7.8 Inverse of a Matrix. Gauss–Jordan Elimination

Given a **square** matrix **A:**

- The **inverse** of an $n \times n$ matrix $\mathbf{A} = [a_{jk}]$ is denoted by $\mathbf{A}^{-1}$ and is an $n \times n$ matrix such that

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

where **I** is the $n \times n$ unit matrix (see Sec. 7.2).

- If **A** has an inverse, then **A** is called a **nonsingular matrix**. If **A** has no inverse, then **A** is called a **singular matrix**.

- If **A** has an inverse, the inverse is *unique*: if both **B** and **C** are inverses of **A**, then $\mathbf{AB} = \mathbf{I}$ and $\mathbf{CA} = \mathbf{I}$ so that we obtain the uniqueness from

$$\mathbf{B} = \mathbf{IB} = (\mathbf{CA})\mathbf{B} = \mathbf{C}(\mathbf{AB}) = \mathbf{CI} = \mathbf{C}.$$

THEOREM 1: Existence of the Inverse

- The inverse $\mathbf{A}^{-1}$ of an $n \times n$ matrix $\mathbf{A}$ exists iff rank $\mathbf{A} = n$, or, equivalently, iff det $\mathbf{A} \neq 0$.
- Hence, $\mathbf{A}$ is nonsingular (inverse exist) if rank $\mathbf{A} = n$, and is singular (inverse not exist) if rank $\mathbf{A} < n$
- Idea: conduct elementary row operations on $\mathbf{A}$

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & \cdots & a_{nn} \end{bmatrix} \xrightarrow{\text{Elementary row operations}} \mathbf{R} = \begin{bmatrix} r_{11} & a_{12} & \cdots & \cdots & r_{1n} \\ & r_{22} & \cdots & \cdots & r_{2n} \\ & & \ddots & & \vdots \\ & & & \ddots & \vdots \\ & & & & r_{nn} \end{bmatrix}$$

All r_{ii} nonzero $\Leftrightarrow$ det $\mathbf{A} \neq 0$ $\Leftrightarrow$ Rank $\mathbf{A} = n$

$\Leftrightarrow$ Inverse $\mathbf{A}^{-1}$ exists and $\mathbf{A}$ is nonsingular

EXAMPLE 1:
Finding Matrix Inverse by Gauss–Jordan Elimination

Determine the inverse $\mathbf{A}^{-1}$ of

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

Solution.

Perform elementary row operations on $n \times 2n = 3 \times 6$ augmented matrix $[\mathbf{A} \mid \mathbf{I}]$.

$$[\mathbf{A} \mid \mathbf{I}] = \left[\begin{array}{ccc|ccc} -1 & 1 & 2 & 1 & 0 & 0 \\ 3 & -1 & 1 & 0 & 1 & 0 \\ -1 & 3 & 4 & 0 & 0 & 1 \end{array}\right]$$

$$= \left[\begin{array}{ccc|ccc} -1 & 1 & 2 & 1 & 0 & 0 \\ 0 & 2 & 7 & 3 & 1 & 0 \\ 0 & 2 & 2 & -1 & 0 & 1 \end{array}\right] \quad \begin{array}{l} \\ \text{Row 2} + 3\ \text{Row 1} \\ \text{Row 3} - \text{Row 1} \end{array}$$

EXAMPLE 4 (continued)

Finding Matrix Inverse by Gauss–Jordan Elimination

$$= \left[\begin{array}{ccc|ccc} -1 & 1 & 2 & 1 & 0 & 0 \\ 0 & 2 & 7 & 3 & 1 & 0 \\ 0 & 0 & -5 & -4 & -1 & 1 \end{array}\right] \quad \begin{array}{l} \\ \\ \text{Row 3} - \text{Row 2} \end{array}$$

Take additional Gauss–Jordan steps to reduce **U** to **I**, i.e., to diagonal form with entries 1 on the main diagonal.

$$= \left[\begin{array}{ccc|ccc} 1 & -1 & -2 & -1 & 0 & 0 \\ 0 & 1 & 3.5 & 1.5 & 0.5 & 0 \\ 0 & 0 & 1 & 0.8 & 0.2 & -0.2 \end{array}\right] \quad \begin{array}{r} -\text{Row 1} \\ 0.5\ \text{Row 2} \\ -0.2\ \text{Row 3} \end{array}$$

EXAMPLE 4 (continued) Finding Matrix Inverse by Gauss–Jordan Elimination

$$= \left[\begin{array}{ccc|ccc} 1 & -1 & 0 & 0.6 & 0.4 & -0.4 \\ 0 & 1 & 0 & -1.3 & -0.2 & 0.7 \\ 0 & 0 & 1 & 0.8 & 0.2 & -0.2 \end{array}\right] \quad \begin{array}{l} \text{Row 1} + 2 \text{ Row 3} \\ \text{Row 2} - 3.5 \text{ Row 3} \\ \end{array}$$

$$= \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -0.7 & 0.2 & 0.3 \\ 0 & 1 & 0 & -1.3 & -0.2 & 0.7 \\ 0 & 0 & 1 & 0.8 & 0.2 & -0.2 \end{array}\right] \quad \text{Row 1} + \text{Row 2}$$

The last three columns constitute $\mathbf{A}^{-1}$. Check:

$$\begin{bmatrix} -1 & 1 & 2 \\ 3 & -1 & 1 \\ -1 & 3 & 4 \end{bmatrix} \begin{bmatrix} -0.7 & 0.2 & 0.3 \\ -1.3 & -0.2 & 0.7 \\ 0.8 & 0.2 & -0.2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Hence $\mathbf{AA}^{-1} = \mathbf{I}$. Similarly $\mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$.

EXAMPLE:

Determine the inverse $\mathbf{A}^{-1}$ of $\mathbf{A}$.

$$\mathbf{A} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & -2 & 0 \end{bmatrix} \rightarrow [\mathbf{A} \mid \mathbf{I}] = \left[\begin{array}{ccc|ccc} 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & -2 & 0 & 0 & 0 & 1 \end{array}\right]$$

$$= \left[\begin{array}{ccc|ccc} 1 & -2 & 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \end{array}\right]$$

Row 1 swap with Row 3

$$= \left[\begin{array}{ccc|ccc} 1 & -2 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & -1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \end{array}\right]$$

Row 2 – Row 3

$$= \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -2 & 2 & 1 \\ 0 & 1 & 0 & -1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \end{array}\right]$$

Row 1 +2 Row 2

$$\mathbf{A}^{-1} = \begin{bmatrix} -2 & 2 & 1 \\ -1 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \leftarrow$$

Proof (Optional): Matrix Inverse by Gauss–Jordan Elimination

- Each elementary row operations can be represented by a matrix operation
- Interchange of two rows : Define $\mathbf{RO}_{swap\,(Rj,Rk)}$ such that

$\overline{\mathbf{A}} = \mathbf{RO}_{swap\,(Rj,Rk)}\mathbf{A}$ is the matrix $\mathbf{A}$ with row j and k swapped.

$$\mathbf{A} = \begin{bmatrix} - & \mathbf{a}_{(1)} & - \\ & \vdots & \\ - & \mathbf{a}_{(j)} & - \\ & \vdots & \\ - & \mathbf{a}_{(k)} & - \\ & \vdots & \\ - & \mathbf{a}_{(n)} & - \end{bmatrix} \qquad \overline{\mathbf{A}} = \begin{bmatrix} - & \mathbf{a}_{(1)} & - \\ & \vdots & \\ - & \mathbf{a}_{(k)} & - \\ & \vdots & \\ - & \mathbf{a}_{(j)} & - \\ & \vdots & \\ - & \mathbf{a}_{(n)} & - \end{bmatrix} \text{swap}$$

- Matrix Operator for interchanging Row 2 and Row 4 of a 5x5 matrix: $\mathbf{RO}_{swap\,(R2,R4)} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$

(Optional)

- Addition of a multiple of a row to another row:

  Define $\mathbf{RO}_{add\,(Rj,cRk)}$ such that $\overline{\mathbf{A}} = \mathbf{RO}_{add\,(Rj,cRk)}\mathbf{A}$ is matrix $\mathbf{A}$ with row j added with row k multiplying a constant c .

$$\mathbf{A} = \begin{bmatrix} \text{---} & \mathbf{a}_{(1)} & \text{---} \\ & \vdots & \\ \text{---} & \mathbf{a}_{(j)} & \text{---} \\ & \vdots & \\ \text{---} & \mathbf{a}_{(k)} & \text{---} \\ & \vdots & \\ \text{---} & \mathbf{a}_{(n)} & \text{---} \end{bmatrix}, \quad \overline{\mathbf{A}} = \begin{bmatrix} \text{---} & \mathbf{a}_{(1)} & \text{---} \\ & \vdots & \\ \text{---} & \mathbf{a}_{(j)} + c\,\mathbf{a}_{(k)} & \text{---} \\ & \vdots & \\ \text{---} & \mathbf{a}_{(k)} & \text{---} \\ & \vdots & \\ \text{---} & \mathbf{a}_{(n)} & \text{---} \end{bmatrix}$$

- Matrix Operator for adding c times Row 4 to Row 2 of a 5 by 5 matrix:

$$\mathbf{RO}_{add\,(R2,cR4)} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & c & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

(Optional)

- Multiplication of a row by a nonzero constant c :

Define $\mathbf{RO}_{mult\,(c,Rj)}$ such that $\overline{\mathbf{A}} = \mathbf{RO}_{mult\,(c,Rj)}\mathbf{A}$ is matrix $\mathbf{A}$ with row j added by row k multiplying a constant c .

$$\mathbf{A} = \begin{bmatrix} \text{——} & \mathbf{a}_{(1)} & \text{——} \\ & \vdots & \\ \text{——} & \mathbf{a}_{(j)} & \text{——} \\ & \vdots & \\ \text{——} & \mathbf{a}_{(k)} & \text{——} \\ & \vdots & \\ \text{——} & \mathbf{a}_{(n)} & \text{——} \end{bmatrix}, \quad \overline{\mathbf{A}} = \begin{bmatrix} \text{——} & \mathbf{a}_{(1)} & \text{——} \\ & \vdots & \\ \text{——} & c\mathbf{a}_{(j)} & \text{——} \\ & \vdots & \\ \text{——} & \mathbf{a}_{(k)} & \text{——} \\ & \vdots & \\ \text{——} & \mathbf{a}_{(n)} & \text{——} \end{bmatrix}$$

- Matrix Operator for multiplying Row 2 of a 5 by 5 matrix by constant c:

$$\mathbf{RO}_{mult\,(c,R2)} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & c & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

- Matrix operators for the three elementary row operations defined: $\mathbf{RO}_{swap(Rj,Rk)}$, $\mathbf{RO}_{add(Rj,cRk)}$, $\mathbf{RO}_{mult(c,Rj)}$ **(Optional)**
- Given matrix **A**, applying a series of elementary row operations, $\mathbf{RO}_1$, $\mathbf{RO}_2$, $\mathbf{RO}_3$, ……., $\mathbf{RO}_L$ to convert **A** to become identity matrix **I**, i.e.,

$$\mathbf{RO}_L\ldots\ldots\mathbf{RO}_3\mathbf{RO}_2\mathbf{RO}_1\mathbf{A}=\mathbf{I}$$

Then the inverse $\mathbf{A}^{-1} = (\mathbf{RO}_L\ldots\ldots\mathbf{RO}_3\mathbf{RO}_2\mathbf{RO}_1)$

- For better recording of the **RO**s performed, conduct elementary row operations on the augmented matrix [**A** | **I**]. When **A** on the left becomes **I**, the right hand side yield $\mathbf{A}^{-1}$

$$[\mathbf{A} \mid \mathbf{I}] \longrightarrow [\,\mathrm{I} \mid \overbrace{\mathbf{RO}_L\ldots\ldots\mathbf{RO}_3\mathbf{RO}_2\mathbf{RO}_1}^{\mathbf{A}^{-1}}\,]$$

THEOREM 2: Formulas for Inverses

Inverse of a Matrix by Determinants

- The inverse of a nonsingular $n \times n$ matrix $\mathbf{A} = [a_{jk}]$ is given by

$$(4) \qquad \mathbf{A}^{-1} = \frac{1}{\det \mathbf{A}}\left[C_{jk}\right]^{\mathsf{T}} = \frac{1}{\det \mathbf{A}}\begin{bmatrix} C_{11} & C_{21} & \cdots & C_{n1} \\ C_{12} & C_{22} & \cdots & C_{n2} \\ \cdot & \cdot & \cdots & \cdot \\ C_{1n} & C_{2n} & \cdots & C_{nn} \end{bmatrix},$$

  where C_{jk} is the cofactor of a_{jk} in det $\mathbf{A}$ (see Sec. 7.7).

- Note: In $\mathbf{A}^{-1}$ above, the cofactor C_{jk} occupies the same place as a_{kj} (not a_{jk}) does in $\mathbf{A}$.

- Particularly, the inverse of a 2 × 2 matrix by this formula is:

$$(4^*) \qquad \mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \quad \textit{is} \quad \mathbf{A}^{-1} = \frac{1}{\det \mathbf{A}} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}.$$

EXAMPLE 2

$$\mathbf{A} = \begin{bmatrix} 3 & 1 \\ 2 & 4 \end{bmatrix}, \qquad \mathbf{A}^{-1} = \frac{1}{10} \begin{bmatrix} 4 & -1 \\ -2 & 3 \end{bmatrix} = \begin{bmatrix} 0.4 & -0.1 \\ -0.2 & 0.3 \end{bmatrix}$$

EXAMPLE 3

- Use (4) to find the inverse of $\mathbf{A} = \begin{bmatrix} -1 & 1 & 2 \\ 3 & -1 & 1 \\ -1 & 3 & 4 \end{bmatrix}$.
- ***Solution.***
  Obtain det $\mathbf{A} = -1(-7) - 1 \cdot 13 + 2 \cdot 8 = 10$, and in (4),

$$C_{11} = \begin{vmatrix} -1 & 1 \\ 3 & 4 \end{vmatrix} = -7, \quad C_{21} = -\begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix} = 2, \quad C_{31} = \begin{vmatrix} 1 & 2 \\ -1 & 1 \end{vmatrix} = 3,$$

$$C_{12} = -\begin{vmatrix} 3 & 1 \\ -1 & 4 \end{vmatrix} = -13, \quad C_{22} = \begin{vmatrix} -1 & 2 \\ -1 & 4 \end{vmatrix} = -2, \quad C_{32} = -\begin{vmatrix} -1 & 2 \\ 3 & 1 \end{vmatrix} = 7,$$

$$C_{13} = \begin{vmatrix} 3 & -1 \\ -1 & 3 \end{vmatrix} = 8, \quad C_{23} = -\begin{vmatrix} -1 & 1 \\ -1 & 3 \end{vmatrix} = 2, \quad C_{33} = \begin{vmatrix} -1 & 1 \\ 3 & -1 \end{vmatrix} = -2,$$

EXAMPLE 3 (continued)

By (4), the inverse is

$$\mathbf{A}^{-1} = \begin{bmatrix} -0.7 & 0.2 & 0.3 \\ -1.3 & -0.2 & 0.7 \\ 0.8 & 0.2 & -0.2 \end{bmatrix}.$$

in agreement with Example 1.

Inverse of Matrix Products obtained by taking the inverse of each factor and multiplying these inverses ***in reverse order,***

(7) $$(\mathbf{AC})^{-1} = \mathbf{C}^{-1}\mathbf{A}^{-1}.$$

Hence for more than two factors,

(8) $$(\mathbf{AC} \ldots \mathbf{PQ})^{-1} = \mathbf{Q}^{-1}\mathbf{P}^{-1} \ldots \mathbf{C}^{-1}\mathbf{A}^{-1}.$$

EXAMPLE: Elimination Theory

- Finding common points to some high-dimensional "surfaces".
- A simple, low dimensional example*: Given 2 polynomials

$$f(x) = -2x^4 + 4x^3 - x^2 + 5x - 1 \text{ and } g(x) = x^2 - 6x + 9$$

Question: do they share a common root?

- If they share a common root, denoted as $x = x_c$, then

$$f(x) = (Ax^3 + Bx^2 + Cx + D)(x - x_c) \text{ and } g(x) = (Ex + F)(x - x_c)$$

or $$f(x)(Ex + F) = g(x)(Ax^3 + Bx^2 + Cx + D)$$

- Equating the coefficients of x^5, x^4, x^3, x^2, x and x^0 (constant) on both side:

$$\begin{aligned} A + 2E &= 0 \\ -6A + B - 4E + 2F &= 0 \\ 9A - 6B + C + E - 4F &= 0 \\ 9B - 6C + D - 5E + F &= 0 \\ 9C - 6D + E - 5F &= 0 \\ 9D + F &= 0 \end{aligned} \qquad \text{(E)}$$

*http://aix1.uottawa.ca/~jkhoury/elimination.htm

- We need existence of constant A, B, C, D, E and F so that the equation (E) is satisfied.
- Hence, we desire non-trivial solution for (E))

  → Determinant of coefficient matrix det(**A**)=0
- Check

$$\begin{vmatrix} 1 & 0 & 0 & 0 & 2 & 0 \\ -6 & 1 & 0 & 0 & -4 & 2 \\ 9 & -6 & 1 & 0 & 1 & -4 \\ 0 & 9 & -6 & 1 & -5 & 1 \\ 0 & 0 & 9 & -6 & 1 & -5 \\ 0 & 0 & 0 & 9 & 0 & 1 \end{vmatrix} \neq 0$$

Hence, no common roots for $f(x)$ and $g(x)$.

- How about $f(x) = -2x^4 + 4x^3 - x^2 + 5x - 6$ and $g(x) = x^2 - 6x + 8$? Do they share common root?

Properties of Matrix Multiplication. Cancellation Laws

Generally,

- Matrix multiplication is not commutative:
$$\mathbf{AB} \neq \mathbf{BA}.$$
- $\mathbf{AC} = \mathbf{AD}$ does NOT imply $\mathbf{C} = \mathbf{D}$ (even when $\mathbf{A} \neq \mathbf{0}$), e.g., $\begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix}\begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$
- $\mathbf{AB} = \mathbf{0}$ NOT imply $\mathbf{BA} = \mathbf{0}$ or $\mathbf{A} = \mathbf{0}$ or $\mathbf{B} = \mathbf{0}$

THEOREM 3.
$\mathbf{A}$ is a square matrix.

- If rank $\mathbf{A} < n$ (no inverse exists), then rank$(\mathbf{BA}) < n$ and rank $(\mathbf{AB}) < n$
- If rank $\mathbf{A} = n$ (i.e., $\mathbf{A}^{-1}$ exists), $\mathbf{AC} = \mathbf{AD}$ implies $\mathbf{C} = \mathbf{D}$.
- If rank $\mathbf{A} = n$, $\mathbf{AB} = \mathbf{0}$ or $\mathbf{BA} = \mathbf{0}$ implies $\mathbf{B} = \mathbf{0}$.
- Hence, if $\mathbf{AB} = \mathbf{0}$ but $\mathbf{A} \neq \mathbf{0}$ and $\mathbf{B} \neq \mathbf{0}$, then rank $\mathbf{A} < n$ and rank $\mathbf{B} < n$.)

THEOREM 4: Determinant of a Product of Matrices

For any $n \times n$ matrices $\mathbf{A}$ and $\mathbf{B}$,

$$\det(\mathbf{AB}) = \det(\mathbf{BA}) = \det \mathbf{A} \det \mathbf{B}. \tag{10}$$

Proof (Optional).

- Matrix Operators for elementary row operations:
  - $\det(\mathbf{RO}_{swap(Rj,Rk)}\mathbf{M})=\mathrm{d}_{swap(Rj,Rk)} \det(\mathbf{M})$; $\mathrm{d}_{swap(Rj,Rk)} =(-1)$
  - $\det(\mathbf{RO}_{add(Rj,cRk)}\mathbf{M})=\mathrm{d}_{add(Rj,cRk)} \det(\mathbf{M})$; $\mathrm{d}_{add(Rj,cRk)}=(+1)$
  - $\det(\mathbf{RO}_{mult(c,Rj)}\mathbf{M})=\mathrm{d}_{mult(c,Rj)} \det(\mathbf{M})$; $\mathrm{d}_{mult(c,Rj)}=c$
- Hence, $\det(\mathbf{RO}_i\mathbf{M})=\mathrm{d}_i \det(\mathbf{M})$, for i=*swap*, *add* or *mult*.

- For inverse Matrix Operators for elementary row operations: **(Optional)**

$$\mathbf{RO}^{-1}_{swap\,(Rj,Rk)} = \mathbf{RO}_{swap\,(Rj,Rk)}$$

$$\mathbf{RO}^{-1}_{add\,(Rj,cRk)} = \mathbf{RO}_{add\,(Rj,-cRk)}$$

$$\mathbf{RO}^{-1}_{mult\,(c,Rj)} = \mathbf{RO}_{swap\,(c^{-1}Rj)}$$

- Then, obtain det($\mathbf{RO}_i^{-1}$) using det($\mathbf{RO}_i$) on the RHS

$$\det(\mathbf{RO}^{-1}_{swap\,(Rj,Rk)}) = \det(\mathbf{RO}_{swap\,(Rj,Rk)}) = (-1)$$

$$\det(\mathbf{RO}^{-1}_{add\,(Rj,cRk)}) = \det(\mathbf{RO}_{add\,(Rj,-cRk)}) = (+1)$$

$$\det(\mathbf{RO}^{-1}_{mult\,(c,Rj)}) = \det(\mathbf{RO}_{swap\,(c^{-1}Rj)}) = \frac{1}{c}$$

- Hence, det($\mathbf{RO}_i^{-1}$)=*1*/ d_i , for *i=swap, add* or *mult*.

- Previously, $\mathbf{A}^{-1} = \mathbf{RO}_L \ldots\ldots \mathbf{RO}_3\mathbf{RO}_2\mathbf{RO}_1$ **(Optional)**

  $\rightarrow \mathbf{A} = (\mathbf{RO}_1)^{-1}(\mathbf{RO}_2)^{-1} \cdots (\mathbf{RO}_L)^{-1}$

- Hence

$$\det(\mathbf{A}) = \det\left(\mathbf{RO}_1^{-1}\mathbf{RO}_2^{-1}\cdots\mathbf{RO}_L^{-1}\right) = \frac{1}{d_1}\det\left(\mathbf{RO}_2^{-1}\cdots\mathbf{RO}_L^{-1}\right)$$

$$= \frac{1}{d_1}\frac{1}{d_2}\det\left(\mathbf{RO}_3^{-1}\cdots\mathbf{RO}_L^{-1}\right) = \frac{1}{d_1 d_2 \cdots d_L}$$

- Furthermore, $\det(\mathbf{AB}) = \det\left(\mathbf{RO}_1^{-1}\mathbf{RO}_2^{-1}\cdots\mathbf{RO}_L^{-1}\mathbf{B}\right)$

$$= \frac{1}{d_1}\det\left(\mathbf{RO}_2^{-1}\cdots\mathbf{RO}_L^{-1}\mathbf{B}\right)$$

$$= \frac{1}{d_1}\frac{1}{d_2}\det\left(\mathbf{RO}_3^{-1}\cdots\mathbf{RO}_L^{-1}\mathbf{B}\right)$$

$$= \frac{1}{d_1 d_2 \cdots d_L}\det(\mathbf{B}) = \det(\mathbf{A})\det(\mathbf{B})$$

7.9 Vector Spaces, Inner Product Spaces, Linear Transformations

DEFINITION

Real Vector Space

Given a nonempty set V of elements **a**, **b**, … is called a **real vector space** (or *real linear space*) if V has two algebraic operations (called *vector addition* and *scalar multiplication*) defined as follows.

I. **Vector addition** – for every pair of vectors **a** and **b** of V, there exist a unique vector of V, called the *sum* of **a** and **b** and denoted by $\mathbf{a} + \mathbf{b}$, such that the following axioms are satisfied.

I.1 *Commutativity.* For any two vectors **a** and **b** of V,

$$\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}.$$

I.2 *Associativity.* For any three vectors **a**, **b**, **c** of V,

$$(\mathbf{a} + \mathbf{b}) + \mathbf{c} = \mathbf{a} + (\mathbf{b} + \mathbf{c}) \qquad \text{(written } \mathbf{a} + \mathbf{b} + \mathbf{c}\text{)}.$$

I.3 There is a unique vector in V, called the *zero vector* and denoted by $\mathbf{0}$, such that for every $\mathbf{a}$ in V, $\mathbf{a} + \mathbf{0} = \mathbf{a}$.

I.4 For every $\mathbf{a}$ in V, there is a unique vector in V that is denoted by $-\mathbf{a}$ and is such that $\mathbf{a} + (-\mathbf{a}) = \mathbf{0}$.

II. Scalar multiplication – for every $\mathbf{a}$ in V and every real number c there exist a unique vector of V, called the ***product*** of c and $\mathbf{a}$ and denoted by $c\mathbf{a}$ (or $\mathbf{a}c$) such that the following axioms are satisfied.

II.1 *Distributivity.* For every scalar c and vectors $\mathbf{a}$ and $\mathbf{b}$ in V, $c(\mathbf{a} + \mathbf{b}) = c\mathbf{a} + c\mathbf{b}$.

II.2 *Distributivity.* For all scalars c and k and every $\mathbf{a}$ in V,

$$(c + k)\mathbf{a} = c\mathbf{a} + k\mathbf{a}.$$

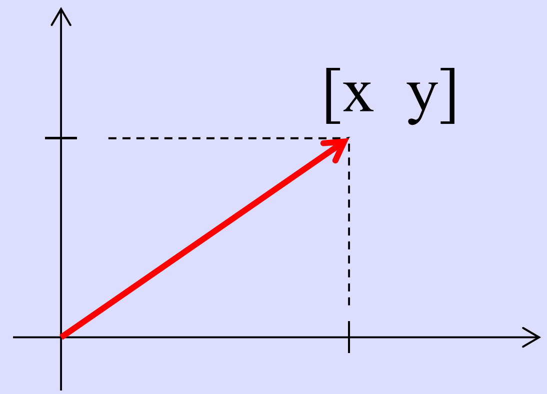
II.3 *Associativity.* For all scalars c and k and every $\mathbf{a}$ in V,

$$c(k\mathbf{a}) = (ck)\mathbf{a} \qquad \text{(written } ck\mathbf{a}\text{)}.$$

II.4 For every $\mathbf{a}$ in V, $1\mathbf{a} = \mathbf{a}$.

Examples of Real Vector Space

- Space $R^2 = \{$row vector $[x \;\; y]$, with x, y finite real numbers$\}$.



- Visualization: arrow from origin to point with coordinate $[x \;\; y]$

- Same visualization for Space $R^2 = \{$column vector $[x \;\; y]^T$, with x, y finite real numbers$\}$.

- Easily extended to Space $R^3 = \{$row vector $[x \;\; y \;\; z]$, with x, y, z finite real numbers$\}$, or generally to

  $R^n = \{$row (or column) vectors with n-components $[x_1 \; x_2 \; \cdots \; x_n]\}$

Note: Vector Space is not only comprising of Vectors

- Example 1: Vector Spaces of Matrix
  - Real 2x2 matrices form a R^4 real vector space

$$\left\{\begin{bmatrix} x_1 & x_2 \\ x_3 & x_4 \end{bmatrix}, x_i \ \text{ real number, for } i = 1,\ldots,4.\right\}$$

  - Dim=4
  - A possible Basis:

$$B_{11} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, B_{11} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, B_{21} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, B_{22} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

- Example 2: Vector space of polynomials
  - The set of polynomial of order 2 or less.

$$\left\{a_3 x^2 + a_2 x^2 + a_1,\ a_i \ \text{ real number, for } i = 1, 2, 3\right\}$$

  - Dim=3
  - A possible basis: $p_1 = 1, p_2 = x, p_3 = x^2$

- Example: Vector Spaces of Functions
  - Real value continuous functions on given interval form a real vector space

    $\{f(x), g(x), \cdots \text{ on given interval } \alpha \le x \le \beta\}$
  - Dim=infinity
  - A possible Basis for $\alpha = 0 \le x \le \beta = 2\pi$

$$\sin x, \sin(2x), \sin(3x), \cdots$$

DEFINITION Inner Product Spaces

Real Inner Product Space

A real vector space V is called a **real inner product space** (or *real pre-Hilbert space*) if for every pair of vectors **a** and **b** in V, there exist an operation called the **inner product** of **a** and **b** and denoted by $(\mathbf{a}, \mathbf{b})$. The inner product is a scalar and satisfies the following axioms.

I. For all scalars q_1 and q_2 and all vectors **a**, **b**, **c** in V,

$$(q_1\mathbf{a} + q_2\mathbf{b}, \mathbf{c}) = q_1(\mathbf{a}, \mathbf{c}) + q_2(\mathbf{b}, \mathbf{c}) \qquad (\textit{Linearity}).$$

II. For all vectors **a** and **b** in V,

$$(\mathbf{a}, \mathbf{b}) = (\mathbf{b}, \mathbf{a}) \qquad (\textit{Symmetry}).$$

III. For every **a** in V,

$$(\mathbf{a}, \mathbf{a}) \geq 0,$$

$$(\mathbf{a}, \mathbf{a}) = 0 \text{ if and only if } \mathbf{a} = \mathbf{0}$$

- Vectors whose inner product is zero are called **orthogonal** to each other.
- The *length* or **norm** of a vector in V is defined by

$$(2) \qquad \|\mathbf{a}\| = \sqrt{(\mathbf{a}, \mathbf{a})} \qquad (\geq 0).$$

- A vector of norm 1 is called a **unit vector**.

From these axioms and from (2) one can derive the basic inequality

$$(3) \qquad \|\mathbf{a}, \mathbf{b}\| \leq \|\mathbf{a}\|\|\mathbf{b}\| \qquad (\textit{Cauchy-Schwarz inequality})$$

From this follows

$$(4) \qquad \|\mathbf{a}+\mathbf{b}\| \leq \|\mathbf{a}\| + \|\mathbf{b}\| \qquad (\textit{Triangle inequality}).$$

A simple direct calculation gives

$$(5) \quad \|\mathbf{a}+\mathbf{b}\|^2 + \|\mathbf{a}-\mathbf{b}\|^2 = 2(\|\mathbf{a}\|^2 + \|\mathbf{b}\|^2) \quad (\textit{Parallelogram equality}).$$

(Proofs of (3), (4), and (5) in Chapter 9)

Example: Inner Product Space

- Space R^2 ={row vector [x y], with x, y finite real numbers}.
- Define for $\mathbf{a} = \begin{bmatrix} x_1 & y_1 \end{bmatrix}$ and $\mathbf{b} = \begin{bmatrix} x_2 & y_2 \end{bmatrix}$ their inner product as

$$(\mathbf{a}, \mathbf{b}) = x_1 x_2 + y_1 y_2$$

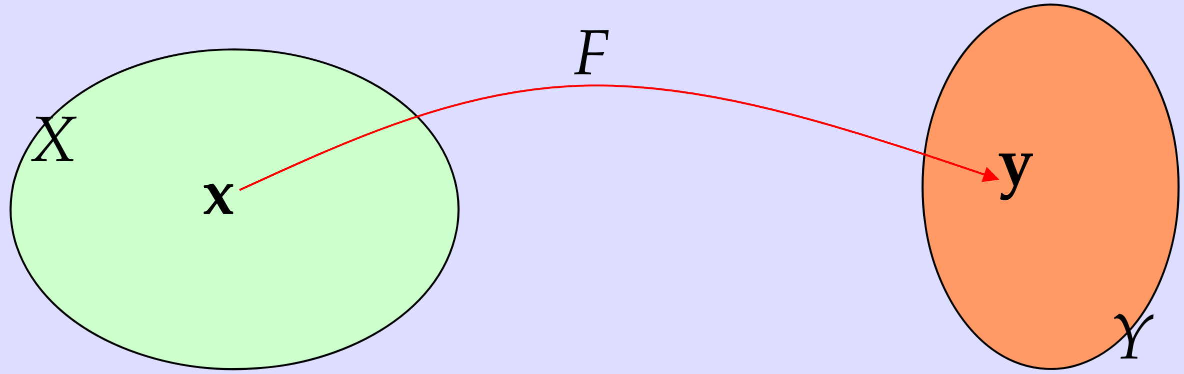
- (**a**,**b**) satisfies the conditions I, II and III in previous slide
- R^2 becomes an inner product space with (**a**,**b**)
- Similarly, R^3 and generally, R^n , are inner product spaces

- Example: Vector Spaces of Functions
  - Real value continuous functions on given interval form a real vector space
    $\{f(x), g(x), \cdots \text{ on given interval } \alpha \le x \le \beta\}$
  - A possible definition of inner product:

$$(f, g) = \int_{\alpha}^{\beta} f(x) g(x) dx$$

Linear Transformations

- Let X and Y are any vector spaces.
- A **mapping** (or **transformation** or **operator**) of X into Y defines for *any* vector $\mathbf{x}$ in X a *unique* vector $\mathbf{y}$ in Y.
- A mapping is denoted by a capital letter, say F.
- The vector $\mathbf{y}$ in Y is the **image** of $\mathbf{x}$ under F and is denoted by $F(x)$ (or just $F\mathbf{x}$).



- X is the domain of F and Y is the range of F.

Linear Transformations (continued)

F is called a **linear mapping** or **linear transformation** if, for all vectors $\mathbf{v}$ and $\mathbf{x}$ in X and scalars c,

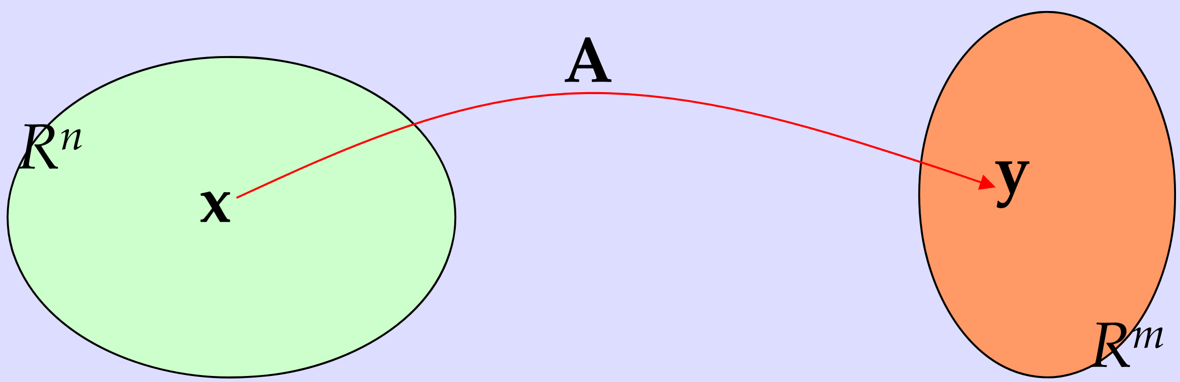
$$(10)\qquad \begin{aligned} F(\mathbf{v}+\mathbf{x}) &= F(\mathbf{v})+F(\mathbf{x}) \\ F(c\mathbf{x}) &= cF(\mathbf{x}). \end{aligned}$$

Example: Linear Transformation of Space R^n into Space R^m

- Let $X = R^n$ and $Y = R^m$. Then any real $m \times n$ matrix $\mathbf{A} = [a_{jk}]$ gives a linear transformation of R^n into R^m,

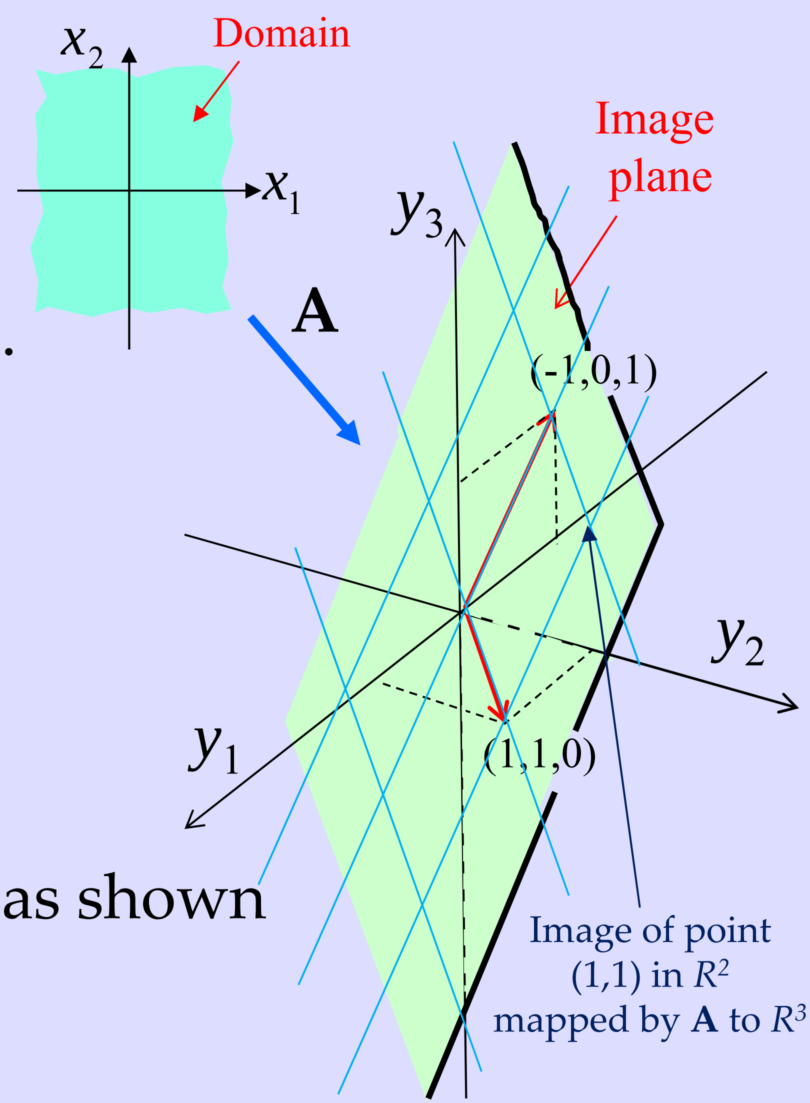
$$\mathbf{y} = \mathbf{A}\mathbf{x} \tag{11}$$

- Transformation is *linear* because $\mathbf{A}(\mathbf{u} + \mathbf{x}) = \mathbf{A}\mathbf{u} + \mathbf{A}\mathbf{x}$ and $\mathbf{A}(c\mathbf{x}) = c\mathbf{A}\mathbf{x}$.



Example: Linear transformation $\mathbf{y} = \mathbf{A}\mathbf{x}$ maps R^2 into R^3.

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



- Domain is whole $x_1 x_2$ plane.
- Image is plane spanned by columns of **A**.

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

- Image of point (1,1) in R^2 is as shown
- The point, say, (3,0,0) is not in image plane

- Note: any real $m \times n$ matrix **A** is a linear transformation of R^n into R^m,

$$\mathbf{y} = \mathbf{A}\mathbf{x} \tag{11}$$

- Actually, any linear transformation from R^n into R^m can be represented by a real $m \times n$ matrix
- **Example:** given a linear transformation F such that for any x in R^3, we know its image y in on R^2. Can we find a matrix representation for F?
  – Yes, by the following way: take the standard basis for R^3,

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \mathbf{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

– Obtain their images on R^2: $\mathbf{f}_1 = F(\mathbf{e}_1), \mathbf{f}_2 = F(\mathbf{e}_2), \mathbf{f}_3 = F(\mathbf{e}_3)$

– Then a 2×3 matrix $\mathbf{A}$ to represent F is

$$\mathbf{A} = \begin{bmatrix} \mathbf{f}_1 & \mathbf{f}_2 & \mathbf{f}_3 \end{bmatrix}$$

– Why? Given any vector in R^3 expressed in the basis of $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$,

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + x_3\mathbf{e}_3$$

– Its image $F(\mathbf{x})$ is given by (noting that F is linear transformation):

$$F\mathbf{x} = F(x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + x_3\mathbf{e}_3) = F(x_1\mathbf{e}_1) + F(x_2\mathbf{e}_2) + F(x_3\mathbf{e}_3)$$

$$= x_1F(\mathbf{e}_1) + x_2F(\mathbf{e}_2) + x_3F(\mathbf{e}_3) = x_1\mathbf{f}_1 + x_2\mathbf{f}_2 + x_3\mathbf{f}_3$$

$$= \begin{bmatrix} \mathbf{f}_1 & \mathbf{f}_2 & \mathbf{f}_3 \end{bmatrix} \mathbf{x}$$

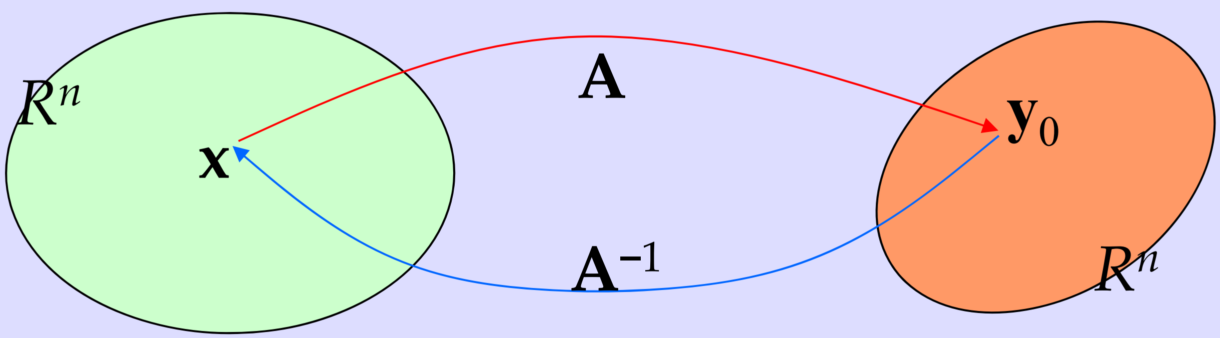
2×3 matrix $\mathbf{A}$ representing F

Example: Linear Transformation of Space R^n into Space R^n

- If **A** is square $n \times n$ matrix, then **A** maps R^n into R^n.
- If **A** is also nonsingular (so $\mathbf{A}^{-1}$ exists), then multiply (11) by $\mathbf{A}^{-1}$ from the left yields

(14) $$\mathbf{A}^{-1}\mathbf{A}\,\mathbf{x} = \mathbf{x} = \mathbf{A}^{-1}\,\mathbf{y}.$$

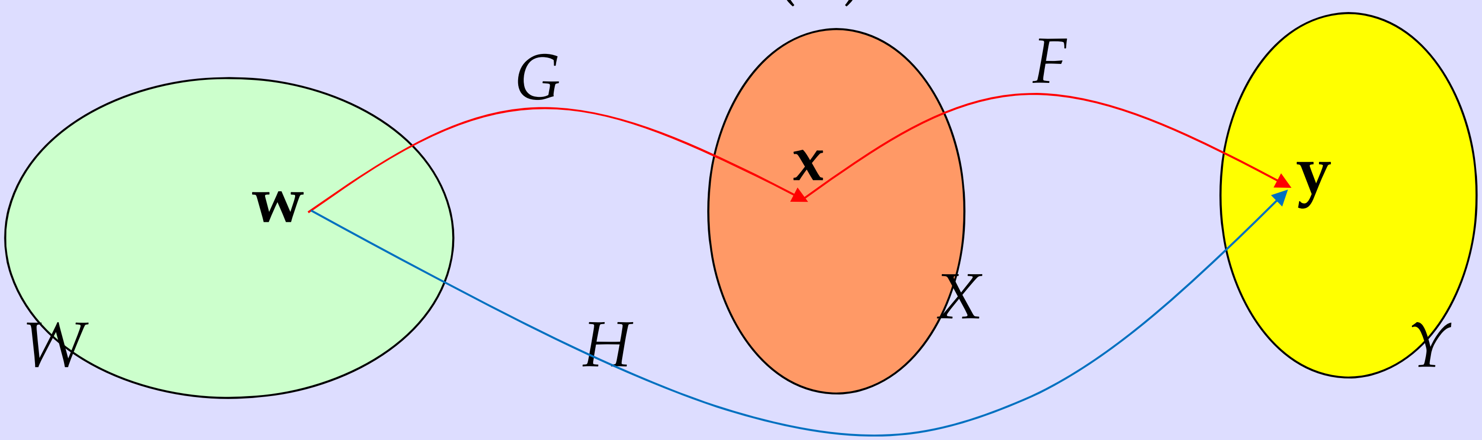
- $\mathbf{A}^{-1}$ is the **inverse transformation** – mapping every $\mathbf{y} = \mathbf{y}_0$ back onto $\mathbf{x}$
- $\mathbf{A}^{-1}$ is also a matrix and hence linear transformation (the inverse of a linear transformation is itself linear).



Composition of Linear Transformations

- Let X, Y, W be general vector spaces.
- G is a linear transformation from W to X, F is a linear transformation from X to Y.
- H, the linear transformation from W to Y, is called the **composition** of F and G,

$$H = F \circ G = FG = F(G).$$



- If F and G are linear transformations, then composition H is also linear transformation