Matrix Theory — Lecture notes
MAT 335, Fall 2022 — D. Ivanšić

2.1 Matrix Multiplication

Let $\mathbf{v}$ be a $p \times 1$ vector, B an $n \times p$ matrix and A an $m \times n$ matrix. These dimensions have been set up so that

$$B\mathbf{v} \text{ is defined, and is an } n \times 1 \text{ vector, and } A(B\mathbf{v}) \text{ is defined, and is an } m \times 1 \text{ vector}$$

We could ask if there is a single matrix C, necessarily with dimensions $m \times p$, so that

$$C\mathbf{v} = A(B\mathbf{v}), \text{ for every vector } \mathbf{v} \text{ in } \mathbf{R}^p$$

Definition. Let A an $m \times n$ matrix and B an $n \times p$ matrix, where $\mathbf{b}_1, \dots \mathbf{b}_p$ are columns of B. We define the matrix product of A and B as the matrix $m \times p$ matrix C with columns

$$AB = C = \begin{bmatrix} A\mathbf{b}_1 & A\mathbf{b}_2 & \dots & A\mathbf{b}_p \end{bmatrix}$$

The columns $A\mathbf{b}_1, \dots, A\mathbf{b}_p$ are $m \times 1$, so AB is an $m \times p$ matrix.

Note. Under this setup, $(AB)\mathbf{v} = A(B\mathbf{v})$ for every $\mathbf{v}$ in $\mathbf{R}^p$, because that was how AB was defined. For dimensions, we write $(m \times n)(n \times p) = (m \times p)$, and the product is defined when the inner dimensions are equal.

In the definition of the product, notice that the (i, j)-entry in the matrix AB is the i-th component of the vector $A\mathbf{b}_j$, which is the dot product of the i-th row of A with the vector $\mathbf{b}_j$, thus

the (i, j)-entry of the matrix AB is the dot product
of the i-th row of A with the j-th column of B

Example. Find the product

$$\begin{bmatrix} 2 & -2 & 3 \\ 1 & 0 & -1 \end{bmatrix} \begin{bmatrix} 1 & 7 & 1 & -3 \\ 0 & -1 & 4 & 3 \\ 4 & 2 & 2 & 5 \end{bmatrix}$$

Example. Recall the data on nutritional value of various foods per 100g serving. Now consider two menus with indicated numbers of servings of chicken, rice and lettuce.

	chicken	rice	lettuce
energy (kcal)	149	359	17
fat (g)	6	1	0
protein (g)	24	7	1
carbohydrates (g)	0	80	3

m. 1	m. 2	
2	3	chicken
1	0	rice
2	3	lettuce

Compute the product of related matrices and interpret the meaning of the resulting matrix.

$$\begin{bmatrix} 149 & 359 & 17 \\ 6 & 1 & 0 \\ 24 & 7 & 1 \\ 0 & 80 & 3 \end{bmatrix} \begin{bmatrix} 2 & 3 \\ 1 & 0 \\ 2 & 3 \end{bmatrix}$$

Theorem 1.6. Let A, B be $m \times n$ matrices, C, D be $n \times p$ matrices and E, F be $p \times q$ matrices, and s a scalar. Then the following statements are true:

(a) $s(AC) = (sA)C = A(sC)$
(b) $A(CE) = (AC)E$ (associativity)
(c) $(A + B)C = AC + BC$ (right distributive law)
(d) $A(C + D) = AC + AD$ (left distributive law)
(e) $I_m A = A = AI_n$
(f) Product of a matrix with a zero matrix is a zero matrix
(g) $(AC)^T = C^T A^T$

Proof. Statements a)–f) are essentially consequences of similar rules for matrix-vector multiplication. Justifying g).

Example. The commutative rule $AB = BA$ is absent, because it is NOT true in general. First, for both AB and BA to be defined and equal sizes, they have to be square. Compute the products below to see $AB \neq BA$ (actually, in this example $AA^T \neq A^T A$).

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

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

Definition. A *block matrix* is a matrix that is thought of as consisting of smaller matrices. A simple example is the matrix $[\ A \quad B\ ]$ that consists of matrices A, B with equal numbers of rows m. If C is a $k \times m$ matrix, then it is easy to see (consider columns of resulting matrices) that

$$C[\ A \quad B\ ] = [\ CA \quad CB\ ]$$

Definition. The (i, j)-entry of a matrix A is called a *diagonal entry* if $i = j$. The diagonal entries form the *diagonal* of A.

A *square* matrix is a *diagonal matrix* if all nondiagonal entries are zero, for example the zero matrix and I_n.

Example. Compute the products and make observations.

$$\begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 1 & 3 & -2 \\ -5 & 0 & 4 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 3 & -2 \\ -5 & 0 & 4 \end{bmatrix} \begin{bmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

Definition. A square matrix A is called

symmetric, if $A^T = A$ skew-symmetric, if $A^T = -A$

Example.

$$\begin{bmatrix} 1 & 3 & -5 \\ 3 & -1 & 2 \\ -5 & 2 & 4 \end{bmatrix} \qquad \begin{bmatrix} 0 & 3 & -5 \\ -3 & 0 & 2 \\ 5 & -2 & 0 \end{bmatrix}$$

Matrix Theory — Lecture notes MAT 335, Fall 2022 — D. Ivanšić

2.3 Invertibility and Elementary Matrices

Definition. An $n \times n$ matrix A is called *invertible* if there exists an $n \times n$ matrix B such that $AB = BA = I_n$. In this case, B is called the *inverse* of A.

Note. If an inverse of A exists, it is unique and we denote it A^{-1}.

Example.

$$\begin{bmatrix} 1 & 4 \\ 3 & 11 \end{bmatrix} \begin{bmatrix} -11 & 4 \\ 3 & -1 \end{bmatrix} = \qquad\qquad \begin{bmatrix} -11 & 4 \\ 3 & -1 \end{bmatrix} \begin{bmatrix} 1 & 4 \\ 3 & 11 \end{bmatrix} =$$

Example. The following matrices do not have an inverse, because no matrix B can multiply them to get I.

zero matrix

any matrix with a zero column

any matrix with two proportional columns

If a matrix A has an inverse, then the system $A\mathbf{x} = \mathbf{b}$ is easy to solve, $\mathbf{x} = A^{-1}\mathbf{b}$.

Example. Solve the system.

$$\begin{bmatrix} 1 & 4 \\ 3 & 11 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 \\ -1 \end{bmatrix}$$

Theorem 2.2. Let A, B be invertible $n \times n$ matrices. Then

(a) A^{-1} is invertible and $(A^{-1})^{-1} = A$

(b) AB is invertible and $(AB)^{-1} = B^{-1}A^{-1}$

(c) A^T is invertible and $(A^T)^{-1} = (A^{-1})^T$

Proof. We simply check that the matrices proposed as inverses satisfy the definition of the inverse:

Definition. An $m \times m$ matrix E is called an *elementary matrix* if it is the result of a single elementary row operation on I_m.

Example. Three elementary matrices corresponding to three elementary row operations.

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

Example. Observe what happens when these matrices multiply a matrix on the *left*.

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 4 & 2 \\ 3 & 7 \\ -2 & 5 \end{bmatrix} = \qquad \begin{bmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 4 & 2 \\ 3 & 7 \\ -2 & 5 \end{bmatrix} =$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 4 & 2 \\ 3 & 7 \\ -2 & 5 \end{bmatrix} =$$

Proposition. Multiplying an $m \times n$ matrix A by an elementary matrix E on the left results in performing the same row operation on A that produced E.

Proposition. Every elementary matrix E is invertible, and its inverse is the elementary matrix resulting from the row operation that reverses the row operation that produced E.

Proof. If F is produced by the row operation that reverses the row operation producing E, then FE will be the matrix with the reversing row operation applied to E, producing I. Therefore, $FE = I$. Similarly $EF = I$ because the row operation that produces E reverses the row operation that produces F.

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -3 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{bmatrix} = \qquad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -3 \\ 0 & 0 & 1 \end{bmatrix} =$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} = \qquad \begin{bmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix} =$$

Theorem 2.3 Let A be an $m \times n$ matrix with reduced row echelon form R. Then there exists an invertible $m \times m$ matrix P such that $PA = R$.

Proof.

Proposition. (Column Correspondence Property) For a matrix A and its reduced row echelon form R, any linear combination of columns of R that is equal to the zero vector is true, with same coefficients, for the corresponding columns of A. In particular, if column j of R is a linear combination of some other columns, then column j of A is a linear combination of the corresponding columns of A, with the same coefficients.

Proof. This follows from the fact that $A\mathbf{x} = \mathbf{0}$ and $R\mathbf{x} = \mathbf{0}$ have the same solutions.

Example. Verify the statement on these matrices and show that the columns of A containing the leading 1's are linearly independent.

$$\begin{bmatrix} -2 & -2 & 10 & -7 & 3 \\ 0 & 1 & 2 & -1 & 0 \\ 1 & 3 & -1 & 2 & 2 \\ -2 & 0 & 14 & -9 & 3 \end{bmatrix} \quad \text{has reduced row echelon form} \quad \begin{bmatrix} 1 & 0 & -7 & 0 & -33 \\ 0 & 1 & 2 & 0 & 7 \\ 0 & 0 & 0 & 1 & 7 \end{bmatrix}$$

2.4 The Inverse of a Matrix

Theorem 2.5. Let A be an $n \times n$ matrix. Then A is invertible if and only if its reduced row echelon form is I_n.

Proof.

The theorem helps us determine whether A is invertible and find its inverse.

Algorithm for Matrix Inversion. Let A be an $n \times n$ matrix, and use row operations to transform $[\ A \quad I_n\ ]$ to form $[\ R \quad B\ ]$, where R is the reduced row echelon form of A. Then either

(a) $R = I_n$, in which case A is invertible and $B = A^{-1}$, or

(b) $R \neq I_n$, in which case A is not invertible.

Example. Find the inverse of the matrix at left.

$$\left[\begin{array}{rrr|rrr} 1 & -4 & 7 & 1 & 0 & 0 \\ 3 & -10 & 26 & 0 & 1 & 0 \\ 1 & -3 & 10 & 0 & 0 & 1 \end{array}\right]$$

Theorem 2.6. **(Invertible Matrix Theorem)** Let A be an $n \times n$ matrix. Then the following statements are equivalent:

(a) A is invertible.

(b) The reduced row echelon form of A is I_n.

(c) $\operatorname{rank} A = n$

(d) Span of columns of A is $\mathbf{R}^n$.

(e) The equation $A\mathbf{x} = \mathbf{b}$ is consistent for every b in $\mathbf{R}^n$.

(f) $\operatorname{nullity} A = 0$

(g) The columns of A are linearly independent.

(h) The only solution to $A\mathbf{x} = \mathbf{0}$ is $\mathbf{0}$.

(i) There exists an $n \times n$ matrix B such that $BA = I_n$.

(j) There exists an $n \times n$ matrix C such that $AC = I_n$.

(k) A is a product of elementary matrices.

Proof. $a \iff b$ by Theorem 2.5, $b \iff c \iff d \iff e \iff f \iff g \iff h$ by Theorems 1.6, 1.8 and the fact that the matrix is $n \times n$. We show $a \iff k$, $a \implies i \implies h \implies a$ and $a \implies j \implies e \implies a$.

Example. Note that $BA = I$ implies that A is invertible only because A is a square matrix. This is not true for a non-square matrix. The product of the matrices below is I_2, but neither is invertible.

$$\begin{bmatrix} 1 & 4 & 2 \\ 3 & 11 & 7 \end{bmatrix} \begin{bmatrix} -11 & 4 \\ 3 & -1 \\ 0 & 0 \end{bmatrix} =$$

Matrix Theory — Lecture notes
MAT 335, Fall 2022 — D. Ivanšić

2.7 Linear Transformations and Matrices

Definition. Let X and Y be sets. A function f from X to Y is a rule that assigns to every element x of X a unique element $f(x)$ of Y. Furthermore, we define these terms:

— the element $f(x)$ is called the *image* of x (under f)

— the set X is called the *domain* of f

— the set Y is called the *codomain* of f

— the *range of* f is the set of images $f(x)$ for all x in X.

Example. Consider $f : \{1, 2, 3, 4\} \to \{4, 7, 9\}$ given by the table. Identify the sets discussed in the definition for this example.

x	1	2	3	4
$f(x)$	9	4	9	4

We will mainly be considering functions $f : \mathbf{R}^n \to \mathbf{R}^m$, they send vectors $\mathbf{v}$ in $\mathbf{R}^n$ to vectors $f(\mathbf{v})$ in $\mathbf{R}^m$.

Example. For the matrix $\begin{bmatrix} 3 & -7 & 1 \\ 4 & 2 & -1 \end{bmatrix}$, consider the function $T_A : \mathbf{R}^3 \to \mathbf{R}^2$ given by $T_A(\mathbf{x}) = A\mathbf{x}$. This function sends vectors from space to vectors in a plane. Write the formula for $T_A\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right)$.

Example. Consider $f : \mathbf{R}^2 \to \mathbf{R}^2$, $f(\mathbf{v})$ = the vector obtained by rotating $\mathbf{v}$ by $\frac{3\pi}{4}$. Then we have seen that

$$f(\mathbf{v}) = \begin{bmatrix} -\frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \end{bmatrix} \mathbf{v}$$

Definition. Let A be an $m \times n$ matrix. The function $T_A : \mathbf{R}^n \to \mathbf{R}^m$ given by $T_A(\mathbf{x}) = A\mathbf{x}$ is called a *matrix transformation induced by A*.

In this course we will mainly be considering functions $T : \mathbf{R}^n \to \mathbf{R}^m$ that are matrix transformations.

Example. What does the matrix transformation induced by the matrix below do?

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

Example. What does the matrix transformation induced by the matrix below do?

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

(shear transformation)

In this course we will mainly be considering functions $T : \mathbf{R}^n \to \mathbf{R}^m$ that have a special property.

Definition. A function $T : \mathbf{R}^n \to \mathbf{R}^m$ is called a *linear transformation* or just *linear* if, for all vectors $\mathbf{u}$, $\mathbf{v}$ in $\mathbf{R}^n$ and all scalars c, we have:

(i) $T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$ (T preserves vector addition)

(ii) $T(c\mathbf{u}) = cT(\mathbf{u})$ (T preserves scalar multiplication)

Example. This is easy: find an example of a function $T : \mathbf{R} \to \mathbf{R}$ which:
a) fails to preserve vector addition
b) fails to preserve scalar multiplication

Example (Theorem 2.7). Show that every matrix transformation $T_A : \mathbf{R}^n \to \mathbf{R}^m$ is a linear transformation.

Theorem 2.8. For all vectors $\mathbf{u}, \mathbf{v}$ in $\mathbf{R}^n$ and all scalars a, b, every linear transformation $T : \mathbf{R}^n \to \mathbf{R}^m$ satisfies:

(a) $T(\mathbf{0}) = \mathbf{0}$

(b) $T(-\mathbf{u}) = -T(\mathbf{u})$

(c) $T(\mathbf{u} - \mathbf{v}) = T(\mathbf{u}) - T(\mathbf{v})$

(d) $T(a\mathbf{u} + b\mathbf{v}) = aT(\mathbf{u}) + bT(\mathbf{v})$

(e) $T(a_1\mathbf{u}_1 + \ldots a_k\mathbf{u}_k) = a_1T(\mathbf{u}_1) + \cdots + a_kT(\mathbf{u}_k)$, for all vectors $\mathbf{u}_1, \ldots \mathbf{u}_k$ in $\mathbf{R}^n$ and all scalars $a_1, \ldots, a_k$.

Proof.

Theorem 2.9. Let $T : \mathbf{R}^n \to \mathbf{R}^m$ be a linear transformation. Then T is a matrix transformation T_A whose matrix

$$A = \begin{bmatrix} T(\mathbf{e}_1) & T(\mathbf{e}_2) & \ldots & T(\mathbf{e}_k) \end{bmatrix}$$

consists of columns that are images under T of standard basis vectors $\mathbf{e}_1, \ldots, \mathbf{e}_n$ of $\mathbf{R}^n$.

Proof.

Definition. The matrix

$$A = \begin{bmatrix} T(\mathbf{e}_1) & T(\mathbf{e}_2) & \ldots & T(\mathbf{e}_k) \end{bmatrix}$$

is called the *standard matrix* of T. (It has the property that $T(\mathbf{v}) = A\mathbf{v}$ for every $\mathbf{v}$ in $\mathbf{R}^n$.

Matrix Theory — Lecture notes
MAT 335, Fall 2022 — D. Ivanšić

2.8 Composition, Invertibility of Linear Transformations

Definition. Let X and Y be sets. A function f from X to Y is said to be:

— *onto*, if the range of f equals Y.

— *one-to-one*, if it sends distinct elements to distinct images, in other words, if $x_1 \neq x_2$, then $f(x_1) \neq f(x_2)$, for all x_1, x_2 in X.

— (equivalently) *onto*, if for every y in Y there is an x in X so that $f(x) = y$.

— (equivalently) *one-to-one*, if $f(x_1) = f(x_2)$, then $x_1 = x_2$, for all x_1, x_2 in X.

Example. Consider functions with codomain $\{4, 7, 9, 11\}$ given by the tables below. Which ones are a) onto? b) one-to-one ?

x	1	2	3
$f(x)$	9	4	7

x	1	2	3	4
$g(x)$	9	7	11	4

x	1	2	3	4
$h(x)$	9	9	11	4

x	1	2	3	4	5
$k(x)$	9	4	9	11	7

Now consider linear transformations $T : \mathbf{R}^n \to \mathbf{R}^m$ (that is, matrix transformations).

Proposition. The range of a linear transformation T is the span of the columns of its standard matrix.

Example. Is the linear transformation $T : \mathbf{R}^2 \to \mathbf{R}^3$ onto?

$$T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} 2x_1 + x_2 \\ -x_1 + 3x_2 \\ 5x_1 - x_2 \end{bmatrix}$$

Theorem 2.10. Let $f : \mathbf{R}^n \to \mathbf{R}^m$ be a linear transformation with standard matrix A. The following statements are equivalent:

(a) T is onto, that is, range of T is $\mathbf{R}^m$.

(b) The columns of A span $\mathbf{R}^m$.

(c) $\operatorname{rank} A = m$

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

Proof.

Definition. Let $T : \mathbf{R}^n \to \mathbf{R}^m$ be a linear transformation. The *null space* of T is the set of all vectors $\mathbf{v}$ in $\mathbf{R}^n$ such that $T(\mathbf{v}) = \mathbf{0}$. Note that $\mathbf{0}$ is always in the null space of T.

Proposition. A linear transformation is one-to-one if and only if its null space contains only $\mathbf{0}$.

Proof.

Example. Is the linear transformation $T : \mathbf{R}^2 \to \mathbf{R}^3$ one-to-one?

$$T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} 2x_1 + x_2 \\ -x_1 + 3x_2 \\ 5x_1 - x_2 \end{bmatrix}$$

Theorem 2.11. Let $f : \mathbf{R}^n \to \mathbf{R}^m$ be a linear transformation with standard matrix A. The following statements are equivalent:

(a) T is one-to-one.

(b) The null space of T consists only of the zero vector.

(c) The columns of A are linearly independent.

(d) $\operatorname{rank} A = n$

(d) The only solution of the equation $A\mathbf{x} = \mathbf{0}$ is $\mathbf{0}$.

Proof.

Definition. Let $f : X \to Y$ and $g : Y \to Z$ be functions. The composition of functions f and g is defined to be the function $g \circ f : X \to Z$ given by

$$(g \circ f)(x) = g(f(x)), \text{ for every } x \text{ in } X$$

Example. Consider functions $f : \{1, 2, 3, 4\} \to \{4, 7, 9\}$ and $g : \{4, 7, 9\} \to \{15, 17, 20, 24\}$ given by the tables. Determine the function $g \circ f$.

x	1	2	3	4
$f(x)$	7	4	9	4

x	4	7	9
$g(x)$	20	15	24

x	1	2	3	4
$(g \circ f)(x)$				

Example. Let $T_A : \mathbf{R}^n \to \mathbf{R}^m$ and $T_B : \mathbf{R}^m \to \mathbf{R}^p$ be linear transformations induced by matrices A and B. Show that $T_B \circ T_A = T_{BA}$.

For this reason, when writing compositions of linear transformations, we usually omit "$\circ$", so $T_B \circ T_A$ is written as $T_B T_A$, thus, the above example reads as $T_B T_A = T_{BA}$.

Theorem 2.12. Let $T : \mathbf{R}^n \to \mathbf{R}^m$ and $U : \mathbf{R}^m \to \mathbf{R}^p$ be linear transformations with standard matrices A and B, respectively. Then $UT : \mathbf{R}^n \to \mathbf{R}^p$ is also linear and its standard matrix is BA.

Example. Compute the composite UT of the linear transformations $T : \mathbf{R}^2 \to \mathbf{R}^3$ and $U : \mathbf{R}^3 \to \mathbf{R}^2$ directly and by using their standard matrices.

$$T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} 2x_1 + x_2 \\ -x_1 + 3x_2 \\ 5x_1 - x_2 \end{bmatrix} \qquad U\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} 3x_1 - x_2 + 2x_3 \\ -3x_1 + 4x_2 + x_3 \end{bmatrix}$$

Definition. A function $f : X \to Y$ is said to be *invertible* if there is a function $g : Y \to X$ such that $g \circ f = id_X$ and $f \circ g = id_Y$, where id_X and id_Y are identity functions on X and Y.

Note. It is easy to see that any invertible function $f : X \to Y$ has to be onto and one-to-one. If f is invertible, the function g from the definition is unique and is called the *inverse of* f, denoted f^{-1}. It is given by:

$$f^{-1}(y) = \text{ the unique } x \text{ that } f \text{ sends to } y, \text{ for every } y \text{ in } Y$$

Theorem 2.13. Let $T : \mathbf{R}^n \to \mathbf{R}^n$ be a linear transformations with standard matrix A. Then T is invertible if and only A is invertible, in which case $T^{-1} = T_{A^{-1}}$. Note this also implies that T^{-1} is linear and its standard matrix is A^{-1}.

Proof.

Table summarizing essential takeaways for a linear tranformation $T : \mathbf{R}^n \to \mathbf{R}^m$ with $m \times n$ standard matrix A.

Property of T	rank of A	Solutions of $A\mathbf{x} = \mathbf{b}$	Columns of A
T is onto	$\operatorname{rank} A = m$	at least one for every $\mathbf{b}$ in $\mathbf{R}^m$	span $\mathbf{R}^m$
T is one-to-one	$\operatorname{rank} A = n$	at most one for every $\mathbf{b}$ in $\mathbf{R}^m$	are linearly independent
T is invertible	$\operatorname{rank} A = m = n$	unique solution for every $\mathbf{b}$ in $\mathbf{R}^m$	span $\mathbf{R}^m$ and are linearly independent