(x_1, \ldots, x_j, \ldots x_{q-1})$. If $r = 2$ we are done, $\{x_j, x_q\}$ is a basis.

If $r \geq 3$: if $x_j, x_q, x_{q+1}$ are dependent, then we disregard $x_{q+1}$ (which belongs to $Sp(x_1, \ldots, x_j, \ldots x_q)$), and check if $x_j, x_q, x_{q+2}$ are dependent etc. Eventually there is a vector $x_p$ so that $x_j, x_q, x_p$ are independent.

The procedure continues until we find $r$ independent vectors. □.

Let $r = \text{rank } M$. Then among the columns $Me_j$ of $M$ there are $r$ linearly independent ones, by Proposition 6; this is called a basis column and is a basis for $R(T_M)$.

2.6. Rank, minors, the dimension of the span of a finite set. Let $M = [M_{ij}]_{i=1,\ldots,n; j=1,\ldots,k}$ be an $n \times k$ matrix. This could be the matrix of a linear transformation $T$, and then the rank of $M$ is $\dim R(T_M)$. Or, the matrix $M$ may be constructed so that its columns represent the coordinates of $k$ vectors $x_1, \ldots, x_k$ in a chosen basis $v_1, \ldots, v_n$: $x_j = \sum_i M_{ij} v_i$; in this case the rank of $M$ is $\dim Sp(x_1, \ldots, x_k)$. In this section we find a practical way to determine the rank of a matrix.

Recall the following properties of determinants:

1) The value of a determinant equals the value of the determinant obtained by turning its lines into columns ($\det M = \det M^T$).

2) A determinant is zero if and only if some column is a linear combination of the others (its columns are linearly dependent).

3) A determinant is zero if and only if some row is a linear combination of the other rows (its rows are linearly dependent).

**Definition.** Let $M$ be an $n \times k$ matrix, and $p \in \mathbb{Z}_+$, $p \leq \min\{k,n\}$. Delete any $n - p$ rows and $k - p$ columns of $M$; we are left with a $p \times p$ sub-matrix. Its determinant is called a minor of order $p$.

Of course, there are many minors of any order.

In the following we assume that $M$ is not the zero matrix.

**Remark 3.** Denote $r = \text{rank } M$. Then any collection of $p \geq r + 1$ columns of $M$ are linearly dependent, therefore any minor of order $\geq r + 1$ is zero.

Then by Proposition 6, the rank of a matrix is the largest number $r$ for which there is a nonzero minor of order $r$ of $M$.

**Examples.**

(i) Determine the dimension of $Sp(x_1, x_2, x_3)$ where $x_1 = (2, 3, 1, 4)^T$, $x_2 = (0, -2, 1, 3)^T$, $x_3 = (2, 5, 0, 1)^T$.

**Solution.** We form the matrix having as columns the coordinates of these
vectors:

\[ M = [x_1 \mid x_2 \mid x_3] = \begin{bmatrix} 2 & 0 & 2 \\ 3 & -2 & 5 \\ 1 & 1 & 0 \\ 4 & 3 & 1 \end{bmatrix} \]

The rank of \( M \) can be at most 3. The matrix has 4 minors of order 3 (obtained by deleting one of the four rows); a direct calculation shows that they are all zero, hence the rank of \( M \) is less than 3. Now we look at the minors of order 2, and we see that the upper left 2 by 2 submatrix has nonzero determinant, hence the rank of \( M \) is 2, and so is \( \text{dim} \, Sp(x_1, x_2, x_3) \).

(ii) Consider the linear transformation \( T : \mathbb{R}^3 \to \mathbb{R}^4 \) given by \( Tx = Mx \) where \( M \) is the matrix in (8). Find the dimension of the set of all vector \( y \) which can be written as \( y = Tx \) for some \( x \).

**Solution.** Of course, this is \( \text{dim} \, \mathbb{R}(T) = \text{rank}(T) = \text{rank}(M) \) which we found to be 2.

2.7. Systems of linear equations, column space and null space.

**Note:**

(0) A homogeneous system \( Mx = 0 \) has the set of solutions equal to \( N(M) \).

(1) The system \( Mx = b \) is soluble (it has solutions) if and only if \( b \in \mathbb{R}(M) \).

(2) Suppose the system \( Mx = b \) is soluble. Then its solution is unique if and only if \( N(M) = \{0\} \).

(3) If \( x_0 \) is a particular solution of \( Mx = b \), then the set of all solutions is

\[ x_0 + N(M) = \{x_0 + z \mid z \in N(M)\} \]

In other words:

**General solution=Particular solution + Homogeneous solutions**

Indeed, on one hand it is clear that all the vectors in \( x_0 + N(M) \) solve the same equation: \( M(x_0 + z) = Mx_0 = b \). Conversely, any two solutions \( x_1, x_2 \) of the same equation: \( Mx_1 = b, Mx_2 = b \), differ by an element \( x_1 - x_2 \in N(M) \).

(4) A soluble system \( Mx = b \) has a unique solution if and only if the homogeneous equation \( Mx = 0 \) has a unique solution (which is 0).

This is clear by (3) and property (iv) of Theorem 5.

2.8. Dimensionality properties.

**Theorem 7.** Let \( T : U \to V \) be a linear transformation between two finite dimensional vector spaces \( U \) and \( V \). Let \( B_U = \{u_1, \ldots, u_n\} \) be a basis of \( U \).

Then:

(i) \( \mathcal{R}(T) = Sp(Tu_1, \ldots, Tu_n) \).

(ii) \( T \) is one-to-one if and only if \( Tu_1, \ldots, Tu_n \) are linearly independent.

(iii) We have

\[ \text{dim} \, U = \text{dim} \, \mathcal{R}(T) + \text{dim} \, N(T) \]

(in other words, \( \text{rank}(T) + \text{nullity}(T) = n \)).
Remarks.
1. Property (i) has already been established for standard bases, see (7).
2. Property (iii) can be interpreted intuitively as: when the operator $T$ acts on $n$ dimensions, some directions collapse to zero ($\dim N(T)$ of them), and the rest span the range (their number is $\dim R(T)$).

2. In the particular case of $T = T_M$ where $T(x) = Mx$ as in §2.5:
   - Property (i) has been established by (7).
   - Property (ii) states that:

**Corollary 8.** A solvable system $Mx = b$ has a unique solution if and only if the columns of $M$ are independent.

If $M$ is an $m \times n$ matrix, its columns are independent if and only if $n \leq m$ (the number of columns does not exceed the number of rows) and the rank $M = n$ (maximal rank).

**Proof of Theorem 7.**

Note that (i) is immediately visible from (3), (4).

For property (ii), note that $P \sum c_j T u_j = 0$ is equivalent to $\sum c_j u_j \in N(T)$. Now, $T$ is one-to-one if and only if $N(T) = \{0\}$ (by Theorem 5 (iii)). If $N(T) = \{0\}$ then $\sum c_j u_j = 0$ which implies all $c_j = 0$ hence $T u_1, \ldots, T u_n$ are independent. Conversely, assume $T u_1, \ldots, T u_n$ are independent and let $x \in N(T)$. Then $x = \sum c_j u_j$ for some $c_j$'s. Then $0 - T x = \sum c_j T u_j$ hence all $c_j = 0$ and therefore $x = 0$, showing that $N(T) = \{0\}$.

Property (iii) is proved as follows.

Let $v_1, \ldots, v_r$ be a basis for $R(T)$. Therefore, $v_j = T(x_j)$ for some $x_j \in U$, for all $j = 1, \ldots, r$.

Now let $z_1, \ldots, z_k$ be a basis for $N(T)$. Property (iii) is proved if we show that $B = \{z_1, \ldots, z_k, x_1, \ldots, x_r\}$ is a basis for $U$.

To show linear independence, consider a linear combination

$$c_1 z_1 + \ldots + c_k z_k + d_1 x_1 + \ldots + d_r x_r = 0 \quad \text{(9)}$$

to which we apply $T$ on both sides, and it gives

$$T(c_1 z_1 + \ldots + c_k z_k + d_1 x_1 + \ldots + d_r x_r) = T 0$$

and by linearity

$$c_1 T z_1 + \ldots + c_k T z_k + d_1 T x_1 + \ldots + d_r T x_r = 0$$

and since $T z_j = 0$, and $v_j = T x_j$

$$d_1 v_1 + \ldots + d_r v_r = 0$$

which imply all $d_j = 0$ since $v_1, \ldots, v_r$ are linearly independent. Then (9) becomes

$$c_1 z_1 + \ldots + c_k z_k = 0$$

which implies all $c_j = 0$ since $z_1, \ldots, z_k$ are linearly independent. In conclusion $B$ is a linearly independent set.
To complete the argument we need to show that \( Sp(B) = U \). For any \( u \in U \) we have

\[
Tu = d_1v_1 + \ldots + d_rv_r
\]

for some scalars \( d_1, \ldots, d_r \). Note that

\[
T(u - d_1x_1 - \ldots - d_rx_r) = Tu - d_1v_1 - \ldots - d_rv_r = 0
\]

therefore \( u - d_1x_1 - \ldots - d_rx_r \in \mathcal{N}(T) \), so \( u - d_1x_1 - \ldots - d_rx_r = c_1z_1 + \ldots + c_kz_k \) which shows that \( u \in Sp(B) \). □

2.8.1. More consequences of the Dimensionality Theorem 7 (iii). Let \( U, V \) be finite dimensional vector spaces, and \( T : U \to V \) be a linear transformation.

1. Suppose \( \dim U = \dim V \). Then \( T \) is one to one \( \iff \) \( T \) is onto.

For systems this means that for a given square matrix \( M \) we have: the homogeneous equation \( Mx = 0 \) has only the zero solution if and only if all \( Mx = b \) are solvable.

Reformulated as: \textit{either all} \( Mx = b \) \textit{are solvable, or} \( Mx = 0 \) \textit{has nonzero solutions, it is the celebrated Fredholm's alternative.}

To prove 1., note that: \( T \) is one to one \( \iff \dim \mathcal{N}(T) = 0 \) (and by Theorem 7 (iii)) \( \iff \dim \mathcal{R}(T) = \dim V \) which means \( \mathcal{R}(T) = V \).

2. If \( \dim U > \dim V \) then \( T \) is not one to one.

Indeed, \( \dim \mathcal{N}(T) = \dim U - \dim \mathcal{R}(T) \geq \dim U - \dim V > 0 \).

3. If \( \dim U < \dim V \) then \( T \) is not onto.

Indeed, \( \dim \mathcal{R}(T) = \dim U - \dim \mathcal{N}(T) \leq \dim U < \dim V \).

2.9. Invertible transformations, isomorphisms.

2.9.1. The inverse function. Recall that if \( f : A \to B \) is a function which is onto to one and onto, then the function \( f \) has an inverse, denoted \( f^{-1} \) defined as \( f^{-1} : B \to A \) by \( f(x) = y \iff x = f^{-1}(y) \).

Recall that the compositions \( f \circ f^{-1} \) and \( f^{-1} \circ f \) equal the identity functions:

\[
(f^{-1} \circ f)(x) = f^{-1}(f(x)) = x, \quad (f \circ f^{-1})(y) = f(f^{-1}(y)) = y
\]

Examples:
1) let \( f : \mathbb{R} \to \mathbb{R}, f(x) = 5x \). Then \( f^{-1} : \mathbb{R} \to \mathbb{R}, f^{-1}(x) = x/5 \);
2) let \( g : \mathbb{R} \to \mathbb{R}, f(x) = x + 3 \). Then \( f^{-1} : \mathbb{R} \to \mathbb{R}, f^{-1}(x) = x - 3 \);
3) let \( h : (-\infty, 0] \to [0, \infty) \) by \( h(x) = x^2 \) then \( h^{-1} : [0, \infty) \to (-\infty, 0], h^{-1}(x) = -\sqrt{x} \).
2.9.2. The inverse of a linear transformation. Let \( T : U \rightarrow V \) be a linear transformation between two vector spaces \( U \) and \( V \). If \( T \) is onto to one and onto, then the function \( T \) has an inverse \( T^{-1} : V \rightarrow U \).

**Exercise.** Show that the inverse of a linear transformation is also a linear transformation.

**Definition 9.** A linear transformation \( T : U \rightarrow V \) which is onto to one and onto is called an **isomorphism of vector spaces**, and \( U \) and \( V \) are called **isomorphic vector spaces**.

Whenever two vector spaces are isomorphic\(^1\) then any property which can be written using the vector space operations that one space has, can be translated, using the isomorphism \( T \), into a similar property for the other space.

The following theorem shows that all the finite dimensional vector spaces are essentially \( \mathbb{R}^n \) or \( \mathbb{C}^n \):

**Theorem 10.**

*Any vector space over \( \mathbb{R} \) of dimension \( n \) is isomorphic to \( \mathbb{R}^n \).*

*Any vector space over \( \mathbb{C} \) of dimension \( n \) is isomorphic to \( \mathbb{C}^n \).*

Indeed, let \( U \) be a vector space of dimension \( n \) over \( F \) (where \( F \) is \( \mathbb{R} \) or \( \mathbb{C} \)). Let \( u_1, \ldots, u_n \) be a basis for \( U \). Define \( T : U \rightarrow F^n \) by \( T(u_j) = e_j \) for \( j = 1, \ldots, n \) and then extended by linearity to all the vectors in \( U \). Clearly \( T \) is onto, therefore it is also one to one, hence it is an isomorphism. \( \square \)

**Example.**

1. As a vector space, \( \mathcal{P}_n \) is essentially \( \mathbb{R}^{n+1} \).

\(^1\)The prefix *iso-* comes from the Greek word *isos*, which means equal, same. Then "isomorphic" means having the same shape.