4.7
Row Space, Column Space, and Null Space

Row Space and Column Space

- Definition
  - If A is an $m \times n$ matrix, then the subspace of R^n spanned by the row vectors of A is called the row space (列空間) of A, and the subspace of R^m spanned by the column vectors is called the column space (行空間) of A.
  - The solution space of the homogeneous system of equation $A\mathbf{x} = \mathbf{0}$, which is a subspace of R^n, is called the null space (零核空間) of A.

Remarks

- In this section we will be concerned with two questions
  - What relationships exist between the solutions of a linear system $A\mathbf{x}=\mathbf{b}$ and the row space, column space, and null space of A.
  - What relationships exist among the row space, column space, and null space of a matrix.

Remarks

- It follows from Formula (10) of Section 1.3

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \qquad \boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

$$A\boldsymbol{x} = x_1\boldsymbol{c}_1 + x_2\boldsymbol{c}_2 + \cdots + x_n\boldsymbol{c}_n = \boldsymbol{b}$$

- We conclude that ***A*****x=b is consistent if and only if b is expressible as a linear combination of the column vectors of *A* or, equivalently, if and only if b is in the column space of *A***.

Theorem 4.7.1

- Theorem 4.7.1
  - A system of linear equations $A\mathbf{x} = \mathbf{b}$ is consistent if and only if $\mathbf{b}$ is in the column space of A.

Example

- Let $A\mathbf{x} = \mathbf{b}$ be the linear system $\begin{bmatrix} -1 & 3 & 2 \\ 1 & 2 & -3 \\ 2 & 1 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ -9 \\ -3 \end{bmatrix}$

  Show that **b** is in the column space of *A*, and express **b** as a linear combination of the column vectors of *A*.

- Solution:
  - Solving the system by Gaussian elimination yields

    $$x_1 = 2,\ x_2 = -1,\ x_3 = 3$$

  - Since the system is consistent, **b** is in the column space of *A*.
  - Moreover, it follows that

    $$2\begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix} - \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} + 3\begin{bmatrix} 2 \\ -3 \\ -2 \end{bmatrix} = \begin{bmatrix} 1 \\ -9 \\ -3 \end{bmatrix}$$

General and Particular Solutions

- Theorem 4.7.2
  - If $\mathbf{x}_0$ denotes any single solution of a consistent linear system $A\mathbf{x} = \mathbf{b}$, and if $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ form a basis for the null space of A, (that is, the solution space of the homogeneous system $A\mathbf{x} = \mathbf{0}$), then every solution of $A\mathbf{x} = \mathbf{b}$ can be expressed in the form

$$\mathbf{x} = \mathbf{x}_0 + c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k$$

Conversely, for all choices of scalars $c_1, c_2, \ldots, c_k$, the vector $\mathbf{x}$ in this formula is a solution of $A\mathbf{x} = \mathbf{b}$.

Proof of Theorem 4.7.2

- Assume that $\mathbf{x}_0$ is any fixed solution of $A\mathbf{x}=\mathbf{b}$ and that $\mathbf{x}$ is an arbitrary solution. Then $A\mathbf{x}_0 = \mathbf{b}$ and $A\mathbf{x} = \mathbf{b}$.
- Subtracting these equations yields

$$A\mathbf{x} - A\mathbf{x}_0 = \mathbf{0} \quad \text{or} \quad A(\mathbf{x}-\mathbf{x}_0)=\mathbf{0}$$

- Which shows that $\mathbf{x}-\mathbf{x}_0$ is a solution of the homogeneous system $A\mathbf{x} = \mathbf{0}$.
- Since $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ is a basis for the solution space of this system, we can express $\mathbf{x}-\mathbf{x}_0$ as a linear combination of these vectors, say $\mathbf{x}-\mathbf{x}_0 = c_1\mathbf{v}_1+c_2\mathbf{v}_2+\ldots+c_k\mathbf{v}_k$. Thus, $\mathbf{x}=\mathbf{x}_0+c_1\mathbf{v}_1+c_2\mathbf{v}_2+\ldots+c_k\mathbf{v}_k$.

Proof of Theorem 4.7.2

- Conversely, for all choices of the scalars $c_1, c_2, \ldots, c_k$, we have

$$A\mathbf{x} = A(\mathbf{x}_0 + c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \ldots + c_k\mathbf{v}_k)$$

$$A\mathbf{x} = A\mathbf{x}_0 + c_1(A\mathbf{v}_1) + c_2(A\mathbf{v}_2) + \ldots + c_k(A\mathbf{v}_k)$$

- But $\mathbf{x}_0$ is a solution of the nonhomogeneous system, and $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ are solutions of the homogeneous system, so the last equation implies that

$$A\mathbf{x} = \mathbf{b} + \mathbf{0} + \mathbf{0} + \ldots + \mathbf{0} = \mathbf{b}$$

- Which shows that $\mathbf{x}$ is a solution of $A\mathbf{x} = \mathbf{b}$.

Remark

- Remark
  - The vector $\mathbf{x}_0$ is called a particular solution (特解) of $A\mathbf{x} = \mathbf{b}$.
  - The expression $\mathbf{x}_0 + c_1\mathbf{v}_1 + \cdot \cdot \cdot + c_k\mathbf{v}_k$ is called the general solution (通解) of $A\mathbf{x} = \mathbf{b}$, the expression $c_1\mathbf{v}_1 + \cdot \cdot \cdot + c_k\mathbf{v}_k$ is called the general solution of $A\mathbf{x} = \mathbf{0}$.
  - The general solution of $A\mathbf{x} = \mathbf{b}$ is the sum of any particular solution of $A\mathbf{x} = \mathbf{b}$ and the general solution of $A\mathbf{x} = \mathbf{0}$.

Example (General Solution of $A\mathbf{x} = \mathbf{b}$)

- The solution to the nonhomogeneous system

$$\begin{aligned} x_1 + 3x_2 - 2x_3 \quad + 2x_5 \quad &= 0 \\ 2x_1 + 6x_2 - 5x_3 - 2x_4 + 4x_5 - 3x_6 &= -1 \\ 5x_3 + 10x_4 \quad + 15x_6 &= 5 \\ 2x_1 + 5x_2 \quad + 8x_4 + 4x_5 + 18x_6 &= 6 \end{aligned}$$

is

$x_1 = -3r - 4s - 2t$, $x_2 = r$,
$x_3 = -2s$, $x_4 = s$,
$x_5 = t$, $x_6 = 1/3$

- The result can be written in vector form as

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix} = \begin{bmatrix} -3r-4s-2t \\ r \\ -2s \\ s \\ t \\ 1/3 \end{bmatrix} = \underbrace{\overbrace{\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1/3 \end{bmatrix}}^{\boldsymbol{x}_0}}_{\mathbf{x}_0} + \underbrace{\overbrace{r\begin{bmatrix} -3 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + s\begin{bmatrix} -4 \\ 0 \\ -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t\begin{bmatrix} -2 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}}^{\boldsymbol{x}}}_{\mathbf{x}}$$

which is the general solution.

- The vector $\mathbf{x}_0$ is a **particular solution** of nonhomogeneous system, and the linear combination $\mathbf{x}$ is the general solution of the homogeneous system.

Elementary Row Operation

- Performing an elementary row operation on an augmented matrix does not change the solution set of the corresponding linear system.
- It follows that applying an elementary row operation to a matrix A does not change the solution set of the corresponding linear system $A\mathbf{x}=\mathbf{0}$, or stated another way, it does not change the null space of A.

The solution space of the homogeneous system of equation $A\mathbf{x} = \mathbf{0}$, which is a subspace of R^n, is called the null space of A.

Example

- Find a basis for the nullspace of $A = \begin{bmatrix} 2 & 2 & -1 & 0 & 1 \\ -1 & -1 & 2 & -3 & 1 \\ 1 & 1 & -2 & 0 & -1 \\ 0 & 0 & 1 & 1 & 1 \end{bmatrix}$

- Solution
  - The nullspace of *A* is the solution space of the homogeneous system

$$\begin{aligned} 2x_1 + 2x_2 - x_3 + x_5 &= 0 \\ -x_1 - x_2 - 2x_3 - 3x_4 + x_5 &= 0 \\ x_1 + x_2 - 2x_3 - x_5 &= 0 \\ x_3 + x_4 + x_5 &= 0 \end{aligned}$$

  - In Example 10 of Section 4.5 we showed that the vectors

$$\mathbf{v}_1 = \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \text{ and } \mathbf{v}_2 = \begin{bmatrix} -1 \\ 0 \\ -1 \\ 0 \\ 1 \end{bmatrix}$$

form a basis for the nullspace.

Theorems 4.7.3 and 4.7.4

- Theorem 4.7.3
  - Elementary row operations do not change the nullspace of a matrix.
- Theorem 4.7.4
  - Elementary row operations do not change the row space of a matrix.

Proof of Theorem 4.7.4

- Suppose that the row vectors of a matrix A are $\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_m$, and let B be obtained from A by performing an elementary row operation. (We say that A and B are row equivalent.)
- We shall show that every vector in the row space of B is also in that of A, and that every vector in the row space of A is in that of B.
- If the row operation is a row interchange, then B and A have the same row vectors and consequently have the same row space.

Proof of Theorem 4.7.4

- If the row operation is multiplication of a row by a nonzero scalar or a multiple of one row to another, then the row vector $\mathbf{r}_1',\mathbf{r}_2',\ldots,\mathbf{r}_m'$ of B are linear combination of $\mathbf{r}_1,\mathbf{r}_2,\ldots,\mathbf{r}_m$; thus they lie in the row space of A.
- Since a vector space is closed under addition and scalar multiplication, all linear combination of $\mathbf{r}_1',\mathbf{r}_2',\ldots,\mathbf{r}_m'$ will also lie in the row space of A. Therefore, each vector in the row space of B is in the row space of A.

Proof of Theorem 4.7.4

- Since B is obtained from A by performing a row operation, A can be obtained from B by performing the inverse operation (Sec. 1.5).
- Thus the argument above shows that the row space of A is contained in the row space of B.

Remarks

- Do elementary row operations change the column space?
  - Yes!
- The second column is a scalar multiple of the first, so the column space of A consists of all scalar multiplies of the first column vector.

$$A = \begin{bmatrix} 1 & 3 \\ 2 & 6 \end{bmatrix} \longrightarrow B = \begin{bmatrix} 1 & 3 \\ 0 & 0 \end{bmatrix}$$

Add -2 times the first row to the second

- Again, the second column is a scalar multiple of the first, so the column space of B consists of all scalar multiples of the first column vector. This is not the same as the column space of A.

Theorem 4.7.5

- Theorem 4.7.5
  - If a matrix *R* is in row echelon form, then the row vectors with the leading 1's (i.e., the nonzero row vectors) form a basis for the row space of *R*, and the column vectors with the leading 1's of the row vectors form a basis for the column space of *R*.

Bases for Row and Column Spaces

The matrix

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

is in row-echelon form. From Theorem 5.5.6 the vectors

$$\mathbf{r}_1 = [1\ \text{-}2\ 5\ 0\ 3]$$

$$\mathbf{r}_2 = [0\ 1\ 3\ 0\ 0]$$

$$\mathbf{r}_3 = [0\ 0\ 0\ 1\ 0]$$

form a basis for the row space of R, and the vectors

$$\mathbf{c}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{c}_2 = \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{c}_4 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

form a basis for the column space of R.

Example

- Find bases for the row and column spaces of

$$A = \begin{bmatrix} 1 & -3 & 4 & -2 & 5 & 4 \\ 2 & -6 & 9 & -1 & 8 & 2 \\ 2 & -6 & 9 & -1 & 9 & 7 \\ -1 & 3 & -4 & 2 & -5 & -4 \end{bmatrix}$$

- Solution:
  - Since elementary row operations do not change the row space of a matrix, we can find a basis for the row space of *A* by finding a basis that of any row-echelon form of A.
  - Reducing *A* to row-echelon form we obtain

$$R = \begin{bmatrix} 1 & -3 & 4 & -2 & 5 & 4 \\ 0 & 0 & 1 & 3 & -2 & -6 \\ 0 & 0 & 0 & 0 & 1 & 5 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Example

$$A = \begin{bmatrix} 1 & -3 & 4 & -2 & 5 & 4 \\ 2 & -6 & 9 & -1 & 8 & 2 \\ 2 & -6 & 9 & -1 & 9 & 7 \\ -1 & 3 & -4 & 2 & -5 & -4 \end{bmatrix} \qquad R = \begin{bmatrix} 1 & -3 & 4 & -2 & 5 & 4 \\ 0 & 0 & 1 & 3 & -2 & -6 \\ 0 & 0 & 0 & 0 & 1 & 5 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

- The basis vectors for the row space of *R* and *A*

  $\mathbf{r}_1$ = [1 -3 4 -2 5 4]

  $\mathbf{r}_2$ = [0 0 1 3 -2 -6]

  $\mathbf{r}_3$ = [0 0 0 0 1 5]

- Keeping in mind that *A* and *R* may have different column spaces, we cannot find a basis for the column space of *A* directly from the column vectors of *R*.

Theorem 4.7.6

- Theorem 4.7.6
  - If *A* and *B* are row equivalent matrices, then:
    - A given set of column vectors of *A* is linearly independent if and only if the corresponding column vectors of *B* are linearly independent.
    - A given set of column vectors of *A* forms a basis for the column space of *A* if and only if the corresponding column vectors of *B* form a basis for the column space of *B*.

Example

$$A = \begin{bmatrix} 1 & -3 & 4 & -2 & 5 & 4 \\ 2 & -6 & 9 & -1 & 8 & 2 \\ 2 & -6 & 9 & -1 & 9 & 7 \\ -1 & 3 & -4 & 2 & -5 & -4 \end{bmatrix} \qquad R = \begin{bmatrix} 1 & -3 & 4 & -2 & 5 & 4 \\ 0 & 0 & 1 & 3 & -2 & -6 \\ 0 & 0 & 0 & 0 & 1 & 5 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

- We can find the basis for the column space of R, then *the corresponding column vectors* of A will form a basis for the column space of A.
- Basis for R's column space

$$\mathbf{c}'_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \qquad \mathbf{c}'_3 = \begin{bmatrix} 4 \\ 1 \\ 0 \\ 0 \end{bmatrix} \qquad \mathbf{c}'_5 = \begin{bmatrix} 5 \\ -2 \\ 1 \\ 0 \end{bmatrix}$$

- Basis for A's column space

$$\mathbf{c}_1 = \begin{bmatrix} 1 \\ 2 \\ 2 \\ -1 \end{bmatrix}, \ \mathbf{c}_3 = \begin{bmatrix} 4 \\ 9 \\ 9 \\ -4 \end{bmatrix}, \ \mathbf{c}_5 = \begin{bmatrix} 5 \\ 8 \\ 9 \\ -5 \end{bmatrix}$$

Example (Basis for a Vector Space Using Row Operations)

- Find a basis for the space spanned by the row vectors

$$\mathbf{v}_1 = (1, -2, 0, 0, 3),\ \mathbf{v}_2 = (2, -5, -3, -2, 6),$$
$$\mathbf{v}_3 = (0, 5, 15, 10, 0),\ \mathbf{v}_4 = (2, 6, 18, 8, 6).$$

- Except for a variation in notation, the space spanned by these vectors is the row space of the matrix

$$\begin{bmatrix} 1 & -2 & 0 & 0 & 3 \\ 2 & -5 & -3 & -2 & 6 \\ 0 & 5 & 15 & 10 & 0 \\ 2 & 6 & 18 & 8 & 6 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & -2 & 0 & 0 & 3 \\ 0 & 1 & 3 & 2 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

  - The nonzero row vectors in this matrix are

$$\mathbf{w}_1 = (1, -2, 0, 0, 3),\ \mathbf{w}_2 = (0, 1, 3, 2, 0),\ \mathbf{w}_3 = (0, 0, 1, 1, 0)$$

  - These vectors form a basis for the row space and consequently form a basis for the subspace of R^5 spanned by $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$, and $\mathbf{v}_4$.

Remarks

- Keeping in mind that *A* and *R* may have different column spaces, we cannot find a basis for the column space of *A directly* from the column vectors of *R*.
- However, if we can find a set of column vectors of *R* that forms a basis for the column space of *R*, then the *corresponding* column vectors of *A* will form a basis for the column space of *A*.

- **The basis vectors obtained for the column space of *A* consisted of column vectors of *A*, but the basis vectors obtained for the row space of *A* were not all vectors of *A*.**
- **Transpose of the matrix can be used to solve this problem**.

Example (Basis for the Row Space of a Matrix)

- Find a basis for the row space of

$$A=\begin{bmatrix}1 & -2 & 0 & 0 & 3\\ 2 & -5 & -3 & -2 & 6\\ 0 & 5 & 15 & 10 & 0\\ 2 & 6 & 18 & 8 & 6\end{bmatrix}$$

consisting entirely of row vectors from A.

- Solution:

$$A^T=\begin{bmatrix}1 & 2 & 0 & 2\\ -2 & -5 & 5 & 6\\ 0 & -3 & 15 & 18\\ 0 & -2 & 10 & 8\\ 3 & 6 & 0 & 6\end{bmatrix} \Rightarrow \begin{bmatrix}1 & 2 & 0 & 2\\ 0 & 1 & 5 & -10\\ 0 & 0 & 0 & 1\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\end{bmatrix}$$

- The column space of A^T are

$$\mathbf{c}_1=\begin{bmatrix}1\\ -2\\ 0\\ 0\\ 3\end{bmatrix},\ \mathbf{c}_2=\begin{bmatrix}2\\ -5\\ -3\\ -2\\ 6\end{bmatrix},\ \text{and } \mathbf{c}_4=\begin{bmatrix}2\\ 6\\ 18\\ 8\\ 6\end{bmatrix}$$

- Thus, the basis vectors for the row space of A are

  $\mathbf{r}_1$ = [1 -2 0 0 3]

  $\mathbf{r}_2$ = [2 -5 -3 -2 6]

  $\mathbf{r}_3$ = [2 6 18 8 6]

Example (Basis and Linear Combinations)

- (a) Find a subset of the vectors $\mathbf{v}_1$ = (1, -2, 0, 3), $\mathbf{v}_2$ = (2, -5, -3, 6), $\mathbf{v}_3$ = (0, 1, 3, 0), $\mathbf{v}_4$ = (2, -1, 4, -7), $\mathbf{v}_5$ = (5, -8, 1, 2) that forms a basis for the space spanned by these vectors.
- (b) Express each vector not in the basis as a linear combination of the basis vectors.

- Solution (a):

$$\begin{bmatrix} 1 & 2 & 0 & 2 & 5 \\ -2 & -5 & 1 & -1 & -8 \\ 0 & -3 & 3 & 4 & 1 \\ 3 & 6 & 0 & -7 & 2 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 0 & 2 & 0 & 1 \\ 0 & 1 & -1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\begin{matrix} \uparrow & \uparrow & \uparrow & \uparrow & \uparrow \\ \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 & \mathbf{v}_4 & \mathbf{v}_5 \end{matrix} \qquad \begin{matrix} \uparrow & \uparrow & \uparrow & \uparrow & \uparrow \\ \mathbf{w}_1 & \mathbf{w}_2 & \mathbf{w}_3 & \mathbf{w}_4 & \mathbf{w}_5 \end{matrix}$$

  - Thus, $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_4\}$ is a basis for the column space of the matrix.

Example

- Solution (b):
  - We can express $\mathbf{w}_3$ as a linear combination of $\mathbf{w}_1$ and $\mathbf{w}_2$, express $\mathbf{w}_5$ as a linear combination of $\mathbf{w}_1$, $\mathbf{w}_2$, and $\mathbf{w}_4$ (Why?). By inspection, these linear combination are

$$\mathbf{w}_3 = 2\mathbf{w}_1 - \mathbf{w}_2$$

$$\mathbf{w}_5 = \mathbf{w}_1 + \mathbf{w}_2 + \mathbf{w}_4$$

  - We call these the dependency equations. The corresponding relationships in the original vectors are

$$\mathbf{v}_3 = 2\mathbf{v}_1 - \mathbf{v}_2$$

$$\mathbf{v}_5 = \mathbf{v}_1 + \mathbf{v}_2 + \mathbf{v}_4$$

4.8
Rank, Nullity, and the Fundamental Matrix Spaces

Dimension and Rank

- Theorem 4.8.1
  - If A is any matrix, then the row space and column space of A have the same dimension.
- Proof: Let R be any row-echelon form of A. It follows from Theorem 4.7.4 and 4.7.6b that

$$\dim(\text{row space of } A) = \dim(\text{row space of } R).$$

$$\dim(\text{column space of } A) = \dim(\text{column space of } R)$$

- The dimension of the row space of R is the number of nonzero rows = number of leading 1's = dimension of the column space of R

Rank and Nullity

- Definition
  - The common dimension of the row and column space of a matrix *A* is called the rank (秩) of *A* and is denoted by rank(*A*); the dimension of the nullspace of a is called the nullity (零核維數) of *A* and is denoted by nullity(*A*).

Example (Rank and Nullity)

- Find the rank and nullity of the matrix

$$A=\begin{bmatrix} -1 & 2 & 0 & 4 & 5 & -3 \\ 3 & -7 & 2 & 0 & 1 & 4 \\ 2 & -5 & 2 & 4 & 6 & 1 \\ 4 & -9 & 2 & -4 & -4 & 7 \end{bmatrix}$$

- Solution:
  - The reduced row-echelon form of A is

$$\begin{bmatrix} 1 & 0 & -4 & -28 & -37 & 13 \\ 0 & 1 & -2 & -12 & -16 & 5 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

  - Since there are two nonzero rows (two leading 1's), the row space and column space are both two-dimensional, so rank(A) = 2.

Example (Rank and Nullity)

- To find the nullity of A, we must find the dimension of the solution space of the linear system $A\mathbf{x}=\mathbf{0}$.
- The corresponding system of equations will be

$$x_1 - 4x_3 - 28x_4 - 37x_5 + 13x_6 = 0$$
$$x_2 - 2x_3 - 12x_4 - 16x_5 + 5x_6 = 0$$

- It follows that the general solution of the system is

$$x_1 = 4r + 28s + 37t - 13u,\ x_2 = 2r + 12s + 16t - 5u,$$
$$x_3 = r,\ x_4 = s,\ x_5 = t,\ x_6 = u$$

or

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix} = r\begin{bmatrix} 4 \\ 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + s\begin{bmatrix} 28 \\ 12 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t\begin{bmatrix} 37 \\ 16 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} + u\begin{bmatrix} -13 \\ -5 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

Thus, nullity(A) = 4.

Example

- What is the maximum possible rank of an $m \times n$ matrix *A* that is not square?
- Solution: The row space of *A* is at most *n*-dimensional and the column space is at most *m*-dimensional. Since the rank of A is the common dimension of its row and column space, it follows that the rank is at most the smaller of *m* and *n*.

$$rank(A) \leq \min(m, n)$$

Theorem 4.8.2

- Theorem 4.8.2 (Dimension Theorem for Matrices)
  - If A is a matrix with n columns, then rank(A) + nullity(A) = n.
- Proof:
- Since A has n columns, $A\mathbf{x} = \mathbf{0}$ has n unknowns. These fall into two categories: the leading variables and the free variables.

$$\begin{bmatrix} \text{number of} \\ \text{leading variables} \end{bmatrix} + \begin{bmatrix} \text{number of} \\ \text{free variables} \end{bmatrix} = n$$

- The number of leading 1's in the reduced row-echelon form of A is the rank of A

$$\text{rank}(A) + \begin{bmatrix} \text{number of} \\ \text{free variables} \end{bmatrix} = n$$

Theorem 4.8.2

- The number of free variables is equal to the nullity of A. This is so because the nullity of A is the dimension of the solution space of $A\mathbf{x}=\mathbf{0}$, which is the same as the number of parameters in the general solution, which is the same as the number of free variables. Thus

$$\text{rank}(A) + \text{nullity}(A) = n$$

Example

$$A = \begin{bmatrix} -1 & 2 & 0 & 4 & 5 & -3 \\ 3 & -7 & 2 & 0 & 1 & 4 \\ 2 & -5 & 2 & 4 & 6 & 1 \\ 4 & -9 & 2 & -4 & -4 & 7 \end{bmatrix}$$

- This matrix has 6 columns, so rank(A) + nullity(A) = 6
- In previous example, we know rank(A) = 4 and nullity(A) = 2

Theorem 4.8.3

- Theorem 4.8.3
  - If A is an $m \times n$ matrix, then:
    - rank(A) = Number of leading variables in the solution of $A\mathbf{x} = \mathbf{0}$.
    - nullity(A) = Number of parameters in the general solution of $A\mathbf{x} = \mathbf{0}$.

Example

- Find the number of parameters in the general solution of $A\mathbf{x} = \mathbf{0}$ if A is a 5×7 matrix of rank 3.
- Solution:
  - nullity(A) = n – rank(A) = 7 – 3 = 4
  - Thus, there are four parameters.

Theorem 4.8.4 (Equivalent Statements)

- If A is an $n\times n$ matrix, and if $T_A : R^n \to R^n$ is multiplication by A, then the following are equivalent:
  - A is invertible.
  - $A\mathbf{x} = \mathbf{0}$ has only the trivial solution.
  - The reduced row-echelon form of A is I_n.
  - A is expressible as a product of elementary matrices.
  - $A\mathbf{x} = \mathbf{b}$ is consistent for every $n\times 1$ matrix $\mathbf{b}$.
  - $A\mathbf{x} = \mathbf{b}$ has exactly one solution for every $n\times 1$ matrix $\mathbf{b}$.
  - $\det(A)\neq 0$.
  - The column vectors of A are linearly independent.
  - The row vectors of A are linearly independent.
  - The column vectors of A span R^n.
  - The row vectors of A span R^n.
  - The column vectors of A form a basis for R^n.
  - The row vectors of A form a basis for R^n.
  - A has rank n.
  - A has nullity 0.

Overdetermined System

- A linear system with more equations than unknowns is called an overdetermined linear system (超定線性方程組). With fewer unknowns than equations, it's called an underdetermined system.
- Theorem 4.8.5
  - If $A\mathbf{x} = \mathbf{b}$ is a consistent linear system of m equations in n unknowns, and if A has rank r, then the general solution of the system contains $n - r$ parameters.

- If A is a 5×7 matrix with rank 4, and if $A\mathbf{x}=\mathbf{b}$ is a consistent linear system, then the general solution of the system contains 7-4=3 parameters.

Theorem 4.8.6

- Let A be an $m \times n$ matrix
- (a) (Overdetemined Case) If $m > n$, then the linear system $A\mathbf{x}=\mathbf{b}$ is inconsistent for at least one vector **b** in R^n.
- (b) (Underdetermined Case) If $m < n$, then for each vector **b** in R^m the linear system $A\mathbf{x}=\mathbf{b}$ is either inconsistent or has infinitely many solutions.

Proof of Theorem 4.8.6 (a)

- Assume that $m>n$, in which case the column vectors of A cannot span R^m (fewer vectors than the dimension of R^m). Thus, there is at least one vector **b** in R^m that is not in the column space of A, and for that **b** the system $A\mathbf{x}=\mathbf{b}$ is inconsistent by Theorem 4.7.1.

Proof of Theorem 4.8.6 (b)

- Assume that $m<n$. For each vector **b** in R^n there are two possibilities: either the system $A\mathbf{x}=\mathbf{b}$ is consistent or it is inconsistent.
- If it is inconsistent, then the proof is complete.
- If it is consistent, then Theorem 4.8.5 implies that the general solution has n-r parameters, where r=rank(A).
- But rank(A) is the smaller of m and n, so $n-r = n-m > 0$
- This means that the general solution has at least one parameter and hence there are infinitely many solutions.

Example

- What can you say about the solutions of an overdetermined system $A\mathbf{x}=\mathbf{b}$ of 7 equations in 5 unknowns in which A has rank = 4?
- What can you say about the solutions of an underdetermined system $A\mathbf{x}=\mathbf{b}$ of 5 equations in 7 unknowns in which A has rank = 4?
- Solution:
  - (a) the system is consistent for some vector **b** in R^7, and for any such **b** the number of parameters in the general solution is n-r=5-4=1
  - (b) the system may be consistent or inconsistent, but if it is consistent for the vector **b** in R^5, then the general solution has n-r=7-4=3 parameters.

Example

- The linear system

$$\begin{aligned} x_1 - 2x_2 &= b_1 \\ x_1 - x_2 &= b_2 \\ x_1 + x_2 &= b_3 \\ x_1 + 2x_2 &= b_4 \\ x_1 + 3x_2 &= b_5 \end{aligned}$$

is overdetermined, so it cannot be consistent for all possible values of b_1, b_2, b_3, b_4, and b_5. Exact conditions under which the system is consistent can be obtained by solving the linear system by Gauss-Jordan elimination.

$$\begin{bmatrix} 1 & 0 & 2b_2 - b_1 \\ 0 & 1 & b_2 - b_1 \\ 0 & 0 & b_3 - 3b_2 + 2b_1 \\ 0 & 0 & b_4 - 4b_2 + 3b_1 \\ 0 & 0 & b_5 - 5b_2 + 4b_1 \end{bmatrix}$$

Example

- Thus, the system is consistent if and only if b_1, b_2, b_3, b_4, and b_5 satisfy the conditions

$$\begin{aligned} 2b_1 - 3b_2 + b_3 \quad\quad\quad\quad &= 0 \\ 2b_1 - 4b_2 \quad\quad + b_4 \quad\quad &= 0 \\ 4b_1 - 5b_2 \quad\quad\quad\quad + b_5 &= 0 \end{aligned}$$

or, on solving this homogeneous linear system, $b_1=5r-4s$, $b_2=4r-3s$, $b_3=2r-s$, $b_4=r$, $b_5=s$ where r and s are arbitrary.

Fundamental Spaces of a Matrix

- Six important vector spaces associated with a matrix A
- **Row space of A**, row space of A^T
- **Column space of A**, column space of A^T
- **Null space of A, null space of A^T**
- Transposing a matrix converts row vectors into column vectors
  - Row space of A^T = column space of A
  - Column space of A^T = row space of A
- These are called the fundamental spaces of a matrix A

Theorem 4.8.7

- if A is any matrix, then rank(A) = rank(A^T)
- Proof:
  - Rank(A) = dim(row space of A) = dim(column space of A^T) = rank(A^T)
- If A is an $m \times n$ matrix, then rank(A)+nullity(A)=n. rank(A^T)+nullity(A^T) = m
- The dimensions of fundamental spaces

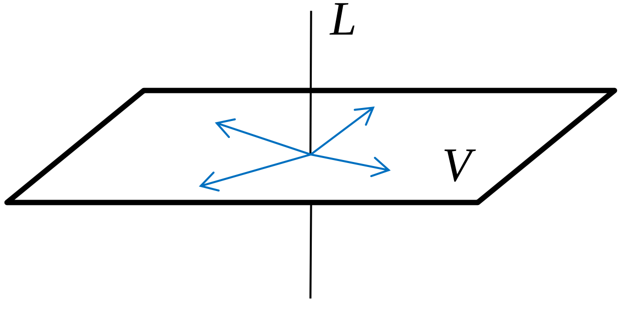
Fundamental Space	Dimension
Row space of A	r
Column space of A	r
Nullspace of A	$n - r$
Nullspace of A^T	$m - r$

Recap

- Theorem 3.4.3: If A is an $m \times n$ matrix, then the solution set of the homogeneous linear system $A\mathbf{x}=\mathbf{0}$ consists of all vectors in R^n that are orthogonal to every row vector of A.
- The null space of A consists of those vectors that are orthogonal to each of the row vectors of A.

Orthogonality

- Definition
  - Let W be a subspace of R^n, the set of all vectors in R^n that are orthogonal to every vector in W is called the orthogonal complement (正交補餘) of W, and is denoted by $W^{\perp}$
  - If V is a plane through the origin of R^3 with Euclidean inner product, then **the set of all vectors that are orthogonal to every vector in V forms the line L through the origin** that is perpendicular to V.

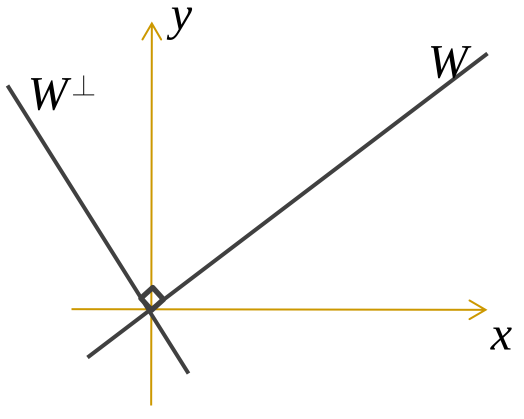


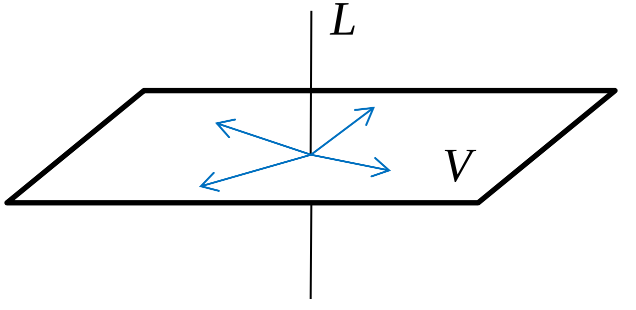
Theorem 4.8.8

- Theorem 4.8.8
- If W is a subspace of a finite-dimensional space R^n, then:
  - $W^{\perp}$ is a subspace of R^n. (read "W perp")
  - The only vector common to W and $W^{\perp}$ is **0**; that is ,$W \cap W^{\perp} = \mathbf{0}$.
  - The orthogonal complement of $W^{\perp}$ is W; that is , $(W^{\perp})^{\perp} = W$.

Example

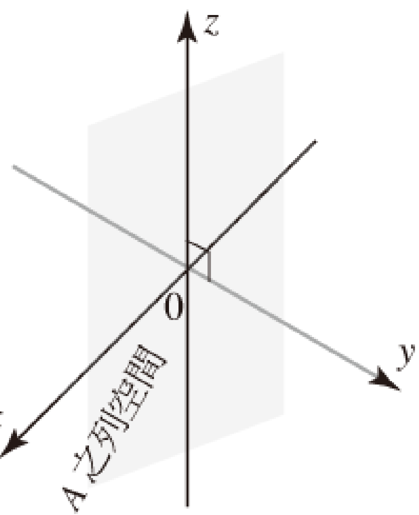
- Orthogonal complements

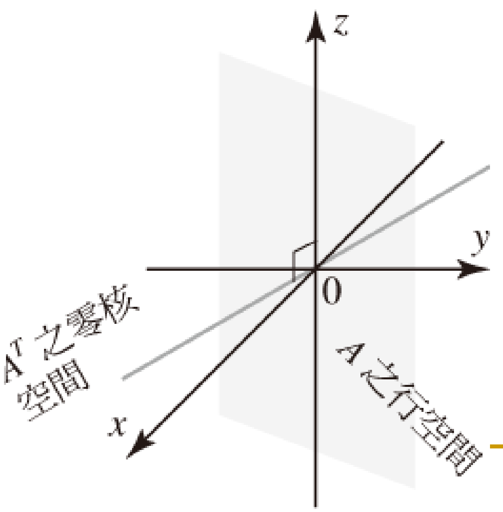




Theorem 4.8.9

- Theorem 4.8.9
  - If A is an $m \times n$ matrix, then:
    - The null space of A and the row space of A are orthogonal complements in R^n.
    - The null space of A^T and the column space of A are orthogonal complements in R^m.





Theorem 4.8.10 (Equivalent Statements)

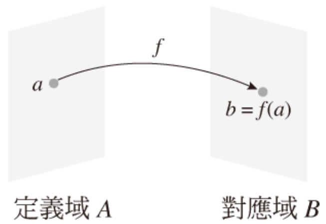
- If A is an $m \times n$ matrix, and if $T_A : R^n \to R^n$ is multiplication by A, then the following are equivalent:
  - A is invertible.
  - $A\mathbf{x} = \mathbf{0}$ has only the trivial solution.
  - The reduced row-echelon form of A is I_n.
  - A is expressible as a product of elementary matrices.
  - $A\mathbf{x} = \mathbf{b}$ is consistent for every $n \times 1$ matrix $\mathbf{b}$.
  - $A\mathbf{x} = \mathbf{b}$ has exactly one solution for every $n \times 1$ matrix $\mathbf{b}$.
  - $\det(A) \neq 0$.
  - The column vectors of A are linearly independent.
  - The row vectors of A are linearly independent.
  - The column vectors of A span R^n.
  - The row vectors of A span R^n.
  - The column vectors of A form a basis for R^n.
  - The row vectors of A form a basis for R^n.
  - A has rank n.
  - A has nullity 0.
  - The orthogonal complement of the nullspace of A is R^n.
  - The orthogonal complement of the row space of A is $\{\mathbf{0}\}$.

Applications of Rank

- Digital data are commonly stored in matrix form.
- Rank plays a role because it measures the “redundancy” in a matrix.
- If A is an $m \times n$ matrix of rank k, then n-k of the column vectors and m-k of the row vectors can be expressed in terms of k linearly independently column or row vectors.
- The essential idea in many data compression schemes is to approximate the original data set by a data set with smaller rank that conveys nearly the same information.

4.9
Matrix Transformations from R^n to R^m

Functions from R^n to R



- A function is a rule f that associates with each element in a set A one and only one element in a set B.
- If f associates the element a with the element b, then we write $b = f(a)$ and say that b is the image of a under f or that $f(a)$ is the value of f at a.
- The set A is called the domain (定義域) of f and the set B is called the codomain (對應域) of f.
- The subset of the codomain B consisting of all possible values for f as a varies over A is called the range (值域) of f.

Examples

Formula	Example	Classification	Description
$f(x)$	$f(x) = x^2$	Real-valued function of a real variable	Function from R to R
$f(x, y)$	$f(x, y) = x^2 + y^2$	Real-valued function of two real variables	Function from R^2 to R
$f(x, y, z)$	$f(x, y, z) = x^2 + y^2 + z^2$	Real-valued function of three real variables	Function from R^3 to R
$f(x_1, x_2, \ldots, x_n)$	$f(x_1, x_2, \ldots, x_n) = x_1^2 + x_2^2 + \ldots + x_n^2$	Real-valued function of *n* real variables	Function from R^n to R

Function from R^n to R^m

- Suppose $f_1, f_2, \ldots, f_m$ are real-valued functions of n real variables, say

$$w_1 = f_1(x_1, x_2, \ldots, x_n)$$
$$w_2 = f_2(x_1, x_2, \ldots, x_n)$$
$$\cdots$$
$$w_m = f_m(x_1, x_2, \ldots, x_n)$$

These m equations assign a unique point $(w_1, w_2, \ldots, w_m)$ in R^m to each point $(x_1, x_2, \ldots, x_n)$ in R^n and thus define a transformation from R^n to R^m. If we denote this transformation by $T: R^n \to R^m$ then

$$T(x_1, x_2, \ldots, x_n) = (w_1, w_2, \ldots, w_m)$$

Function from R^n to R^m

- If $m = n$ the transformation $T: R^n \to R^m$ is called an operator (運算子) on R^n.

Example: A Transformation from R^2 to R^3

$$w_1 = x_1 + x_2$$
$$w_2 = 3x_1x_2$$
$$w_3 = x_1^2 - x_2^2$$

- Define a transform T: $R^2 \longrightarrow R^3$
- With this transformation, the image of the point (x_1, x_2) is

$$T(x_1, x_2) = (x_1 + x_2, 3x_1x_2, x_1^2 - x_2^2)$$

- Thus, for example, T(1,-2) = (-1, -6, -3)

Linear Transformations from R^n to R^m

- A linear transformation (or a linear operator if $m = n$) $T: R^n \to R^m$ is defined by equations of the form

$$\begin{aligned} w_1 &= a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n \\ w_2 &= a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n \\ &\vdots \\ w_m &= a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n \end{aligned}$$

or

$$\begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_m \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{13} \\ a_{21} & a_{22} & \cdots & a_{23} \\ \vdots & \vdots & & \vdots \\ a_{mn} & a_{mn} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix}$$

or

$$\mathbf{w} = A\mathbf{x}$$

- The matrix $A = [a_{ij}]$ is called the standard matrix for the linear transformation T, and T is called multiplication by A.

Example (Transformation and Linear Transformation)

- The linear transformation $T : R^4 \rightarrow R^3$ defined by the equations

$$w_1 = 2x_1 - 3x_2 + x_3 - 5x_4$$
$$w_2 = 4x_1 + x_2 - 2x_3 + x_4$$
$$w_3 = 5x_1 - x_2 + 4x_3$$

the standard matrix for T (i.e., $\mathbf{w} = A\mathbf{x}$) is $A = \begin{bmatrix} 2 & -3 & 1 & -5 \\ 4 & 1 & -2 & 1 \\ 5 & -1 & 4 & 0 \end{bmatrix}$

$$\begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} = \begin{bmatrix} 2 & -3 & 1 & -5 \\ 4 & 1 & -2 & 1 \\ 5 & -1 & 4 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$$

Notations

- Notations:
  - If it is important to emphasize that A is the standard matrix for T, we denote the linear transformation T: $R^n \rightarrow R^m$ by T_A: $R^n \rightarrow R^m$. Thus,

$$T_A(x) = A\mathbf{x}$$

  - We can also denote the standard matrix for T by the symbol $[T]$, or

$$T(\mathbf{x}) = [T]\mathbf{x}$$

Theorem 4.9.1

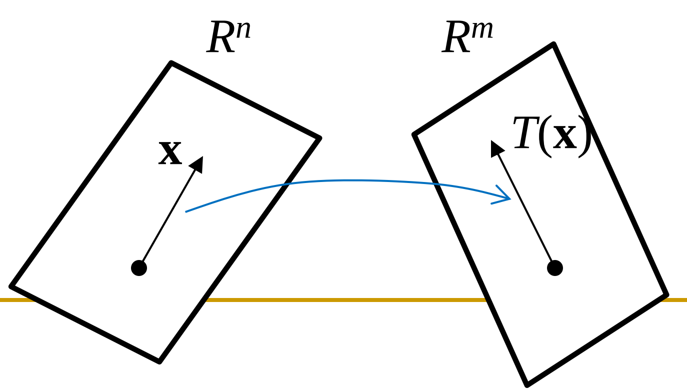
- For every matrix A the matrix transformation $T_A\colon R^n \rightarrow R^m$ has the following properties for all vectors $\mathbf{u}$ and $\mathbf{v}$ in R^n and for every scalar k
  - (a) $T_A(\mathbf{0}) = \mathbf{0}$
  - (b) $T_A(k\mathbf{u}) = kT_A(\mathbf{u})$ [Homogeneity property]
  - $T_A(\mathbf{u}+\mathbf{v}) = T_A(\mathbf{u}) + T_A(\mathbf{v})$ [Additivity property]
  - $T_A(\mathbf{u}-\mathbf{v}) = T_A(\mathbf{u}) - T_A(\mathbf{v})$
- Proof: $A\mathbf{0} = \mathbf{0}$, $A(k\mathbf{u}) = k(A\mathbf{u})$, $A(\mathbf{u}+\mathbf{v}) = A\mathbf{u} + A\mathbf{v}$, $A(\mathbf{u}-\mathbf{v})=A\mathbf{u}-A\mathbf{v}$

Remark

- A matrix transformation maps linear combinations of vectors in R^n into the corresponding linear combinations in R^m in the sense that

$$T_A(k_1\mathbf{u}_1+k_2\mathbf{u}_2+\ldots+k_r\mathbf{u}_r) = k_1T_A(\mathbf{u}_1)+k_2T_A(\mathbf{u}_2)+\ldots+k_rT_A(\mathbf{u}_r)$$

- Depending on whether n-tuples and m-tuples are regarded as vectors or points, the geometric effect of a matrix transformation $T_A:R^n \to R^m$ is to map each vector (point) in R^n into a vector in R^m



Theorem 4.9.2

- If $T_A: R^n \to R^m$ and $T_B: R^n \to R^m$ are matrix transformations, and if $T_A(\mathbf{x}) = T_B(\mathbf{x})$ for every vector $\mathbf{x}$ in R^n, then $A=B$.
- Proof:
  - To say that $T_A(\mathbf{x}) = T_B(\mathbf{x})$ for every vector $\mathbf{x}$ in R^n is the same as saying that $A\mathbf{x} = B\mathbf{x}$ for every vector $\mathbf{x}$ in R^n.
  - This is true, in particular, if $\mathbf{x}$ is any of the standard basis vectors $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ for R^n; that is $A\mathbf{e}_j = B\mathbf{e}_j$ ($j=1,2,\ldots,n$)
  - Since every entry of $\mathbf{e}_j$ is 0 except for the jth, which is 1, it follows from Theorem 1.3.1 that $A\mathbf{e}_j$ is the jth column of A, and $B\mathbf{e}_j$ is the jth column of B. Therefore, $A = B$.

Zero Transformation

- Zero Transformation from R^n to R^m
  - If 0 is the $m \times n$ zero matrix and $\mathbf{0}$ is the zero vector in R^n, then for every vector $\mathbf{x}$ in R^n

$$T_0(\mathbf{x}) = 0\mathbf{x} = \mathbf{0}$$

  - So multiplication by zero maps every vector in R^n into the zero vector in R^m. We call T_0 the zero transformation from R^n to R^m.

Identity Operator

- Identity Operator on R^n
  - If I is the $n \times n$ identity, then for every vector $\mathbf{x}$ in R^n

$$T_I(\mathbf{x}) = I\mathbf{x} = \mathbf{x}$$

  - So multiplication by I maps every vector in R^n into itself.
  - We call T_I the identity operator on R^n.

A Procedure for Finding Standard Matrices

- To find the standard matrix A for a matrix transformations from R^n to R^m:
- $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ are the standard basis vectors for R^n.
- Suppose that the images of these vectors under the transformation T_A are

$$T_A(\mathbf{e}_1)=A\mathbf{e}_1,\ T_A(\mathbf{e}_2)=A\mathbf{e}_2,\ \ldots,\ T_A(\mathbf{e}_n) = A\mathbf{e}_n$$

- $A\mathbf{e}_j$ is just the jth column of the matrix A, Thus,

$$A = [T] = [T(\mathbf{e}_1) \mid T(\mathbf{e}_2) \mid \ldots \mid T(\mathbf{e}_n)]$$

Reflection Operators

- In general, operators on R^2 and R^3 that map each vector into its symmetric image about some line or plane are called reflection (倒影) operators.
- Such operators are linear.

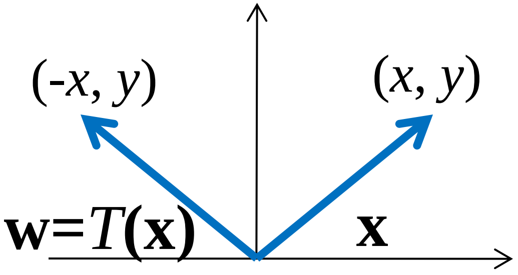
Example

- If we let **w**=T(**x**), then the equations relating the components of **x** and **w** are

$$w_1 = -x = -x + 0y$$
$$w_2 = y = 0x + y$$

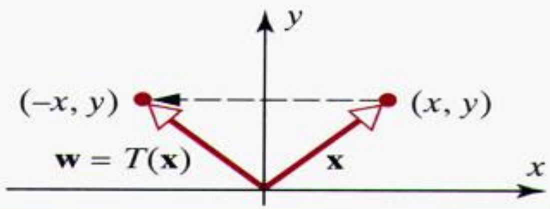
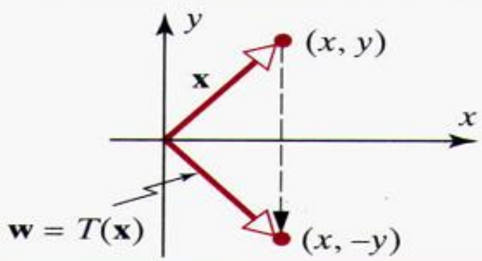
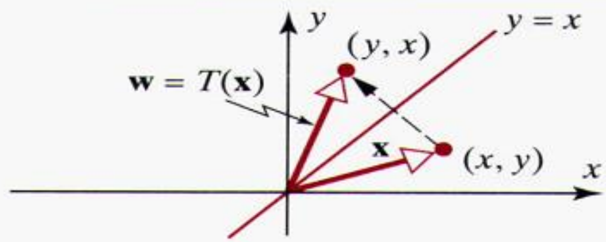
or, in matrix form

$$\begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

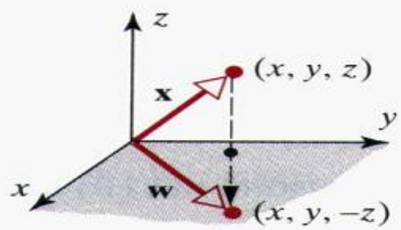
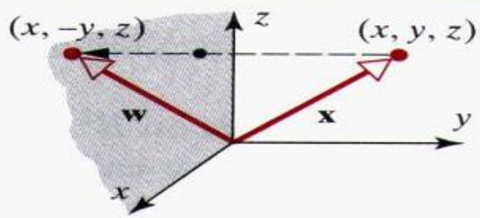
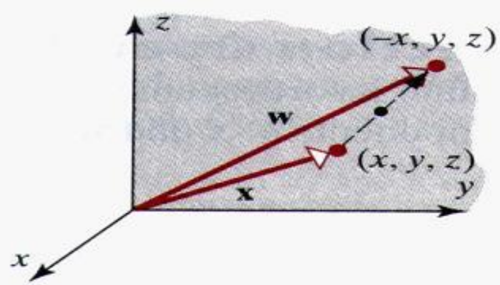


- The standard matrix for T is $\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$

Reflection Operators (2-Space)

Operator	Illustration	Equations	Standard Matrix
Reflection about the y-axis		$w_1 = -x$ $w_2 = y$	$\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$
Reflection about the x-axis		$w_1 = x$ $w_2 = -y$	$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$
Reflection about the line $y = x$		$w_1 = y$ $w_2 = x$	$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$

Reflection Operators (3-Space)

Operator	Illustration	Equations	Standard Matrix
Reflection about the xy-plane		$w_1 = x$ $w_2 = y$ $w_3 = -z$	$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$
Reflection about the xz-plane		$w_1 = x$ $w_2 = -y$ $w_3 = z$	$\begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$
Reflection about the yz-plane		$w_1 = -x$ $w_2 = y$ $w_3 = z$	$\begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

Projection Operators

- In general, a projection operator (or more precisely an orthogonal projection operator) on R^2 or R^3 is any operator that maps each vector into its orthogonal projection on a line or plane through the origin.
- The projection operators are linear.

Example

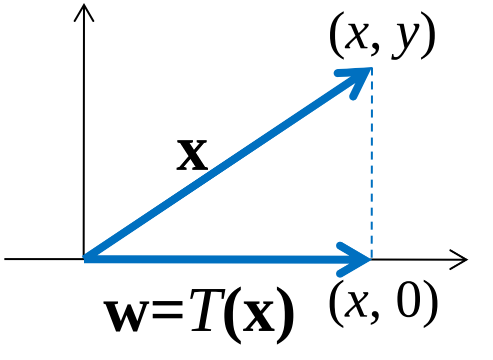
- Consider the operator $T: R^2 \rightarrow R^2$ that maps each vector into its orthogonal projection on the x-axis. The equations relating the components of **x** and **w**=T(**x**) are

$$w_1 = x = x + 0y$$
$$w_2 = 0 = 0x + 0y$$

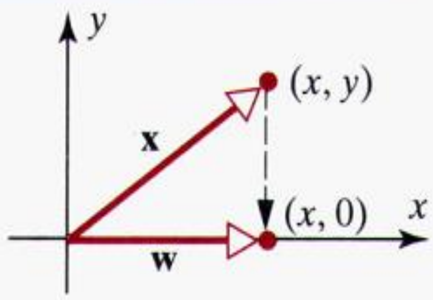
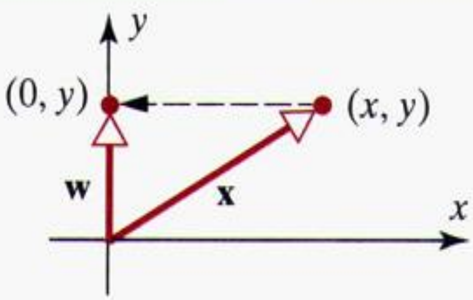
or, in matrix form

$$\begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

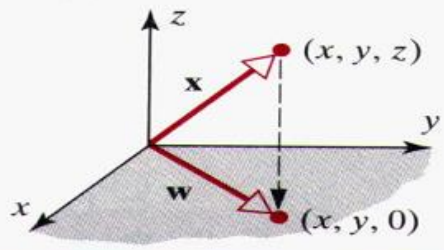
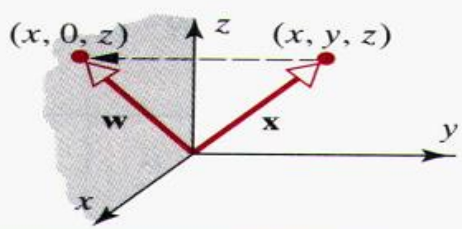
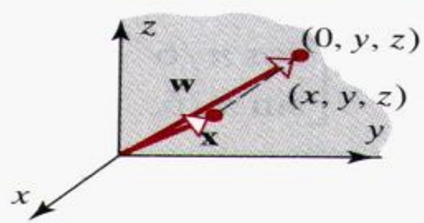
- The standard matrix for T is $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$



Projection Operators

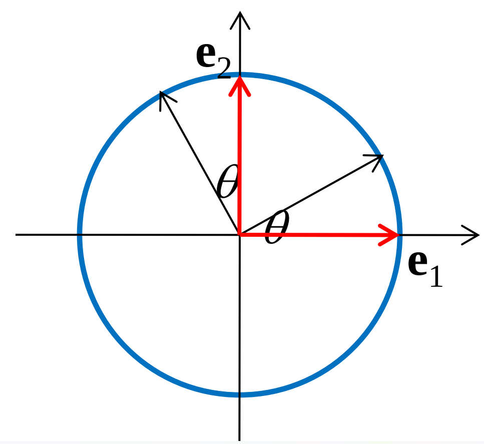
Operator	Illustration	Equations	Standard Matrix
Orthogonal projection on the x-axis		$w_1 = x$ $w_2 = 0$	$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$
Orthogonal projection on the y-axis		$w_1 = 0$ $w_2 = y$	$\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$

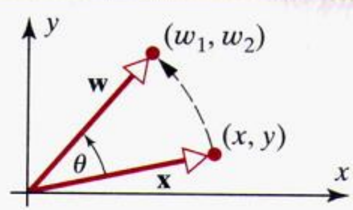
Projection Operators

Operator	Illustration	Equations	Standard Matrix
Orthogonal projection on the xy-plane		$w_1 = x$ $w_2 = y$ $w_3 = 0$	$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$
Orthogonal projection on the xz-plane		$w_1 = x$ $w_2 = 0$ $w_3 = z$	$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$
Orthogonal projection on the yz-plane		$w_1 = 0$ $w_2 = y$ $w_3 = z$	$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

Rotation Operators

- **The rotation operator T:$R^2 \to R^2$ moves points counterclockwise about the origin through an angle θ**
- Find the standard matrix
- $T(\mathbf{e}_1) = T(1,0) = (\cos\theta, \sin\theta)$
- $T(\mathbf{e}_2) = T(0,1) = (-\sin\theta, \cos\theta)$



Operator	Illustration	Equations	Standard Matrix
Rotation through an angle θ		$w_1 = x\cos\theta - y\sin\theta$ $w_2 = x\sin\theta + y\cos\theta$	$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$

Example

- If each vector in R^2 is rotated through an angle of $\pi/6$ (30°) ,then the image **w** of a vector

$$\mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix}$$

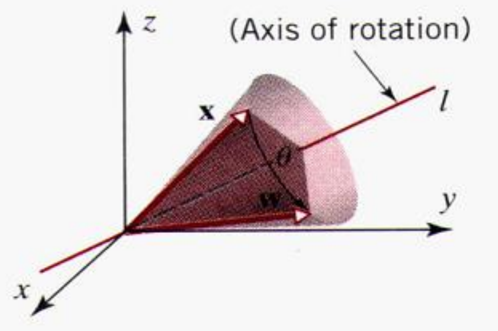
is $$\mathbf{w} = \begin{bmatrix} \cos \pi/6 & -\sin \pi/6 \\ \sin \pi/6 & \cos \pi/6 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \sqrt{3}/2 & -1/2 \\ 1/2 & \sqrt{3}/2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \sqrt{3}/2\, x - 1/2\, y \\ 1/2\, x + \sqrt{3}/2\, y \end{bmatrix}$$

- For example, the image of the vector

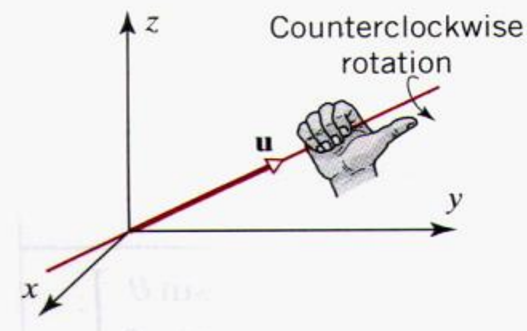
$$\mathbf{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \text{ is } \mathbf{w} = \begin{bmatrix} \dfrac{\sqrt{3}-1}{2} \\ \dfrac{1+\sqrt{3}}{2} \end{bmatrix}$$

A Rotation of Vectors in R^3

- A rotation of vectors in R^3 is usually described in relation to a ray emanating from (發源自) the origin, called the axis of rotation.
- As a vector revolves around the axis of rotation it sweeps out some portion of a cone (圓錐體).
- The angle of rotation is described as "clockwise" or "counterclockwise" in relation to a viewpoint that is along the axis of rotation *looking toward the origin.*
- The axis of rotation can be specified by a nonzero vector **u** that runs along the axis of rotation and has its initial point at the origin.
- The counterclockwise direction for a rotation about its axis can be determined by a "right-hand rule".

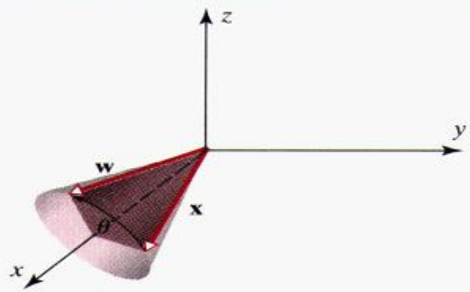
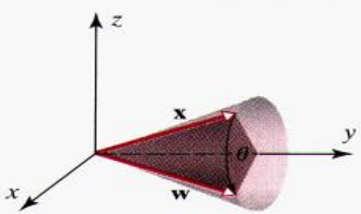
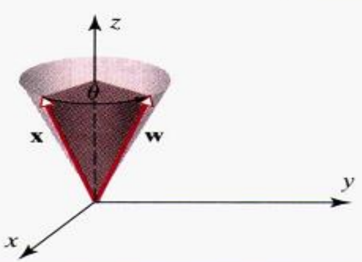


(a) Angle of rotation



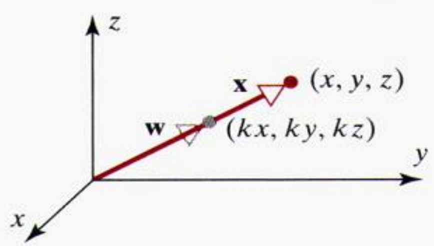
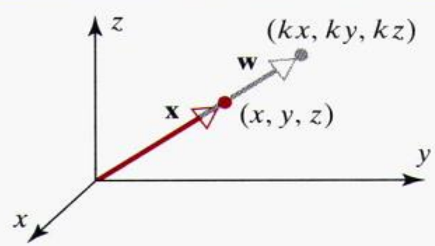
(b) Right-hand rule

A Rotation of Vectors in R^3

Operator	**Illustration**	**Equations**	**Standard Matrix**
Counterclockwise rotation about the positive x-axis through an angle θ		$w_1 = x$ $w_2 = y\cos\theta - z\sin\theta$ $w_3 = y\sin\theta + z\cos\theta$	$\begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix}$
Counterclockwise rotation about the positive y-axis through an angle θ		$w_1 = x\cos\theta + z\sin\theta$ $w_2 = y$ $w_3 = -x\sin\theta + z\cos\theta$	$\begin{bmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{bmatrix}$
Counterclockwise rotation about the positive z-axis through an angle θ		$w_1 = x\cos\theta - y\sin\theta$ $w_2 = x\sin\theta + y\cos\theta$ $w_3 = z$	$\begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$

Dilation and Contraction Operators

- If k is a nonnegative scalar, the operator on R^2 or R^3 is called a contraction with factor k if $0 \leq k \leq 1$ (以因素k收縮) and a dilation with factor k if $k \geq 1$ (以因素k膨脹).

Operator	**Illustration**	**Equations**	**Standard Matrix**
Contraction with factor k on R^3 $(0 \leq k \leq 1)$		$w_1 = kx$ $w_2 = ky$ $w_3 = kz$	$\begin{bmatrix} k & 0 & 0 \\ 0 & k & 0 \\ 0 & 0 & k \end{bmatrix}$
Dilation with factor k on R^3 $(k \geq 1)$		$w_1 = kx$ $w_2 = ky$ $w_3 = kz$	

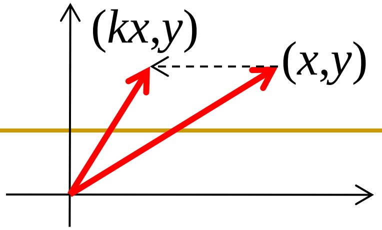
Compression or Expansion

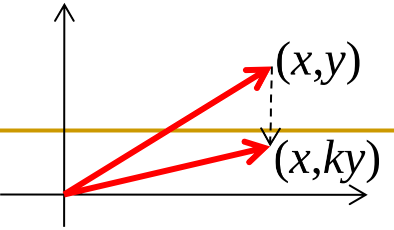
- If $T: R^2 \to R^2$ is a compression ($0<k<1$) or expansion ($k>1$) in the x-direction with factor k, then

$$T(\boldsymbol{e}_1) = T\left(\begin{bmatrix}1\\0\end{bmatrix}\right) = \begin{bmatrix}k\\0\end{bmatrix} \qquad T(\boldsymbol{e}_2) = T\left(\begin{bmatrix}0\\1\end{bmatrix}\right) = \begin{bmatrix}0\\1\end{bmatrix}$$

so the standard matrix for T is $\begin{bmatrix}k & 0\\0 & 1\end{bmatrix}$.

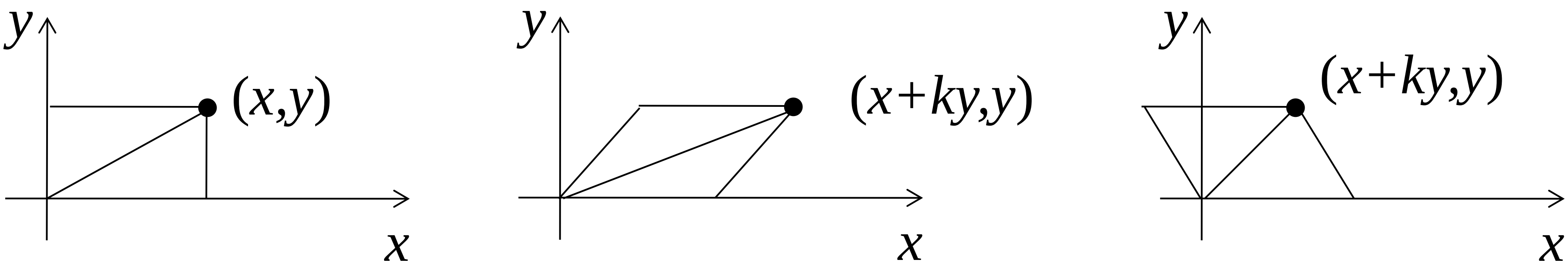
- Similarly, the standard matrix for a compression or expansion in the y-direction is $\begin{bmatrix}1 & 0\\0 & k\end{bmatrix}$





Shears

- **A shear (剪) in the *x*-direction with factor *k*** is a transformation that moves each point (x,y) parallel to the x-axis by an amount ky to the new position $(x+ky,y)$.
- Points farther from the x-axis move a greater distance than those closer.



$k > 0$ $k < 0$

Shears

- If $T: R^2 \longrightarrow R^2$ is a shear with factor k in the x-direction, then

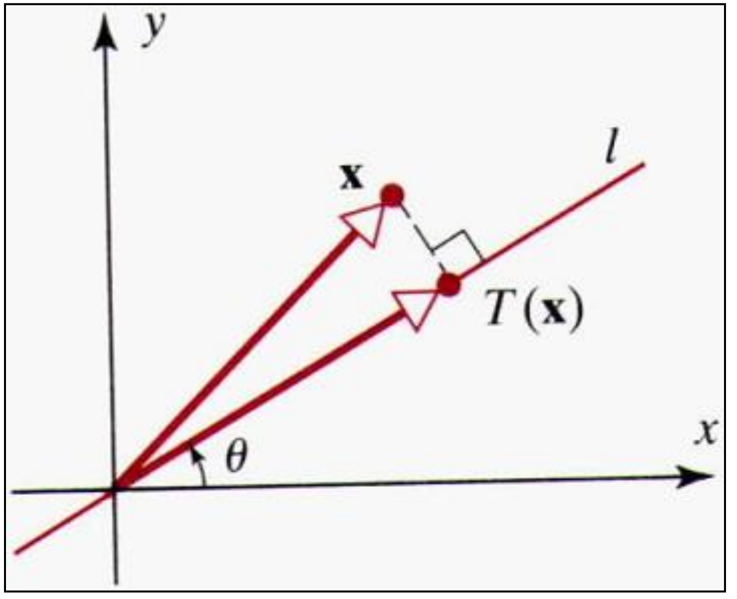
$$T(\boldsymbol{e}_1) = T\left(\begin{bmatrix}1\\0\end{bmatrix}\right) = \begin{bmatrix}x+ky\\y\end{bmatrix} = \begin{bmatrix}1+k0\\0\end{bmatrix} = \begin{bmatrix}1\\0\end{bmatrix}$$

$$T(\boldsymbol{e}_2) = T\left(\begin{bmatrix}0\\1\end{bmatrix}\right) = \begin{bmatrix}x+ky\\y\end{bmatrix} = \begin{bmatrix}0+k1\\1\end{bmatrix} = \begin{bmatrix}k\\1\end{bmatrix}$$

- The standard matrix for T is $\begin{bmatrix}1 & k\\0 & 1\end{bmatrix}$
- Similarly, the standard matrix for a shear in the y-direction with factor k is $\begin{bmatrix}1 & 0\\k & 1\end{bmatrix}$

Example (Standard Matrix for a Projection Operator)

- Let l be the line in the xy-plane that passes through the origin and makes an angle θ with the positive x-axis, where $0 \leq \theta \leq \pi$. Let $T: R^2 \rightarrow R^2$ be a linear operator that maps each vector into orthogonal projection on l.
  - Find the standard matrix for T.
  - Find the orthogonal projection of the vector $\mathbf{x} = (1,5)$ onto the line through the origin that makes an angle of $\theta = \pi/6$ with the positive x-axis.



Example

- The standard matrix for T can be written as

$$[T] = [T(\mathbf{e}_1) \mid T(\mathbf{e}_2)]$$

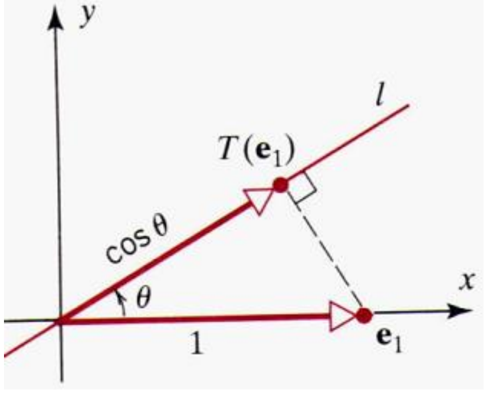
- Consider the case $0 \le \theta \le \pi/2$.
  - $\|T(\mathbf{e}_1)\| = \cos\theta$

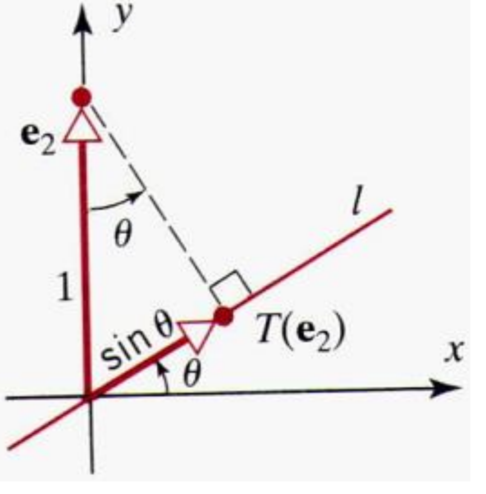
$$\Rightarrow \quad T(\mathbf{e}_1) = \begin{bmatrix} \|T(\mathbf{e}_1)\|\cos\theta \\ \|T(\mathbf{e}_1)\|\sin\theta \end{bmatrix} = \begin{bmatrix} \cos^2\theta \\ \sin\theta\cos\theta \end{bmatrix}$$

  - $\|T(\mathbf{e}_2)\| = \sin\theta$

$$\Rightarrow \quad T(\mathbf{e}_2) = \begin{bmatrix} \|T(\mathbf{e}_2)\|\cos\theta \\ \|T(\mathbf{e}_2)\|\sin\theta \end{bmatrix} = \begin{bmatrix} \sin\theta\cos\theta \\ \sin^2\theta \end{bmatrix}$$

$$\Rightarrow \quad [T] = \begin{bmatrix} \cos^2\theta & \sin\theta\cos\theta \\ \sin\theta\cos\theta & \sin^2\theta \end{bmatrix}$$





Example

$$[T]=\begin{bmatrix}\cos^2\theta & \sin\theta\cos\theta\\ \sin\theta\cos\theta & \sin^2\theta\end{bmatrix}$$

- Since sin (π/6) = 1/2 and cos (π/6) = $\sqrt{3}$/2, it follows from part (a) that the standard matrix for this projection operator is

$$[T]=\begin{bmatrix}3/4 & \sqrt{3}/4\\ \sqrt{3}/4 & 1/4\end{bmatrix}$$

Thus,

$$T\left(\begin{bmatrix}1\\5\end{bmatrix}\right)=\begin{bmatrix}3/4 & \sqrt{3}/4\\ \sqrt{3}/4 & 1/4\end{bmatrix}\begin{bmatrix}1\\5\end{bmatrix}=\begin{bmatrix}\dfrac{3+5\sqrt{3}}{4}\\ \dfrac{\sqrt{3}+5}{4}\end{bmatrix}$$

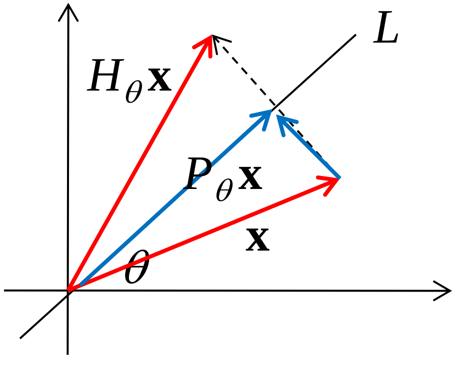
Reflections About Lines Through the Origin

- Let P_θ denote the standard matrix of orthogonal projections on lines through the origin

$P_\theta \mathbf{x} - \mathbf{x} = (1/2)(H_\theta \mathbf{x} - \mathbf{x})$, or equivalently $H_\theta \mathbf{x} = (2P_\theta - I)\mathbf{x}$

- $H_\theta = (2P_\theta - I)$

$$H_\theta = \begin{bmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{bmatrix}$$



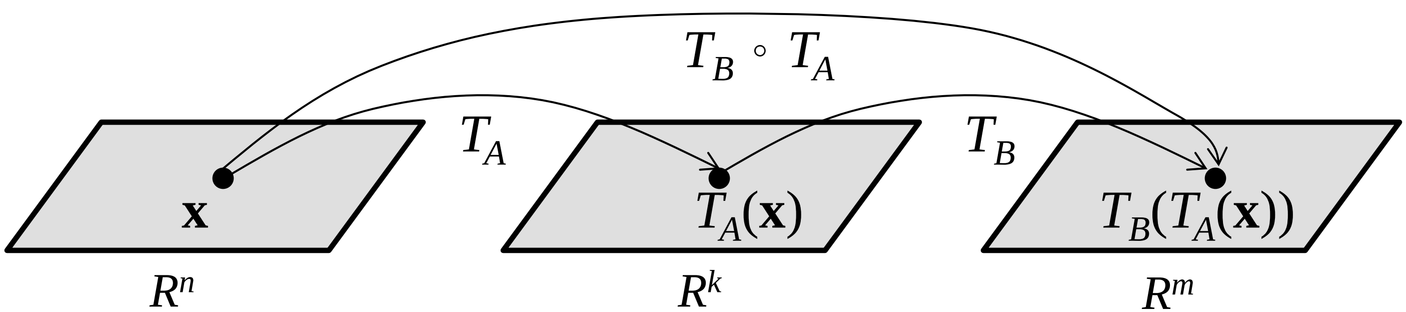
4.10 Properties of Matrix Transformations

Composition of T_B with T_A

- Definition
  - If $T_A : R^n \to R^k$ and $T_B : R^k \to R^m$ are linear transformations, *the composition of T_B with T_A*, denoted by $T_B \circ T_A$ (read "T_B circle T_A"), is the function defined by the formula

$$(T_B \circ T_A)(\mathbf{x}) = T_B(T_A(\mathbf{x}))$$

where $\mathbf{x}$ is a vector in R^n.



Composition of T_B with T_A

- This composition is itself a matrix transformation since

$$(T_B \circ T_A)(\mathbf{x})=(T_B(T_A(\mathbf{x}))=B(T_A(\mathbf{x}))=B(A\mathbf{x})=(BA)\mathbf{x}$$

- It is multiplication by BA, i.e. $T_B \circ T_A = T_{BA}$
- The compositions can be defined for more than two linear transformations.
- For example, if $T_1 : U \rightarrow V$ and $T_2 : V \rightarrow W$,and $T_3 : W \rightarrow Y$ are linear transformations, then the composition $T_3 \circ T_2 \circ T_1$ is defined by $(T_3 \circ T_2 \circ T_1)(\mathbf{u}) = T_3(T_2(T_1(\mathbf{u})))$

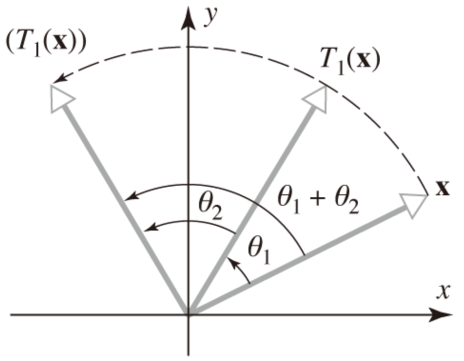
Remark

- It is not true, in general, that $AB = BA$
- So it is not true, in general, that $T_B \circ T_A = T_A \circ T_B$

Example

- Let $T_1 : R^2 \to R^2$ and $T_2 : R^2 \to R^2$ be the matrix operators that rotate vectors through the angles θ_1 and θ_2, respectively.
- The operation $(T_2 \circ T_1)(\mathbf{x}) = T_2(T_1(\mathbf{x}))$ first rotates $\mathbf{x}$ through the angle θ_1, then rotates $T_1(\mathbf{x})$ through the angle θ_2.

$$[T_1] = \begin{bmatrix} \cos\theta_1 & -\sin\theta_1 \\ \sin\theta_1 & \cos\theta_1 \end{bmatrix} \qquad [T_2] = \begin{bmatrix} \cos\theta_2 & -\sin\theta_2 \\ \sin\theta_2 & \cos\theta_2 \end{bmatrix}$$



$$[T_2 \circ T_1] = \begin{bmatrix} \cos(\theta_1+\theta_2) & -\sin(\theta_1+\theta_2) \\ \sin(\theta_1+\theta_2) & \cos(\theta_1+\theta_2) \end{bmatrix}$$

$$[T_2][T_1] = \begin{bmatrix} \cos\theta_2 & -\sin\theta_2 \\ \sin\theta_2 & \cos\theta_2 \end{bmatrix} \begin{bmatrix} \cos\theta_1 & -\sin\theta_1 \\ \sin\theta_1 & \cos\theta_1 \end{bmatrix}$$

$$= \begin{bmatrix} \cos\theta_2\cos\theta_1 - \sin\theta_2\sin\theta_1 & -(\cos\theta_2\sin\theta_1 + \sin\theta_2\cos\theta_1) \\ \sin\theta_2\cos\theta_1 + \cos\theta_2\sin\theta_1 & -\sin\theta_2\sin\theta_1 + \cos\theta_2\cos\theta_1 \end{bmatrix}$$

$$= \begin{bmatrix} \cos(\theta_1+\theta_2) & -\sin(\theta_1+\theta_2) \\ \sin(\theta_1+\theta_2) & \cos(\theta_1+\theta_2) \end{bmatrix} = [T_2 \circ T_1]$$

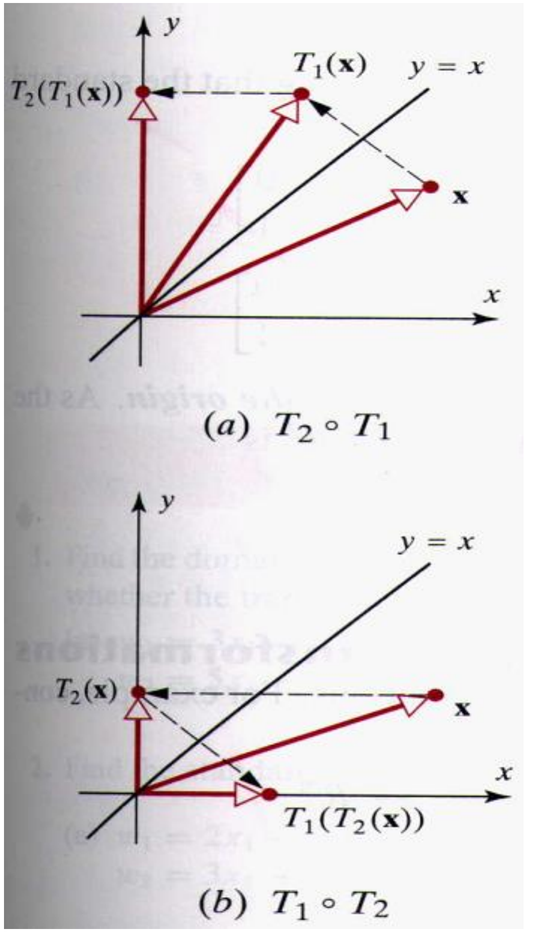
Composition is Not Commutative

- Let T_1 be the reflection operator
- Let T_2 be the orthogonal projection on the y-axis

$$[T_1 \circ T_2] = [T_1][T_2] = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

$$[T_2 \circ T_1] = [T_2][T_1] = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$$

so $[T_1 \circ T_2] \neq [T_2 \circ T_1]$



Composition of Two Reflections

- Let T_1 be the reflection about the y-axis, and let T_2 be the reflection about the x-axis. In this case, $T_1 \circ T_2$ and $T_2 \circ T_1$ are the same.

$$(T_1 \circ T_2)(x, y) = T_1(x, -y) = (-x, -y)$$

$$(T_2 \circ T_1)(x, y) = T_2(-x, y) = (-x, -y)$$

$$[T_1 \circ T_2] = [T_1]\ [T_2] = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$$

$$[T_2 \circ T_1] = [T_2]\ [T_1] = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$$

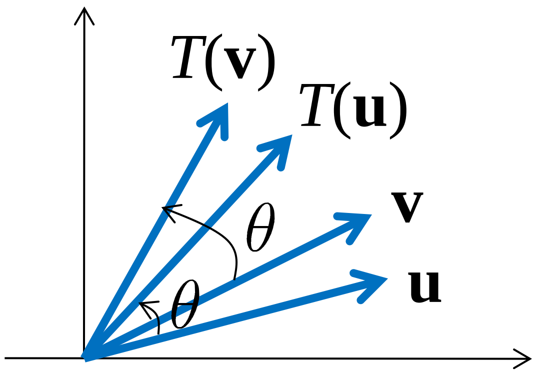
One-to-One Linear transformations

- Definition
  - A linear transformation $T : R^n \rightarrow R^m$ is said to be one-to-one if T maps distinct vectors (points) in R^n into distinct vectors (points) in R^m

- Remark:
  - That is, for each vector **w** in the range of a one-to-one linear transformation T, there is exactly one vector **x** such that $T(\mathbf{x}) = \mathbf{w}$.

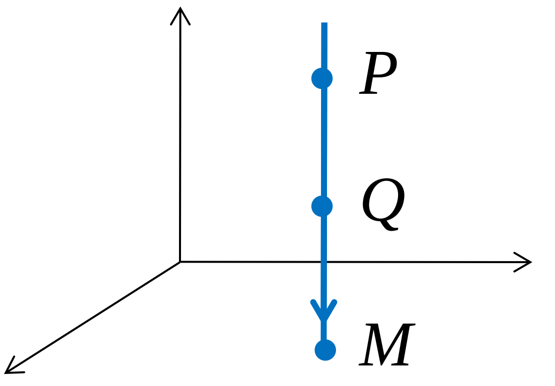
Example

One-to-one linear transformation



Distinct vectors **u** and **v** are rotated into distinct vectors $T(\mathbf{u})$ and $T(\mathbf{v})$.

Not one-to-one linear transformation



The distinct points P and Q are mapped into the same point M.

Theorem 4.10.1 (Equivalent Statements)

- If A is an $n \times n$ matrix and $T_A : R^n \rightarrow R^n$ *is* multiplication by A, then the following statements are equivalent.
  - A is invertible
  - The range of T_A is R^n
  - T_A is one-to-one

Proof of Theorem 4.10.1

- (a)→(b): Assume A is invertible. $A\mathbf{x}=\mathbf{b}$ is consistent for every $n \times 1$ matrix $\mathbf{b}$ in R^n. This implies that T_A maps $\mathbf{x}$ into the arbitrary vector $\mathbf{b}$ in R^n, which implies the range of T_A is R^n.
- (b)→(c): Assume the range of T_A is R^n. For every vector $\mathbf{b}$ in R^n there is some vector $\mathbf{x}$ in R^n for which $T_A(\mathbf{x})=\mathbf{b}$ and hence the linear system $A\mathbf{x}=\mathbf{b}$ is consistent for every vector $\mathbf{b}$ in R^n. But we know $A\mathbf{x}=\mathbf{b}$ has a unique solution, and hence for every vector $\mathbf{b}$ in the range of T_A there is exactly one vector $\mathbf{x}$ in R^n such that $T_A(\mathbf{x})=\mathbf{b}$.

Example

- The rotation operator $T : R^2 \to R^2$ is one-to-one
  - The standard matrix for T is $$[T] = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$
  - $[T]$ is invertible since $$\det\begin{vmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{vmatrix} = \cos^2\theta + \sin^2\theta = 1 \neq 0$$

Example

- The projection operator $T: R^3 \rightarrow R^3$ is not one-to-one
  - The standard matrix for T is

$$[T] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

  - $[T]$ is not invertible since $\det[T] = 0$

Inverse of a One-to-One Linear Operator

- Suppose $T_A : R^n \rightarrow R^n$ is a one-to-one linear operator

  $\Rightarrow$ The matrix A is invertible.

  $\Rightarrow$ $T_A^{-1} : R^n \rightarrow R^n$ is itself a linear operator; it is called the inverse of T_A.

  $\Rightarrow$ $T_A(T_A^{-1}(\mathbf{x})) = AA^{-1}\mathbf{x} = I\mathbf{x} = \mathbf{x}$ and

  $T_A^{-1}(T_A(\mathbf{x})) = A^{-1}A\mathbf{x} = I\mathbf{x} = \mathbf{x}$

  $\Rightarrow$ $T_A \circ T_A^{-1} = T_{AA^{-1}} = T_I$ and

  $T_A^{-1} \circ T_A = T_{A^{-1}A} = T_I$

Inverse of a One-to-One Linear Operator

- If **w** is the image of **x** under T_A, then T_A^{-1} maps **w** back into **x**, since

$$T_A^{-1}(\mathbf{w}) = T_A^{-1}(T_A(\mathbf{x})) = \mathbf{x}$$

- When a one-to-one linear operator on R^n is written as $T : R^n \rightarrow R^n$, then the inverse of the operator T is denoted by T^{-1}.
- Thus, by the standard matrix, we have $[T^{-1}]=[T]^{-1}$

Example

- Let $T : R^2 \to R^2$ be the operator that rotates each vector in R^2 through the angle θ:

$$[T] = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

- Undo the effect of T means rotate each vector in R^2 through the angle -θ.

- This is exactly what the operator T^{-1} does: the standard matrix T^{-1} is

$$[T^{-1}] = [T]^{-1} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} = \begin{bmatrix} \cos(-\theta) & -\sin(-\theta) \\ \sin(-\theta) & \cos(-\theta) \end{bmatrix}$$

- The only difference is that the angle θ is replaced by -θ

Example

- Show that the linear operator $T : R^2 \rightarrow R^2$ defined by the equations

$$w_1 = 2x_1 + x_2$$
$$w_2 = 3x_1 + 4x_2$$

is one-to-one, and find $T^{-1}(w_1, w_2)$.

- Solution:

$$\begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \Rightarrow [T] = \begin{bmatrix} 2 & 1 \\ 3 & 4 \end{bmatrix} \Rightarrow [T^{-1}] = [T]^{-1} = \begin{bmatrix} \frac{4}{5} & -\frac{1}{5} \\ -\frac{3}{5} & \frac{2}{5} \end{bmatrix}$$

$$\Rightarrow [T^{-1}] \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} \frac{4}{5} & -\frac{1}{5} \\ -\frac{3}{5} & \frac{2}{5} \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} \frac{4}{5} w_1 - \frac{1}{5} w_2 \\ -\frac{3}{5} w_1 + \frac{2}{5} w_2 \end{bmatrix}$$

$$\Rightarrow T^{-1}(w_1, w_2) = \left(\frac{4}{5} w_1 - \frac{1}{5} w_2, -\frac{3}{5} w_1 + \frac{2}{5} w_2\right)$$

Linearity Properties

- Theorem 4.10.2 (Properties of Linear Transformations)
  - A transformation $T : R^n \rightarrow R^m$ is linear if and only if the following relationships hold for all vectors **u** and **v** in R^n and every scalar c.
    - $T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$
    - $T(c\mathbf{u}) = cT(\mathbf{u})$

Proof of Theorem 4.10.2

- Conversely, assume that properties (a) and (b) hold for the transformation T. We can prove that T is linear by finding a matrix A with the property that $T(\mathbf{x}) = A\mathbf{x}$ for all vectors $\mathbf{x}$ in R^n.
- The property (a) can be extended to three or more terms. $T(\mathbf{u}+\mathbf{v}+\mathbf{w}) = T(\mathbf{u}+(\mathbf{v}+\mathbf{w})) = T(\mathbf{u})+T(\mathbf{v}+\mathbf{w}) = T(\mathbf{u})+T(\mathbf{v})+T(\mathbf{w})$
- More generally, for any vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ in R^n, we have

$$T(\mathbf{v}_1+\mathbf{v}_2+\ldots+\mathbf{v}_k) = T(\mathbf{v}_1) + T(\mathbf{v}_2) + \ldots + T(\mathbf{v}_k)$$

Proof of Theorem 4.10.2

- Now, to find the matrix A, let $\mathbf{e}_1$, $\mathbf{e}_2$, …, $\mathbf{e}_n$ be the vectors

$$\boldsymbol{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad \boldsymbol{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad \cdots\cdots \quad \boldsymbol{e}_n = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

- Let A be the matrix whose successive column vectors are $T(\mathbf{e}_1)$, $T(\mathbf{e}_2)$, …, $T(\mathbf{e}_n)$; that is
  $A = [T(\mathbf{e}_1) \mid T(\mathbf{e}_2) \mid \ldots \mid T(\mathbf{e}_n)]$

Proof of Theorem 4.10.2

- If $\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ is any vector in R^n, then as discussed in

  Section 1.3, the product $A\mathbf{x}$ is a linear combination of the column vectors of A with coefficients $\mathbf{x}$, so

$$
\begin{aligned}
A\mathbf{x} &= x_1T(\mathbf{e}_1) + x_2T(\mathbf{e}_2) + \ldots + x_nT(\mathbf{e}_n) \\
&= T(x_1\mathbf{e}_1) + T(x_2\mathbf{e}_2) + \ldots + T(x_n\mathbf{e}_n) \\
&= T(x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + .. + x_n\mathbf{e}_n) \\
&= T(\mathbf{x})
\end{aligned}
$$

Theorem 4.10.3

- **Every linear transformation from R^n to R^m is a matrix transformation, and conversely, every matrix transformation from R^n to R^m is a linear transformation.**

Theorem 4.10.4 (Equivalent Statements)

- If A is an $m \times n$ matrix, and if $T_A : R^n \rightarrow R^n$ is multiplication by A, then the following are equivalent:
  - A is invertible.
  - $A\mathbf{x} = \mathbf{0}$ has only the trivial solution.
  - The reduced row-echelon form of A is I_n.
  - A is expressible as a product of elementary matrices.
  - $A\mathbf{x} = \mathbf{b}$ is consistent for every $n \times 1$ matrix $\mathbf{b}$.
  - $A\mathbf{x} = \mathbf{b}$ has exactly one solution for every $n \times 1$ matrix $\mathbf{b}$.
  - $\det(A) \neq 0$.
  - The column vectors of A are linearly independent.
  - The row vectors of A are linearly independent.
  - The column vectors of A span R^n.
  - The row vectors of A span R^n.
  - The column vectors of A form a basis for R^n.
  - The row vectors of A form a basis for R^n.
  - A has rank n.
  - A has nullity 0.
  - The orthogonal complement of the nullspace of A is R^n.
  - The orthogonal complement of the row space of A is $\{\mathbf{0}\}$.
  - The range of T_A is R^n.
  - T_A is one-to-one.

4.11
Geometry of Matrix Operations

Example: Transforming with Diagonal Matrices

- Suppose that the *xy*-plane first is compressed or expanded by a factor of k_1 in the *x*-direction and then is compressed or expanded by a factor of k_2 in the *y*-direction. Find a single matrix operator that performs both operations.

$$\begin{bmatrix} k_1 & 0 \\ 0 & 1 \end{bmatrix}$$

x-compression (expansion)

$$\begin{bmatrix} 1 & 0 \\ 0 & k_2 \end{bmatrix}$$

y-compression (expansion)

$$A = \begin{bmatrix} 1 & 0 \\ 0 & k_2 \end{bmatrix} \begin{bmatrix} k_1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} k_1 & 0 \\ 0 & k_2 \end{bmatrix}$$

- If $k_1=k_2=k$, this is a contraction or dilation. $A = \begin{bmatrix} k & 0 \\ 0 & k \end{bmatrix}$

Example

- Find a matrix transformation from R^2 to R^2 that first shears by a factor of 2 in the *x*-direction and then reflects about $y = x$.
- The standard matrix for the shear is $A_1 = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$

  and for the reflection is $A_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$

- Thus the standard matrix for the sear followed by the reflection is

$$A_2 A_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 2 \end{bmatrix}$$

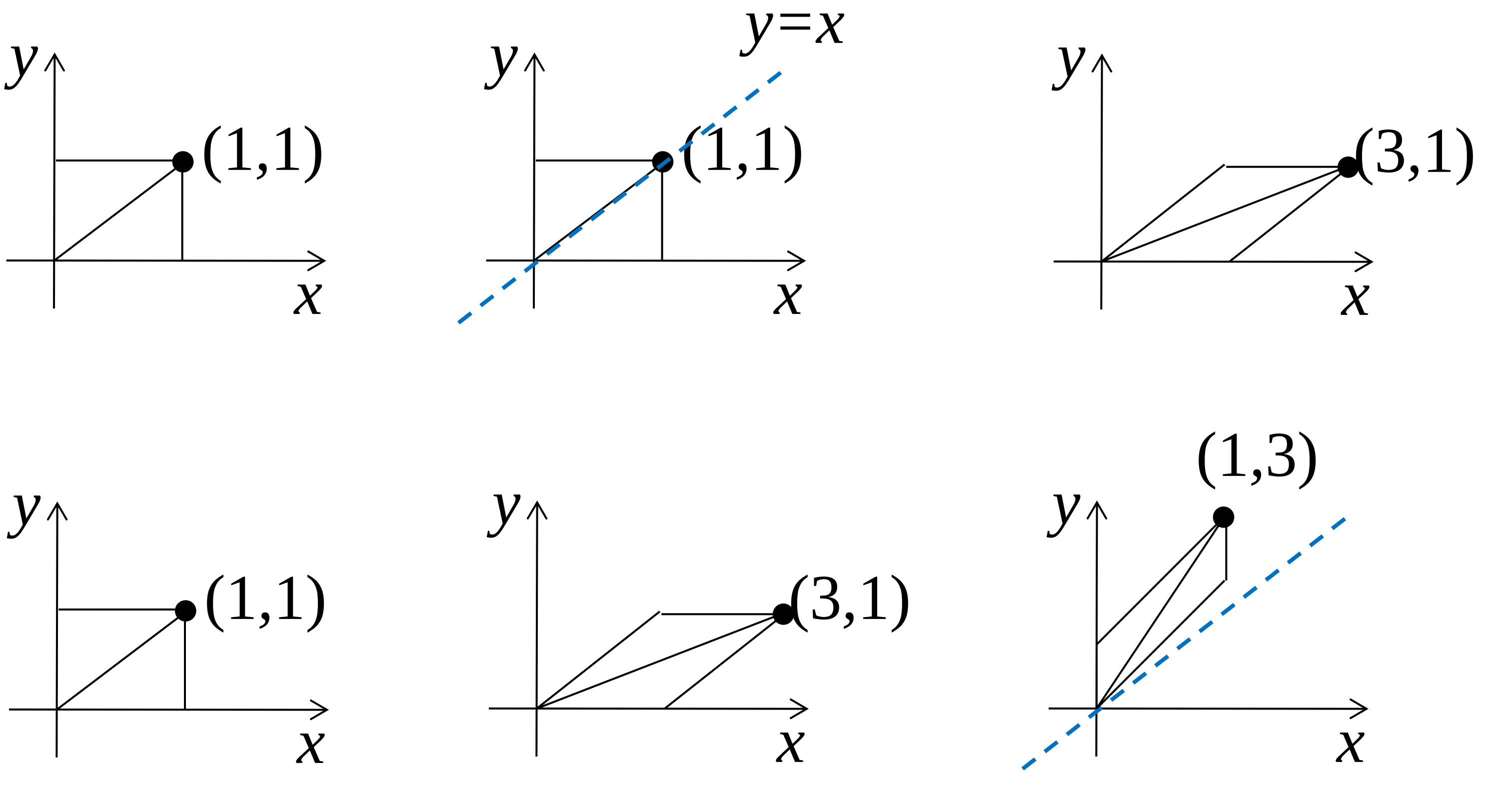
Example

- Find a matrix transformation from R^2 to R^2 that first reflects about $y = x$ and then shears by a factor of 2 in the x-direction.

$$A_1 A_2 = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 0 \end{bmatrix}$$

- Note that $A_1 A_2 \neq A_2 A_1$

Geometry



Geometry of One-to-One Matrix Operators

- A matrix transformation T_A is **one-to-one** if and only if A is **invertible** and **can be expressed as a product of elementary matrices**.

$$A = E_1 E_2 \cdots E_r$$

$$T_A = T_{E_1 E_2 \cdots E_r} = T_{E_1} \circ T_{E_2} \circ \cdots \circ T_{E_r}$$

- **Theorem 4.11.1:** If E is an elementary matrix, then T_E: $R^2 \rightarrow R^2$ is one of the following:
  - A shear along a coordinate axis
  - A reflection about y=x
  - A compression along a coordinate axis
  - An expansion along a coordinate axis
  - A reflection about a coordinate axis
  - A compression or expansion along a coordinante axis followed by a reflection about a coordinate axis

Proof of Theorem 4.11.1

- Because a 2×2 elementary matrix results from performing a single elementary row operation on the 2×2 identity matrix, it must have one of the following forms:

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

- $\begin{bmatrix} 1 & 0 \\ k & 1 \end{bmatrix}$ and $\begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix}$ represent **shears** along coordinates axes.
- $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ represents a **reflection** about $y = x$.

Proof of Theorem 4.11.1

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

- If $k > 0$, $\begin{bmatrix}k & 0\\ 0 & 1\end{bmatrix}$ and $\begin{bmatrix}1 & 0\\ 0 & k\end{bmatrix}$ represent compressions or expansion along coordinate axes, depending on whether $0 \leq k \leq 1$ (compression) or $k \geq 1$ (expansion).
- If ***k* < 0**, and if we express k in the form $k=-k_1$, where **$k_1>0$**, then

$$\begin{bmatrix}k & 0\\ 0 & 1\end{bmatrix} = \begin{bmatrix}-k_1 & 0\\ 0 & 1\end{bmatrix} = \begin{bmatrix}-1 & 0\\ 0 & 1\end{bmatrix}\begin{bmatrix}k_1 & 0\\ 0 & 1\end{bmatrix}$$

$$\begin{bmatrix}1 & 0\\ 0 & k\end{bmatrix} = \begin{bmatrix}1 & 0\\ 0 & -k_1\end{bmatrix} = \begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}\begin{bmatrix}1 & 0\\ 0 & k_1\end{bmatrix}$$

Proof of Theorem 4.11.1

$$\begin{bmatrix} k & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} -k_1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} k_1 & 0 \\ 0 & 1 \end{bmatrix}$$

- It represents a **compression or expansion** along the *x*-axis followed by a **reflection** about the *y*-axis.

$$\begin{bmatrix} 1 & 0 \\ 0 & k \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -k_1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & k_1 \end{bmatrix}$$

- It represents a **compression or expansion** along the *y*-axis followed by a **reflection** about the *x*-axis.

Theorem 4.11.2

- **If $T_A\!:\!R^2 \rightarrow R^2$ is multiplication by an invertible matrix A, then the geometric effect of T_A is the same as an appropriate succession of shears, compressions, expansions, and reflections.**

Example: Geometric Effect of Multiplication by a Matrix

- Assuming that k_1 and k_2 are positive, express the diagonal matrix $A = \begin{bmatrix} k_1 & 0 \\ 0 & k_2 \end{bmatrix}$ as a product of elementary matrices, and describe the geometric effect of multiplication by A in terms of compressions and expansions.
- We know

interchangeable!

$$A = \begin{bmatrix} k_1 & 0 \\ 0 & k_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & k_2 \end{bmatrix} \begin{bmatrix} k_1 & 0 \\ 0 & 1 \end{bmatrix}$$

which shows the geometric effect of compressing or expanding by a factor of k_1 in the x-direction and then compressing or expanding by a factor of k_2 in the y-direction.

Example

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

- Express A as a product of elementary matrices, and then describe the geometric effect of multiplication by A in terms of shears, compressions, expansion, and reflections.
- A can be reduced to I as follows:

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

Add -3 times the first row to the second

Multiply the second row by -1/2

Add -2 times the second row to the first

- The three successive row operations can be performed by multiplying on the left successively by

$$E_1 = \begin{bmatrix} 1 & 0 \\ -3 & 1 \end{bmatrix} \qquad E_2 = \begin{bmatrix} 1 & 0 \\ 0 & -\frac{1}{2} \end{bmatrix} \qquad E_3 = \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix}$$

Example

- Inverting these matrices

3 2 1

$$A = E_1^{-1} E_2^{-1} E_3^{-1} = \begin{bmatrix} 1 & 0 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -2 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$$

- Reading from right to left and noting that

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

it follows that the effect of multiplying by A is equivalent to
1. shearing by a factor of 2 in the x-direction,
2. then expanding by a factor of 2 in the y-direction,
3. then reflecting about the x-axis,
4. then shearing by a factor of 3 in the y-direction.

Theorem 4.11.3

- If T:$R^2 \rightarrow R^2$ is multiplication by an invertible matrix, then
  - (a) the image of a straight line is a straight line.
  - (b) the image of a straight line through the origin is a straight line through the origin.
  - (c) the images of parallel straight lines are parallel straight lines.
  - (d) the images of the line segment joining points *P* and *Q* is the line segment joining the images of *P* and *Q*.
  - (e) the images of three points lie on a line if and only if the points themselves line on some line.

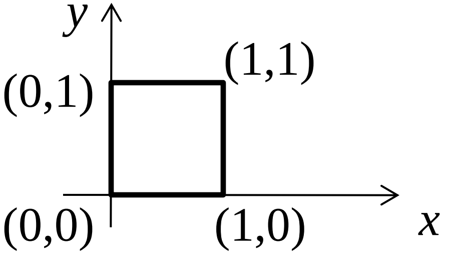
Example: Image of a Square

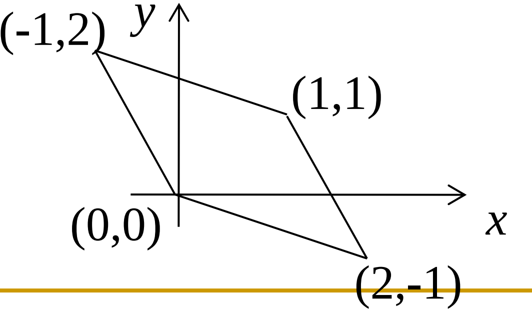
- Sketch **the images of the unit square** under multiplication by

$$A = \begin{bmatrix} -1 & 2 \\ 2 & -1 \end{bmatrix}$$

- Since $\begin{bmatrix} -1 & 2 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ $\quad \begin{bmatrix} -1 & 2 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$

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





Example: Image of a Line

$$A = \begin{bmatrix} 3 & 1 \\ 2 & 1 \end{bmatrix}$$

- The invertible matrix maps the **line *y*=2*x*+1** into another line. Find its equation.
- Let (x,y) be a point on the line y=2x+1, and let (x',y') be its image under multiplication by A. Then

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} 3 & 1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \Rightarrow \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 3 & 1 \\ 2 & 1 \end{bmatrix}^{-1} \begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ -2 & 3 \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix}$$

- So $\begin{aligned} x &= x' - y' \\ y &= -2x' + 3y' \end{aligned} \Rightarrow y' = \frac{4}{5}x' + \frac{1}{5}$
- Thus (x', y') satisfies $y = \frac{4}{5}x + \frac{1}{5}$, which is the equation we want.