Ch 1. Linear Equations (2/2)

Sim, Min Kyu, Ph.D., mksim@seoultech.ac.kr

서울과학기술대학교
SEOULTECH SEOUL NATIONAL UNIVERSITY OF SCIENCE & TECHNOLOGY

1.5. Solution Sets of Linear Systems

- A system of linear equations is said to be *homogeneous* if it can be written in the form $A\mathbf{x} = \mathbf{0}$, where A is an $m \times n$ matrix and $\mathbf{0}$ is the zero vector in $\mathbb{R}^m$.
- Such a system $A\mathbf{x} = \mathbf{0}$ always has at least one solution, namely, $\mathbf{x} = \mathbf{0}$ (the zero vector in $\mathbb{R}^n$). This zero solution is usually called the *trivial solution*.
- There being a trivial solution, the important question for a homogeneous system is whether there exists a *nontrivial solution*, that is, a nonzero vector $\mathbf{x}$ that satisfies $A\mathbf{x} = \mathbf{0}$.

Homogeneous Linear Systems

Example 1. Determine if the following homogeneous system has a nontrivial solution. Then describe the solution set.

$$\begin{aligned} 3x_1 + 5x_2 - 4x_3 &= 0 \\ -3x_1 - 2x_2 + 4x_3 &= 0 \\ 6x_1 + x_2 - 8x_3 &= 0 \end{aligned}$$

Solution: Let A be the matrix of coefficients of the system and row reduce the augmented matrix $[A\ \mathbf{0}]$ to echelon form.

$$\begin{pmatrix} 3 & 5 & -4 & 0 \\ -3 & -2 & 4 & 0 \\ 6 & 1 & -8 & 0 \end{pmatrix} \sim \begin{pmatrix} 3 & 5 & -4 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & -9 & 0 & 0 \end{pmatrix} \sim \begin{pmatrix} 3 & 5 & -4 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

- Since x_3 is a free variable, we know that there exists a nontrivial solution. That is, setting x_3 equal to any arbitrary number will still generate a legit solution.

- Continue the row reduction of $[A \ \mathbf{0}]$ to *reduced* echelon form:

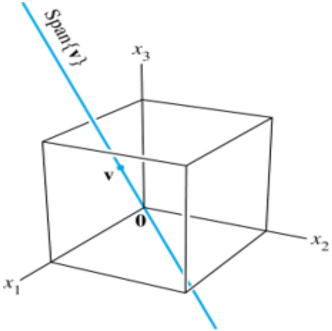
$$\begin{pmatrix} 1 & 0 & -\frac{4}{3} & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\begin{aligned} x_1 - \frac{4}{3}x_3 &= 0 \\ x_2 &= 0 \\ 0 &= 0 \end{aligned}$$

- Solving for basic variables, x_1, x_2 to obtain $x_1 = \frac{4}{3}x_3$, $x_2 = 0$, with x_3 free. As a vector, this solution can be written as

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} \frac{4}{3}x_3 \\ 0 \\ x_3 \end{pmatrix} = x_3 \begin{pmatrix} \frac{4}{3} \\ 0 \\ 1 \end{pmatrix} = x_3\mathbf{v}, \quad \text{where } \mathbf{v} = \begin{pmatrix} \frac{4}{3} \\ 0 \\ 1 \end{pmatrix}$$

- Here x_3 is factored out of the expression for the general solution vector.
- This shows that every solution of $A\mathbf{x} = \mathbf{0}$ in this case is a scalar multiple of $\mathbf{v}$.
- The trivial solution is obtained by choosing $x_3 = 0$.
- Geometrically, the solution set is a line through 0 in $\mathbb{R}^3$.



Parametric vector form

- The equation of the form $\mathbf{x} = s\mathbf{u} + t\mathbf{v}$ ($s, t \in \mathbb{R}$) is called *a parametric vector equation* of the plane.
- In **Example 1**, the equation $\mathbf{x} = x_3\mathbf{v}$ (with x_3 free), or $\mathbf{x} = t\mathbf{v}$ (with $t \in \mathbb{R}$) is a parametric vector equation of a line.
- Whenever a solution set is described explicitly with vectors as in Example 1, we say that the solution is in *parametric vector form*.

Solutions of nonhomogenous systems

- When a nonhomogeneous linear system has many solutions, the general solution can be written in parametric vector form as one vector plus an arbitrary linear combination of vectors that satisfy the corresponding homogeneous system.

Example 3. Describe all solutions of $A\mathbf{x} = \mathbf{b}$, where

$$A = \begin{pmatrix} 3 & 5 & -4 \\ -3 & -2 & 4 \\ 6 & 1 & -8 \end{pmatrix} \text{ and } \mathbf{b} = \begin{pmatrix} 7 \\ -1 \\ -4 \end{pmatrix}$$

Solution Row operations on $[A \ \mathbf{b}]$ produce

$$[A \ \mathbf{b}] \sim \begin{pmatrix} 3 & 5 & -4 & 7 \\ -3 & -2 & 4 & -1 \\ 6 & 1 & -8 & -4 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & -\frac{4}{3} & -1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

- Thus, $x_1 = -1 + \frac{4}{3}x_3,\ x_2 = 2$, and x_3 is free.
- As a vector, the general solution of $A\mathbf{x} = \mathbf{b}$ has the form

$$x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -1 + \frac{4}{3}x_3 \\ 2 \\ x_3 \end{pmatrix} =$$

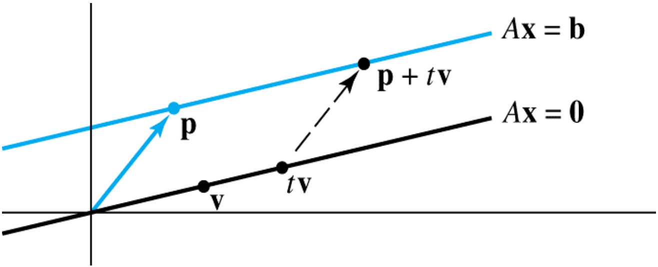
- The equation $x = \mathbf{p} + x_3\mathbf{v}$, or, writing t as a general parameter,

$$x = \mathbf{p} + t\mathbf{v} \ (t \in \mathbb{R})$$

  describes the solution set in parametric vector form.
- The vector $\mathbf{p}$ itself is a solution to the nonhomogeneous system, by letting $t = 0$.
- The solution to nonhomogenous system is composed of *a constant multiple of homogeneous solution* ($\mathbf{v}$) and *particular solution* ($\mathbf{p}$).

$$x = \mathbf{p} + t\mathbf{v} \ (t \in \mathbb{R})$$

- *The solution is, therefore, the line through* $\mathbf{p}$ *parallel to* $\mathbf{v}$.
- The solution is, a line through $\mathbf{p}$ parallel to the solution set of $A\mathbf{x} = \mathbf{0}$.



- The relation between the solution sets of $A\mathbf{x} = \mathbf{b}$ and $A\mathbf{x} = \mathbf{0}$ shown in the figure above generalizes to any consistent equation $A\mathbf{x} = \mathbf{b}$, although the solution set will be larger than a line when there are several free variables.

Theorem

Suppose the equation $A\mathbf{x} = \mathbf{b}$ is consistent for some given $\mathbf{b}$, and let $\mathbf{p}$ be a solution. Then the solution set of $A\mathbf{x} = \mathbf{b}$ is the set of all vectors of the form $\mathbf{w} = \mathbf{p} + \mathbf{v_h}$, where v_h is any solution of the homogeneous equation.

- If $A\mathbf{x} = \mathbf{b}$ has a solution, then the solution set is obtained by translating the solution set of $A\mathbf{x} = \mathbf{0}$, using any particular solution $\mathbf{p}$ of $A\mathbf{x} = \mathbf{b}$ for the translation.
- (*Translating* means moving toward a specific direction).

Writing a solution set (of a consistent system) in parametric vector form

1. Row reduce the augmented matrix to reduced echelon form.
2. Express each basic variable in terms of any free variables appearing in an equation.
3. Write a typical solution $\mathbf{x}$ as a vector whose entries depend on the free variables, if any.
4. Decompose $\mathbf{x}$ into a linear combination of vectors (with numeric entries) using the free variables as parameters.

1.7. Linear Independence

1.7. Linear Independence

Definition (Linear independence)

An indexed set of vectors $\{v_1, ..., v_p\}$ in $\mathbb{R}^n$ is said to be *linearly independent* if the vector equation

$$x_1\mathbf{v_1} + x_2\mathbf{v_2} + ... + x_p\mathbf{v_p} = \mathbf{0} \tag{1}$$

has only the trivial solution.

Definition (Linear dependence)

An indexed set of vectors $\{v_1, ..., v_p\}$ in $\mathbb{R}^n$ is said to be *linearly dependent* if the vector equation

$$x_1\mathbf{v_1} + x_2\mathbf{v_2} + ... + x_p\mathbf{v_p} = \mathbf{0} \tag{2}$$

has only the nontrivial solution. (some of $x_i,\ 1 \leq i \leq p$, is not zero.)

$$x_1\mathbf{v_1} + x_2\mathbf{v_2} + ... + x_p\mathbf{v_p} = \mathbf{0}$$

- Adopting the perspective of matrix multiplication, it follows

$$\begin{pmatrix} | & | & & | \\ \mathbf{v_1} & \mathbf{v_2} & \cdots & \mathbf{v_p} \\ | & | & & | \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ ... \\ x_p \end{pmatrix} = \mathbf{0}$$

- Each linear dependence relation among the columns of A corresponds to a nontrivial solution of $A\mathbf{x} = \mathbf{0}$
- The columns of matrix A are linearly independent if and only if the equation $A\mathbf{x} = \mathbf{0}$ has only the trivial solution.
- *In order to check the linear dependence/independence of vectors, arrange them into column vectors of a matrix, and look for a homogeneous solution. If any nontrivial solution, they are linearly dependent, otherwise linearly independent.*

Sets of one vector

- A set containing only one vector – say, $\mathbf{v}$ – is linearly independent if and only if $\mathbf{v}$ is not the zero vector.
- This is because the vector equation $x_1\mathbf{v} = \mathbf{0}$ has only the trivial solution when $\mathbf{v} \neq \mathbf{0}$.
- The zero vector is linearly dependent because $x_1\mathbf{0} = \mathbf{0}$ has many nontrivial solutions.

Sets of two vectors

- A set of two vectors $\{\mathbf{v_1}, \mathbf{v_2}\}$ is linearly dependent if at least one of the vectors is a multiple of the other.
- The set is linearly independent if and only if neither of the vectors is a multiple of the other.

Sets of two or more vectors

Theorem (Characterization of linearly dependent sets)

An indexed set $S = \{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_p}\}$ of two or more vectors is linearly dependent if and only if at least one of the vectors in S is a linear combination of the others. In fact, if S is linearly dependent and $\mathbf{v_1} \neq \mathbf{0}$, then some $\mathbf{v_j}$ (with $j > 1$) is a linear combination of the preceding vectors, $v_1, \ldots, v_{j-1}$.

Theorem

If a set contains more vectors than there are entries in each vector, then the set is linearly dependent. That is, any set $\{\mathbf{v_1}, \ldots, \mathbf{v_p}\}$ in $\mathbb{R}^n$ is linearly dependent if $p > n$.

$$\begin{array}{cc} & p \\ n & \begin{bmatrix} * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \end{bmatrix} \end{array}$$

- If $p > n$, the columns are linearly dependent.

Theorem

If a set $S = \{\mathbf{v_1}, \mathbf{v_2}, ..., \mathbf{v_p}\}$ *in* $\mathbb{R}^n$ *contains the zero vector, then the set is linearly dependent.*

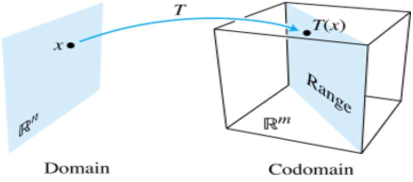
Theorem

There can be at most n *linearly independent vectors where each vector is in* $\mathbb{R}^n$*.*

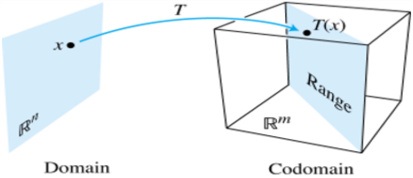
1.8. Introduction to Linear Transformations

A few definitions

- A *transformation* (or *function* or *mapping*) T from $\mathbb{R}^n$ to $\mathbb{R}^m$ is a rule that assigns to each vector $\mathbf{x}$ in $\mathbb{R}^n$ a vector $T(\mathbf{x})$ in $\mathbb{R}^m$.
- The set $\mathbb{R}^n$ is called *domain*(정의역) of T
- The set $\mathbb{R}^m$ is called the *codomain*(공역) of T.
- The notation $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ indicates that the domain of T is $\mathbb{R}^n$ and the codomain is $\mathbb{R}^m$.
- For x in $\mathbb{R}^n$, the vector $T(\mathbf{x})$ in $\mathbb{R}^m$ is called the *image*(치역) of $\mathbf{x}$ (under the action of T).
- The set of all images $T(\mathbf{x})$ is called the range of T. See figure below.



Domain, codomain, and range of $T : \mathbb{R}^n \to \mathbb{R}^m$.



Domain, codomain, and range of $T : \mathbb{R}^n \to \mathbb{R}^m$.

- For each $\mathbf{x} \in \mathbb{R}^n$, $T(\mathbf{x})$ is computed as $A\mathbf{x}$, where A is an matrix.
- For simplicity, we denote such a *matrix transformation* by $\mathbf{x} \mapsto A\mathbf{x}$ ("maps to' ').
- Observe that the domain of T is $\mathbb{R}^n$ when A has n columns and the codomain of T is $\mathbb{R}^m$ when each column of A has m entries.
- The range of T is the set of all linear combinations of the columns of A, because each image $T(\mathbf{x})$ is of the form $A\mathbf{x}$.

Example 1. Let

$$A = \begin{pmatrix} 1 & -3 \\ 3 & 5 \\ -1 & 7 \end{pmatrix}, \ \mathbf{u} = \begin{pmatrix} 2 \\ -1 \end{pmatrix}, \ \mathbf{b} = \begin{pmatrix} 3 \\ 2 \\ -5 \end{pmatrix}, \ \mathbf{c} = \begin{pmatrix} 3 \\ 2 \\ 5 \end{pmatrix}$$

and define a transformation $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ by $T(\mathbf{x}) = A\mathbf{x}$, so that

$$T(x) = A\mathbf{x} = \begin{pmatrix} 1 & -3 \\ 3 & 5 \\ -1 & 7 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 - 3x_2 \\ 3x_1 + 5x_2 \\ -x_1 + 7x_2 \end{pmatrix}$$

- **a.** Find $T(\mathbf{u})$, the image of u under the transformation T.

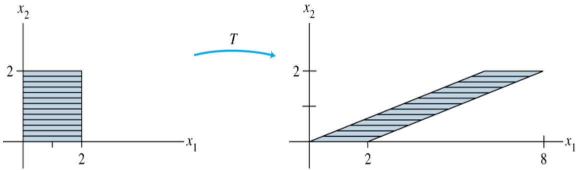
- **b.** Find an **x** in $\mathbb{R}^2$ whose image under T is $\mathbf{b}$.

- **c.** Is there more than one **x** whose image under T is $\mathbf{b}$?

- **d.** Determine if **c** is in the range of the transformation T.

Shear transformation

Example 3. Let $A = \begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix}$. The transformation $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ defined by $T(\mathbf{x}) = A\mathbf{x}$ is called a *shear transformation*.



- The key idea is to show that T maps line segments onto line segments and then to check that the *corners of the square* map onto the vertices of the parallelogram.
- For instance, the image of the point For instance, the image of the point $\mathbf{u} = \begin{pmatrix} 0 \\ 2 \end{pmatrix}$ is $T(\mathbf{u}) = \begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 2 \end{pmatrix} = \begin{pmatrix} 6 \\ 2 \end{pmatrix}$.

Linear transformation

Definition (Linear transformation)

A transformation (or mapping) T is linear if:

- $T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$ for all $\mathbf{u}$, $\mathbf{v}$ in the domain of T
- $T(c\mathbf{u}) = cT(\mathbf{u})$ for all scalars c and all $\mathbf{u}$ in the domain of T.

Linear transformations *preserve the operations of vector addition and scalar multiplication*.

- It follows both $T(\mathbf{0}) = \mathbf{0}$ and $T(c\mathbf{u} + d\mathbf{v}) = cT(\mathbf{u}) + dT(\mathbf{v})$.
- In engineering and physics, the following equation is referred to as a *superposition principle*.

$$T(c_1\mathbf{v_1} + \cdots + c_p\mathbf{v_p}) = c_1T(\mathbf{v_1}) + \cdots + c_pT(\mathbf{v_p})$$

- Given a scalar r, define $T : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ by $T(\mathbf{x}) = r\mathbf{x}$. What is the corresponding matrix?
- T is called a *contraction*(수축) when $0 \leq r \leq 1$ and a *dilation*(확대) when $r > 1$.

1.9. The Matrix of Linear Transformation

The matrix of linear transformation

Theorem

Let $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ be a linear transformation. Then there exists a unique matrix A such that $T(\mathbf{x}) = A\mathbf{x}$ for all $\mathbf{x} \in \mathbb{R}^n$.

Remark

- In fact, A is the $m \times n$ matrix whose j-th column is the vector $T(\mathbf{e_j})$ where $\mathbf{e_j}$ is the j-th column of identity matrix in $\mathbb{R}^n$. (or a j-th *unit vector*). Namely,

$$A = \begin{pmatrix} | & & | \\ T(\mathbf{e_1}) & \cdots & T(\mathbf{e_n}) \\ | & & | \end{pmatrix}$$

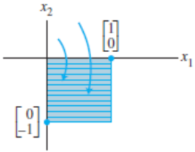
Remark

- This matrix A is called the *standard matrix for the linear transformation*.
- Every *linear transformation* from $\mathbb{R}^n$ to $\mathbb{R}^m$ can be viewed as a *matrix transformation*, and vice versa.
- The term *linear transformation* focuses on a property of a mapping, while matrix transformation describes how such a mapping is implemented.

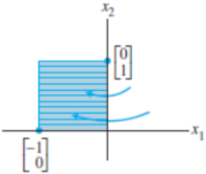
Geometric linear transformation in $\mathbb{R}^2$

Reflections

TABLE 1 Reflections

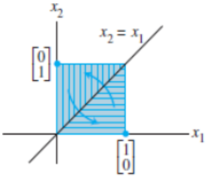
Transformation	**Image of the Unit Square**	**Standard Matrix**
Reflection through the x_1-axis		$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$

Reflection through the x_2-axis



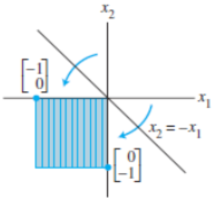
$$\begin{bmatrix}-1 & 0\\ 0 & 1\end{bmatrix}$$

Reflection through the line $x_2 = x_1$



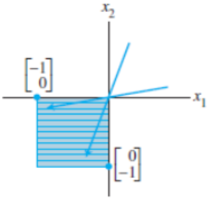
$$\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}$$

Reflection through the line $x_2 = -x_1$



$$\begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}$$

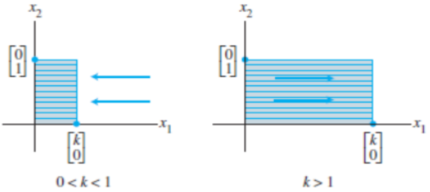
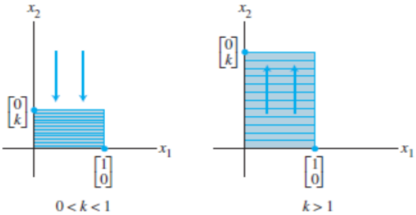
Reflection through the origin



$$\begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$$

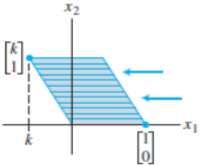
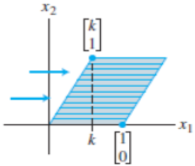
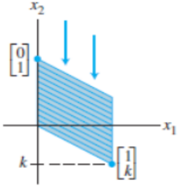
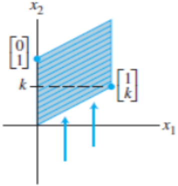
Contractions and expansions

TABLE 2 Contractions and Expansions

Transformation	Image of the Unit Square	Standard Matrix
Horizontal contraction and expansion		$\begin{bmatrix} k & 0 \\ 0 & 1 \end{bmatrix}$
Vertical contraction and expansion		$\begin{bmatrix} 1 & 0 \\ 0 & k \end{bmatrix}$

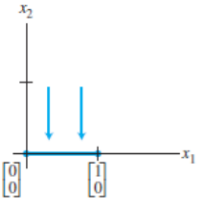
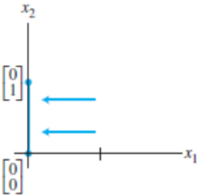
Shears

TABLE 3 Shears

Transformation	Image of the Unit Square		Standard Matrix
Horizontal shear	 $k < 0$	 $k > 0$	$\begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix}$
Vertical shear	 $k < 0$	 $k > 0$	$\begin{bmatrix} 1 & 0 \\ k & 1 \end{bmatrix}$

Projections

TABLE 4 Projections

Transformation	Image of the Unit Square	Standard Matrix
Projection onto the x_1-axis		$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$
Projection onto the x_2-axis		$\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$

Definition (onto)

A mapping $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is said to be **onto** $\mathbb{R}^n$ if each $\mathbf{b} \in \mathbb{R}^m$ is the image of at least one $\mathbf{x} \in \mathbb{R}^n$.

Remark

- If *onto*, *codomain* is same set as *range*.
- If *not onto*, there is some $\mathbf{b} \in \mathbb{R}^m$ for which the equation $T(\mathbf{x}) = \mathbf{b}$ has no solution.

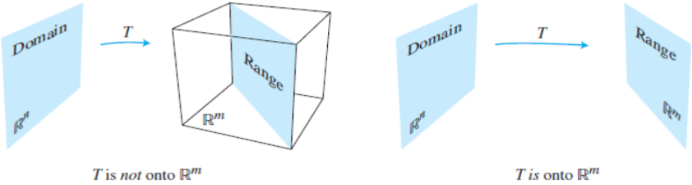


FIGURE 3 Is the range of T all of $\mathbb{R}^m$?

Definition (one-to-one)

A mapping $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is said to be **one-to-one** if each $\mathbf{b} \in \mathbb{R}^n$ is the image of **at most one** $\mathbf{x} \in \mathbb{R}^m$.

Example 4. Let T be the linear transformation whose standard matrix is

$$A = \begin{pmatrix} 1 & -4 & 8 & 1 \\ 0 & 2 & -1 & 3 \\ 0 & 0 & 0 & 5 \end{pmatrix}$$

- Does T map $\mathbb{R}^4$ onto $\mathbb{R}^3$?
- Is T a one-to-one mapping?

Theorem

*Let $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ be a linear transformation. Then T is **one-to-one** if and only if the equation $T(\mathbf{x}) = \mathbf{0}$ has only the trivial solution.*

Theorem

Let $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ be a linear transformation and let A be the standard matrix for T. Then,

- *T maps $\mathbb{R}^n$ onto $\mathbb{R}^m$ if and only if the columns of A span $\mathbb{R}^m$.*
- *T is one-to-one if and only if the columns of A are linearly independent.*

Suggested Exercises

- 1.5.14, 1.5.18
- 1.8.1, 1.8.9, 1.8.13, 1.8.14, 1.8.15, 1.8.16

```
"Man can learn nothing unless he proceeds from the known to the unknown. - Claude Bernard"
```