Chapter 1: Matrices, Vectors, and Systems of Linear Equations

- Gaussian elimination
- Generating sets and linear independence

  Information about the existence and uniqueness of solutions of a system of linear equations

1.1 Matrices and Vectors

Definitions: A **matrix** is a rectangular array of *scalars*. If the matrix has m rows and n columns, the **size** of the matrix is m by n, written $m \times n$. The matrix is **square** if $m=n$. The scalar in the i-th row and j-th column is called the (i,j)-**entry** of the matrix.

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

- Let a_{ij} and b_{ij} be the (i,j)-entry of matrices A and B, respectively. A and B are **equal**, $A=B$, if they have the same size and $a_{ij} = b_{ij}$ $\forall$ i and j.
- A **submatrix** of a matrix M is obtained by deleting from M entire rows or entire columns, or both.

Matrix sums and scalar multiplication

- If A and B are $m \times n$ matrices, the **sum** of A and B, denoted by $A+B$, is the $m \times n$ matrix whose (i,j)-entry is $a_{ij} + b_{ij}$.

- Let c be a scalar. The **scalar multiple** cA is the $m \times n$ matrix whose (i,j)-entry is ca_{ij}.
  - $1A=A$
  - $(-1)A$, denoted by $-A$
  - $0A$, denoted by O, the $m \times n$ **zero matrix**, in which each entry is 0.
- **Substraction**: For any matrices A and B of the same size, define $A-B$ to be the matrix whose (i,j)-entry is $a_{ij} - b_{ij}$.
  - $A-A=O$ $\forall$ matrices A.

Theorem 1.1 (Properties of Matrix Addition and Scalar Multiplication)

Let A, B, and C be $m \times n$ matrices, and let s and t be any scalars. Then

a) $A+B = B+A$. (commutative law of matrix addition)

b) $(A+B)+C = A + (B+C)$. (associate law of matrix addition)

c) $A+O = A$.

d) $A+(-A) = O$.

e) $(st)A = s(tA)$.

f) $s(A+B) = sA + sB$.

g) $(s+t)A = sA + tA$.

Proof of Theorem 1.1(b):

- The (i, j) entry of $(A + B) + C$ is ….. $\left(a_{i,j} + b_{i,j}\right) + c_{i,j}$.
- The (i, j) entry of $\mathrm{A} + (B + C)$ is ….. $a_{i,j} + \left(b_{i,j} + c_{i,j}\right)$.

By the associate law holds for addition, +, of scalars,
$\left(a_{i,j} + b_{i,j}\right) + c_{i,j} = a_{i,j} + \left(b_{i,j} + c_{i,j}\right)$.
Therefore the (i, j) entry of $(A + B) + C$ equals the (i, j) entry of $\mathrm{A} + (B + C), \forall\ 1 \le i \le m, 1 \le j \le n$, proving (b).#

Instead, it is often written as $A + B + C$.

Examples of inventory matrices

B and *C* for July and August sales at two stores

$$B = \begin{bmatrix} 6 & 8 \\ 15 & 20 \\ 45 & 64 \end{bmatrix} \text{ and } C = \begin{bmatrix} 7 & 9 \\ 18 & 31 \\ 52 & 68 \end{bmatrix}$$

- $b_{2,1} + c_{2,1} = 15 + 18 = 33$
- $B + C = \begin{bmatrix} 13 & 17 \\ 33 & 51 \\ 97 & 132 \end{bmatrix}$
- $C^T = \begin{bmatrix} 7 & 18 & 52 \\ 9 & 31 & 68 \end{bmatrix}$
- $1B = B$; $(-1)B = -B$; $0B = \mathbf{0} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$, a zero matrix
- $\boldsymbol{b}_1 = \begin{bmatrix} 6 \\ 15 \\ 45 \end{bmatrix}$, the 1st column of *B*, a **column** vector, a **submatrix**
- $[7 \quad 9]$, the 1st row of *C*, a **row** vector, a **submatrix**

Matrix Transposes

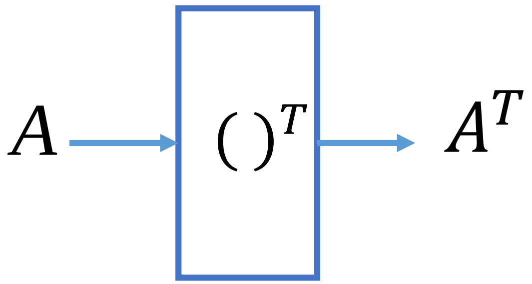
- The **transpose** of an $m \times n$ matrix A is the $n \times m$ matrix denoted by A^T whose (i,j)-entry is the (j,i)-entry of A.

Theorem 1.2 (Properties of the Transpose) Let A and B be $m \times n$ matrices. Then

a) $(A+B)^T = A^T + B^T$.

b) $(sA)^T = sA^T$.

c) $(A^T)^T = A$.



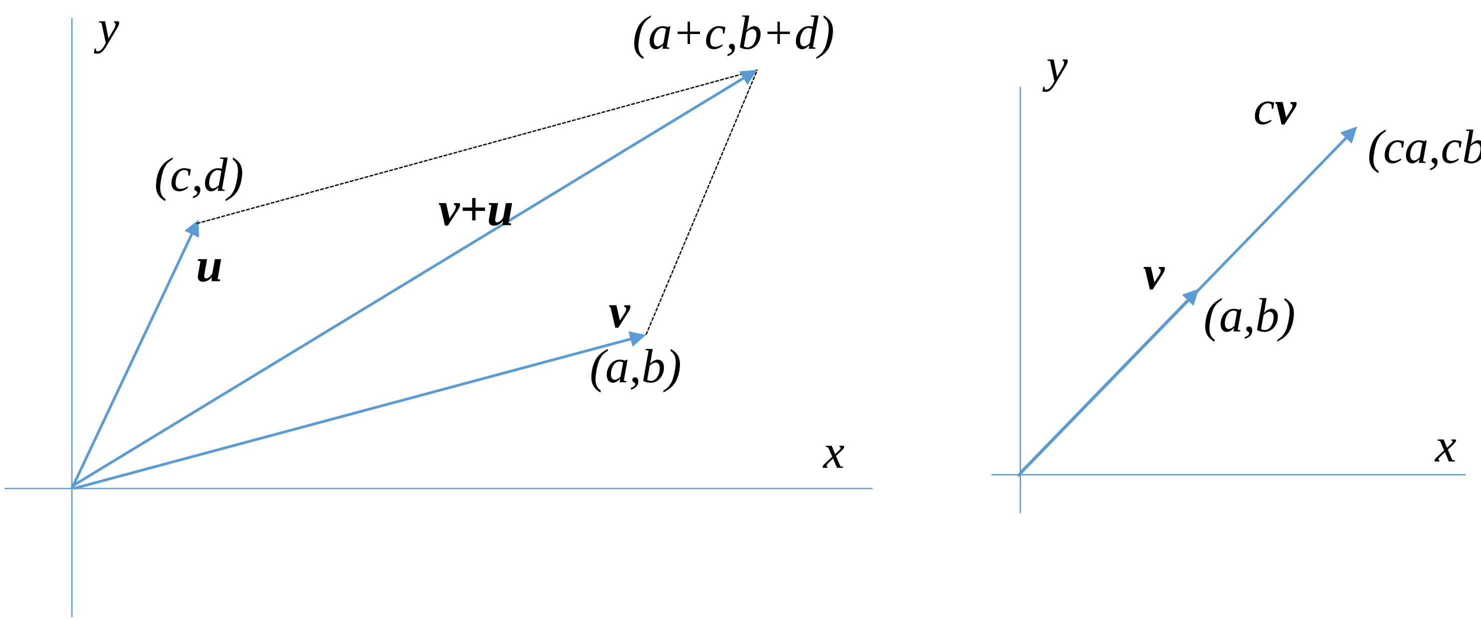
Vectors

- A matrix having exactly one row is called a **row vector.**
- A matrix having exactly one column is called a **column vector.**
- *Vector*, referring to either a row vector or a column vector.
- The entries of a vector are called **components.**
- R^n, the set of all column vectors with n components.
- Two arithmetic operations on vectors, **vector addition** and **scalar multiplication**, satisfy the properties in Theorem 1.1

Denote the i-th component of the vector $\boldsymbol{u} \in \boldsymbol{R}^n$ by u_i.

Thus, $\boldsymbol{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}$.

Occasionally vector $\boldsymbol{u}$ is identified by an n-tuple $(u_1, u_2, \ldots, u_n)$.



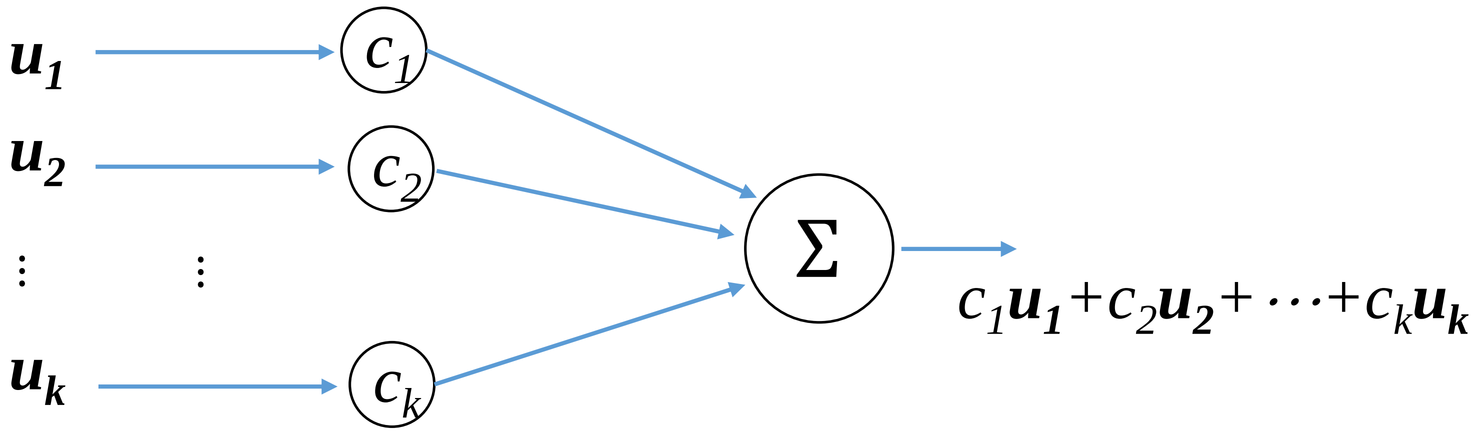
Geometry of vectors, vector addition, the parallelogram law, and scalar multiplication

1.2 Linear combinations, Matrix-vector products, and special matrices

Definitions A **linear combination** of vectors $\boldsymbol{u_1}, \boldsymbol{u_2}, \ldots, \boldsymbol{u_k}$ is a vector of the form

$$c_1\boldsymbol{u_1} + c_2\boldsymbol{u_2} + \cdots + c_k\boldsymbol{u_k},$$

where $c_1, c_2, \cdots, c_k$ are scalars. The scalars are called the **coefficients** of the linear combination.



Examples:

1. $\begin{bmatrix}2\\8\end{bmatrix} = (-3)\begin{bmatrix}1\\1\end{bmatrix} + 4\begin{bmatrix}1\\3\end{bmatrix} + 1\begin{bmatrix}1\\-1\end{bmatrix}$

$\leftarrow$ $\begin{bmatrix}2\\8\end{bmatrix}$ is a linear combination of $\boldsymbol{u_1} = \begin{bmatrix}1\\1\end{bmatrix}$, $\boldsymbol{u_2} = \begin{bmatrix}1\\3\end{bmatrix}$, and $\boldsymbol{u_3} = \begin{bmatrix}1\\-1\end{bmatrix}$ with coefficient $c_1 = -3$, $c_2 = 4$, and $c_3 = 1$.

2. Whether $\begin{bmatrix}3\\4\end{bmatrix}$ is a linear combination of $\begin{bmatrix}3\\2\end{bmatrix}$ and $\begin{bmatrix}4\\6\end{bmatrix}$?

$\leftarrow$ Is there a solution of $\begin{bmatrix}3\\4\end{bmatrix} = x_1\begin{bmatrix}3\\2\end{bmatrix} + x_2\begin{bmatrix}6\\4\end{bmatrix}$ or the system of equations $\begin{matrix}3x_1 + 6x_2 = 3\\2x_1 + 4x_2 = 4\end{matrix}$?

Standard Vectors

Define the **standard vector** of R^n by

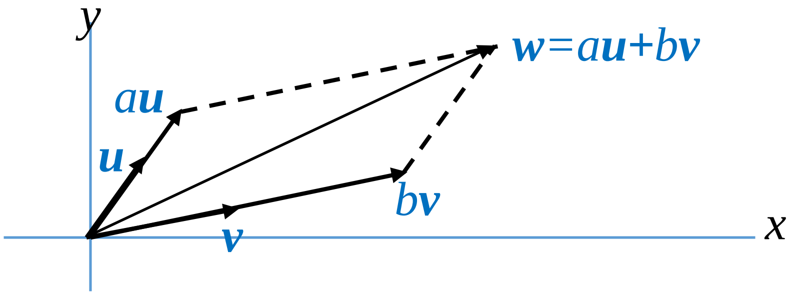
$$\boldsymbol{e}_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \boldsymbol{e}_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \cdots, \boldsymbol{e}_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}.$$

Thus, for any vector $\boldsymbol{v}$ in R^n,

$$\boldsymbol{v} = v_1\boldsymbol{e}_1 + v_2\boldsymbol{e}_2 + \cdots + v_n\boldsymbol{e}_n \ ,$$

a linear combination of standard vectors of R^n.

If $\boldsymbol{u}$ and $\boldsymbol{v}$ are any nonparallel vectors in R^2, then every vector in R^2 is a linear combination of $\boldsymbol{u}$ and $\boldsymbol{v}$.



Matrix-Vector products

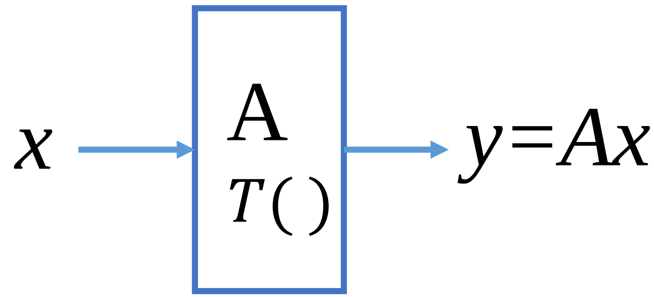
Definition Let A be an $m \times n$ matrix and $\boldsymbol{v}$ be an $n \times 1$ vector. Define the **matrix-vector product** of A and $\boldsymbol{v}$, denoted by $A\boldsymbol{v}$,

$$A\boldsymbol{v} = v_1\boldsymbol{a}_1 + v_2\boldsymbol{a}_2 + \cdots + v_n\boldsymbol{a}_n \quad \left(\text{or } A\boldsymbol{v} = \sum_{i=1}^{n} v_i\boldsymbol{a}_i\right),$$

where $\boldsymbol{a}_i$ is the i-th column of A and v_i is the i-th component of $\boldsymbol{v}$.

$$A\boldsymbol{v} = [\boldsymbol{a}_1 \;\; \boldsymbol{a}_2 \;\; \cdots \;\; \boldsymbol{a}_n]\begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}\begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

$$= \begin{bmatrix} a_{11}v_1 + a_{12}v_2 + \cdots + a_{1n}v_n \\ a_{21}v_1 + a_{22}v_2 + \cdots + a_{2n}v_n \\ \vdots \\ a_{m1}v_1 + a_{m2}v_2 + \cdots + a_{mn}v_n \end{bmatrix}$$



A linear system without perturbation

Thus, the i-th component of $A\boldsymbol{v}$ is $\sum_{j=1}^{n} a_{ij}v_j$.

“Stochastic matrix” ? (page 21)

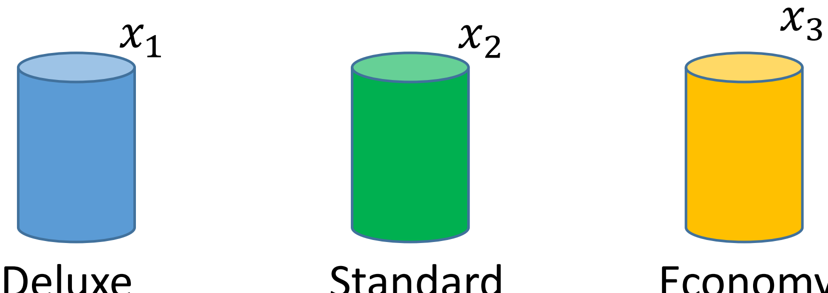
Two examples

1. A garden supply store sells three mixtures of grass seed

$$\begin{array}{cc} & \begin{array}{ccc} \mathit{deluxe} & \mathit{standard} & \mathit{economy} \end{array} \\ B = & \begin{bmatrix} .80 & .60 & .40 \\ .20 & .40 & .60 \end{bmatrix} \begin{array}{c} \mathit{bluegrass} \\ \mathit{rye} \end{array} \end{array}$$

You purchase a blend of grass seed containing 5 lb of bluegrass and 3 lb of rye. …. A possible solution is to solve a system of two linear equations in three unknowns

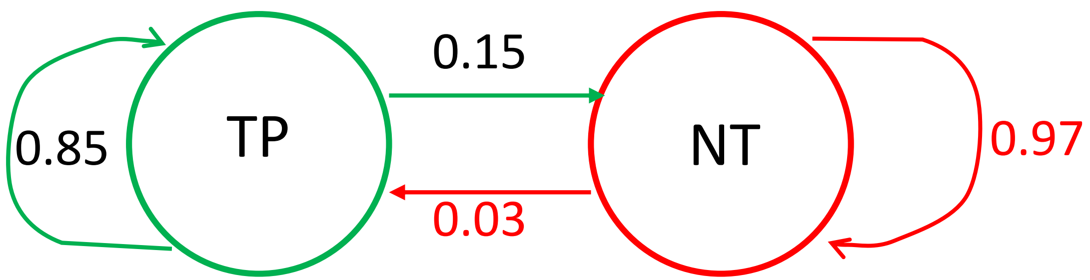
$$B\boldsymbol{x} = \begin{bmatrix} 5 \\ 3 \end{bmatrix},\ B\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 5 \\ 3 \end{bmatrix} \text{ or } \begin{bmatrix} .80x_1 + .60x_2 + .40x_3 \\ .20x_1 + .40x_2 + .60x_3 \end{bmatrix} = \begin{bmatrix} 5 \\ 3 \end{bmatrix}$$



2. The population change between Cities TP and NT is represented by a stochastic matrix $A = \begin{bmatrix} .85 & .03 \\ .15 & .97 \end{bmatrix}$.

Let $\boldsymbol{p} = \begin{bmatrix} 250 \\ 400 \end{bmatrix}$ be the vector of current populations of TP and NT.

- The population in the next year is $A\boldsymbol{p} = \begin{bmatrix} .85 & .03 \\ .15 & .97 \end{bmatrix} \begin{bmatrix} 250 \\ 400 \end{bmatrix} = \begin{bmatrix} 224.5 \\ 425.5 \end{bmatrix}$
- The population in two years is $A(A\boldsymbol{p}) = A^2\boldsymbol{p} = \begin{bmatrix} 203.59 \\ 446.41 \end{bmatrix}$
- $\lim_{\boldsymbol{n} \to \infty} A^{\boldsymbol{n}} \boldsymbol{p} = ?$



A two-state Markov Chain

Identity Matrices

Definition For each positive integer n, the $n \times n$ **identity matrix** I_n is the $n \times n$ matrix whose respective columns are the standard vectors $\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n$ in R^n.

$$I_2 = [\boldsymbol{e}_1 \quad \boldsymbol{e}_2] = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}; \; I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}; \ldots; \; I_n = [\boldsymbol{e}_1 \quad \boldsymbol{e}_2 \quad \cdots \quad \boldsymbol{e}_n]$$

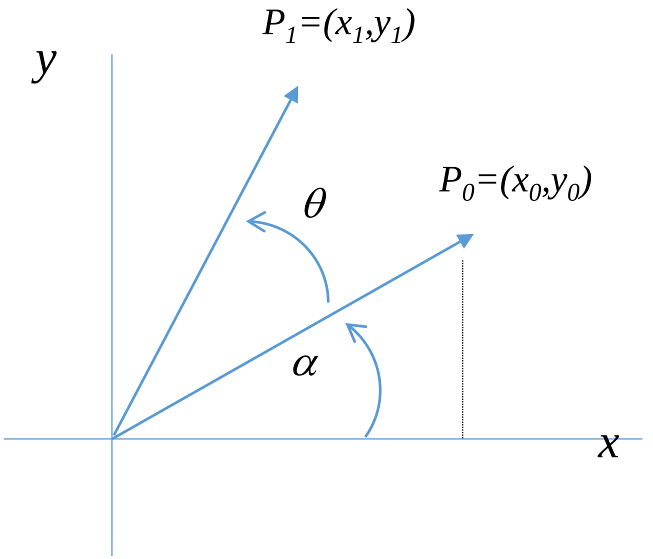
ROTATION Matrices

- Consider a point $P_0=(x_0,y_0)$ in R^2 with polar coordinate (r, α), where $r \geq 0$ and α is the angle between the segment $\overline{0P_0}$ and the positive x-axis.
- Suppose that $\overline{0P_0}$ is rotated by an angle θ to the segment $\overline{0P_1}$, where $P_1=(x_1, y_1)$. Then, $(r, \alpha+\theta)$ is the polar coordinate for P_1.

Thus, $\begin{aligned} x_0 &= r\cos\alpha \\ y_0 &= r\sin\alpha \end{aligned}$

and, $\begin{aligned} x_1 &= r\cos(\alpha+\theta) = r\cos\alpha\cos\theta - r\sin\alpha\sin\theta = x_0\cos\theta - y_0\sin\theta \\ y_1 &= r\sin(\alpha+\theta) = r\sin\alpha\cos\theta + r\cos\alpha\sin\theta = y_0\cos\theta + x_0\sin\theta \end{aligned}$

Therefore $\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x_0 \\ y_0 \end{bmatrix}$



Define $A_\theta = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$

which is called the θ-**rotation matrix**, or simply a **rotation matrix**.

Theorem 1.3 (Properties of Matrix-Vector product) Let A and B be an $m \times n$ matrices, and let $\boldsymbol{u}$ and $\boldsymbol{v}$ be vectors in R^n. Then

a) $A(\boldsymbol{u}+\boldsymbol{v}) = A\boldsymbol{u} + A\boldsymbol{v}$.
b) $A(c\boldsymbol{u}) = c(A\boldsymbol{u}) = (cA)\boldsymbol{u}$ for every scalar c.
c) $(A+B)\boldsymbol{u} = A\boldsymbol{u} + B\boldsymbol{u}$.
d) $A\boldsymbol{e}_j = \boldsymbol{a}_j$ for $j=1,2, \cdots, n$ where $\boldsymbol{e_j}$ is the j-th standard vector in R^n.
e) If B is an $m \times n$ matrix such that $B\boldsymbol{w} = A\boldsymbol{w}$ for all $\boldsymbol{w}$ in R^n, then $B=A$.
f) $A\boldsymbol{0}$ is the $m \times 1$ zero vector.
g) If O is the $m \times n$ zero matrix, then $O\boldsymbol{v}$ is the $m \times 1$ zero vector.
h) $I_n\boldsymbol{v} = \boldsymbol{v}$.

For any $m \times n$ matrix A, any scalars, c_1, c_2, $\cdots$, c_k, and any vector $\boldsymbol{u}_1$, $\boldsymbol{u}_2$, $\cdots$, $\boldsymbol{u}_k$ in R^n,

$$A(c_1\boldsymbol{u}_1 + c_2\boldsymbol{u}_2 + \cdots c_k\boldsymbol{u}_k) = c_1A\boldsymbol{u}_1 + c_2A\boldsymbol{u}_2 + \cdots + c_kA\boldsymbol{u}_k.$$

Proof of Theory 1.3 (a) : $A(\boldsymbol{u}+\boldsymbol{v}) = A\boldsymbol{u} + A\boldsymbol{v}$

The i-th component of $\boldsymbol{u}+\boldsymbol{v}$ is $u_i + v_i$.

Then

$A(\boldsymbol{u}+\boldsymbol{v}) = \sum_{i=1}^{n}(u_i + v_i)\boldsymbol{a}_i$ ← The def of Matrix-Vector product

$= \sum_{i=1}^{n}(u_i\boldsymbol{a}_i + v_i\boldsymbol{a}_i)$ by Theorem 1.1 (g)

$= \sum_{i=1}^{n} u_i\boldsymbol{a}_i + \sum_{i=1}^{n} v_i\boldsymbol{a}_i$ by Theorem 1.1 (a) and (b)

$= A\boldsymbol{u} + A\boldsymbol{v}$ by the def of Matrix-Vector product

Instead, we'll write $A\vec{u} + A\vec{v}$ on the blackboard of classroom.

1.3 Systems of Linear Equations

- A **linear equation** in the variables $x_1, x_2, \cdots, x_n$:

$$a_1x_1 + a_2x_2 + \cdots + a_xx_n = b\ ,$$

where $a_1, a_2, \cdots, a_n$ and b are real numbers.

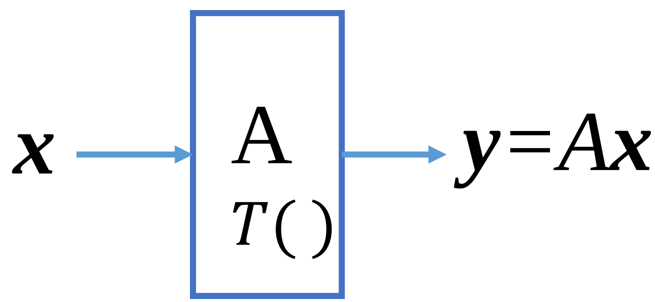
  - ✓ The scalars $a_1, a_2, \cdots, a_n$ are the **coefficients.**
  - ✓ b is the **constant term** of the equation.

- A **system of linear equations** is a set of m linear equations in the same n variables:

$$\begin{aligned}
a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\
a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\
&\vdots \\
a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m
\end{aligned}$$

That is, $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$

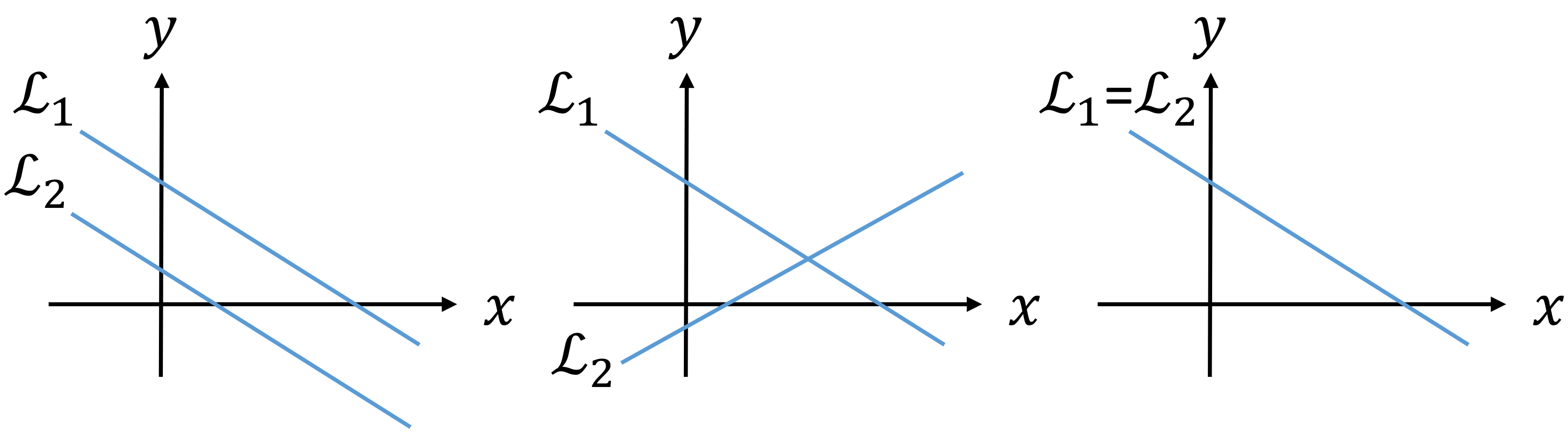
- ✓ A **solution** of a system of linear equations in the variables $x_1, x_2, \cdots, x_n$ is a vector $\begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ s_n \end{bmatrix}$ in R^n such that every equation in the system is satisfied when each x_i is replaced by s_i.
- ✓ The set of all solutions of a system of linear equations is called the **solution set** of that system.



A linear system without perturbation

Systems of 2 Linear Equations in 2 Variables:

$$Line\ \mathcal{L}_1: \quad a_1x + b_1y = c_1$$
$$Line\ \mathcal{L}_2: \quad a_2x + b_2y = c_2$$



Every system of linear equations has *no solution*, *exactly one solution*, or *infinitely many solutions*.

- A system of linear equations that has one or more solutions is called **consistent**;
- Otherwise, it is called **inconsistent**.

Elementary row operations

- Two systems of linear equations that have exactly the same solutions are called **equivalent**.
- A system of linear equations can be expressed as the matrix equation $A\boldsymbol{x}=\boldsymbol{b}$.
  - ✓ A is called the **coefficient matrix** (or the matrix of coefficients).
  - ✓ $[A\ \boldsymbol{b}]=\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{bmatrix}$, called the **augmented matrix** of the system.

Two elementary row operations:

$$\begin{aligned} x_1 - 2x_2 - x_3 &= 3 \quad \cdots \ (1) \\ 3x_1 - 6x_2 - 5x_3 &= 3 \quad \cdots \ (2) \\ 2x_1 - x_2 + x_3 &= 0 \quad \cdots \ (3) \end{aligned}$$

1. $(-3) \times \mathrm{eqn}(1) + \mathrm{eqn}(2)$ replaces $\mathrm{eqn}(2)$
2. $(-2) \times \mathrm{eqn}(1) + \mathrm{eqn}(3)$ replaces $\mathrm{eqn}(3)$

- Finishing steps (1) and (2), we have

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

3. Interchanging eqn (5) and eqn (6), we have

$$\begin{aligned} x_1 - 2x_2 - x_3 &= 3 \quad \cdots \ (7) \\ 0 + 3x_2 + 3x_3 &= -6 \quad \cdots \ (8) \\ 0 - 0 - 2x_3 &= -6 \quad \cdots \ (9) \end{aligned}$$

Associated operations on Augmented matrix

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

1. $.-3\boldsymbol{r}_1 + \boldsymbol{r}_2 \rightarrow \boldsymbol{r}_2$
2. $-2\boldsymbol{r}_1 + \boldsymbol{r}_3 \rightarrow \boldsymbol{r}_3$

……

$$\begin{bmatrix} 1 & -2 & -1 & 3 \\ 0 & 0 & -2 & -6 \\ 0 & 3 & 3 & -6 \end{bmatrix}$$

3. $\boldsymbol{r}_2 \leftrightarrow \boldsymbol{r}_3$

$$\begin{bmatrix} 1 & -2 & -1 & 3 \\ 0 & 3 & 3 & -6 \\ 0 & 0 & -2 & -6 \end{bmatrix}$$

The elementary matrix for step 3 is $E = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$.

Definition Three **elementary row operations** on a matrix:

1. Interchange any two rows of the matrix. (**interchange operation**)
2. Multiply every entry of some row of the matrix by the same nonzero scalar. (**scaling operation**)
3. Add a multiple of one row of the matrix to another row. (**row addition operation**)

- Every elementary row operation can be reversed.
  - ✓ Suppose that an elementary row operation is performed on an augmented matrix $[A\ \boldsymbol{b}]$ to obtain a new matrix $[A'\ \boldsymbol{b}']$.
    By the reversibility of row operation,
    the solutions of $A\mathbf{x}=\boldsymbol{b}$ are the same as those of $A'\mathbf{x}=\boldsymbol{b}'$.
- *Each elementary row operation produces the augmented matrix of an equivalent system of linear equations*.

Reduced Row Echelon Form

To solve a system of linear equations:

1. Represent the system by its augmented matrix.
2. Use elementary row operations to transform the augmented matrix into a matrix having a special form, called *reduced row echelon form*.
   - ✓ The system of linear equations whose augmented matrix has this form is *equivalent* to the original system and is easily solved.

- **Zero row**, a row of a matrix with all 0 entries.
- **Nonzero row**, otherwise↑.
- **Leading entry**, the leftmost nonzero entry of a nonzero row.

- Matrix M is in reduced row echelon form

$$M = \begin{bmatrix} 1 & 0 & * & * & 0 & 0 & * \\ 0 & 1 & * & * & 0 & 0 & * \\ 0 & 0 & 0 & 0 & 1 & 0 & * \\ 0 & 0 & 0 & 0 & 0 & 1 & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

- Matrices A and B are not in reduced row echelon forms

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

  - ✓ B is in row echelon form.

Definitions A matrix is in **row echelon form** if it satisfies the following three conditions:

1. Each nonzero row lies above every zero row.
2. The leading entry of a nonzero row lies in a column to the right of the column containing the leading entry of any preceding row.
3. If a column contains the leading entry of some row, then all entries of that column below the leading entry are 0. (a result of 2)

If a matrix also satisfies the following two additional conditions, it is in **reduced row echelon form**.

4. If a column contains the leading entry of some row, then all the other entries of that column are 0.
5. The leading entry of each nonzero row is 1.

Consider the system of linear equation

$$\begin{aligned} x_1 - 3x_2 \qquad + 2x_4 \qquad &= 7 \\ x_3 + 6x_4 \qquad &= 9 \\ x_5 \quad &= 2 \\ 0 \quad &= 0. \end{aligned}$$

- Its augmented matrix is

$$\begin{bmatrix} 1 & -3 & 0 & 2 & 0 & 7 \\ 0 & 0 & 1 & 6 & 0 & 9 \\ 0 & 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$, which is in reduced row echelon form.

✓ Basic variables: x_1, x_3 and x_5

✓ Free variables: x_2 and x_4

✓ The **general solution** is

$$\begin{aligned} x_1 &= 7 + 3x_2 - 2x_4 \\ x_2 &\quad \text{free} \\ x_3 &= 9 - 6x_4 \\ x_4 &\quad \text{free} \\ x_5 &= 2 \end{aligned}$$

, written in **vector form**

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 7 + 3x_2 - 2x_4 \\ x_2 \\ 9 - 6x_4 \\ x_4 \\ 2 \end{bmatrix} = \begin{bmatrix} 7 \\ 0 \\ 9 \\ 0 \\ 2 \end{bmatrix} + x_2 \begin{bmatrix} 3 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -2 \\ 0 \\ -6 \\ 1 \\ 0 \end{bmatrix}.$$

Some terms:

- **Basic variables**, corresponding to the leading entries of the augmented matrix.
- **Free variables, ….**
- **General solution**

Whenever an augmented matrix contains a row in which the only nonzero entry lies in the last column, the corresponding system of linear equations has no solution.

Theorem 1.4 Every matrix can be transformed into *one and only one* matrix in reduced row echelon form by means of a sequence of elementary row operations.

Let R be the reduced row echelon form of A.

Procedure for solving a system of linear equations

1. Write the augmented matrix $[A\ \boldsymbol{b}]$ of the system.

2. Find the reduced row echelon form $[R\ \boldsymbol{c}]$ of $[A\ \boldsymbol{b}]$.

3. If $[R\ \boldsymbol{c}]$ contains a row in which the only nonzero entry lies in the last column, then $A\boldsymbol{x}=\boldsymbol{b}$ has *no solution*.
   Otherwise, the system has *at least one solution*:
   - ✓ Write the system of linear equations corresponding to matrix $[R\ \boldsymbol{c}]$, and solve the system for the basic variables in terms of the free variables to obtain a general solution of $A\boldsymbol{x}=\boldsymbol{b}$.

1.4 Gaussian Elimination

Let R be the reduced row echelon form of a matrix A.

Terminology :

- The first nonzero entry in a nonzero row of R is the ***leading entry*** of that row.
- The positions that contain the leading entries of the nonzero rows of R are called the **pivot positions** of A.
- A column of A that contains some pivot position of A is called a **pivot column** of A.

1st pivot position

2nd pivot position

3rd pivot position

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

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

pivot rows

Pivot columns

Gaussian elimination algorithm, to find the reduced row echelon form of an nonzero matrix:

1. Determine the leftmost nonzero column, a pivot column. The topmost position in this column is a pivot position.
2. In the pivot column, choose any nonzero entry in a row that is not above the pivot row, and perform the appropriate row interchange to bring this entry into the pivot position.
3. Add an appropriate multiple of the row containing the pivot position to each lower row in order to change each entry below the pivot position into 0.
4. Ignore the row containing the pivot position and all rows above it. If there is a nonzero row that is not ignored, repeat steps 1—4 on the *submatrix* that remains.

(Continue to next page)

The next two steps transform a matrix in row echelon form into a matrix in *reduced* row echelon form. Unlike steps 1—4 which started at the top of the matrix and worked down (called *forward pass*), steps 5 and 6 *start at the last nonzero row* of the matrix and work up (called *backward pass*).

5. If the leading entry of the row is not 1, perform the appropriate scaling operation to make it 1. Then add an appropriate multiple of this row to every preceding row to change each entry above the pivot position into 0.
6. If step 5 was performed on the first row, stop. Otherwise, repeat step 5 on the preceding row.

The forward pass of Gaussian elimination,

1. Determine the first pivot column

Pivot position

$$\begin{bmatrix} 0 & 0 & 2 & -4 & -5 & 2 & 5 \\ 0 & 1 & -1 & 1 & 3 & 1 & -1 \\ 0 & 6 & 0 & -6 & 5 & 16 & 7 \end{bmatrix}$$

1^{st} pivot column

2.&3. Row interchange …. $\boldsymbol{r}_1 \leftrightarrow \boldsymbol{r}_2$ and then $(-6)\boldsymbol{r}_1+\boldsymbol{r}_3 \rightarrow \boldsymbol{r}_3$

Pivot position

$$\begin{bmatrix} 0 & 1 & -1 & 1 & 3 & 1 & -1 \\ 0 & 0 & 2 & -4 & -5 & 2 & 5 \\ 0 & 0 & 6 & -12 & -13 & 10 & 13 \end{bmatrix}$$

2^{nd} pivot column

4. Repeat

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

3^{rd} pivot column

$$\begin{bmatrix} 0 & 1 & 0 & -1 & 0 & 1 & 2 \\ 0 & 0 & 1 & -2 & 0 & 6 & 0 \\ 0 & 0 & 0 & 0 & 1 & 2 & -1 \end{bmatrix}$$, finished by the backward pass of GE, with rank=3, nullity=4.

The rank and nullity of a matrix

Definition The **rank** of an $m\times n$ matrix A, denoted by rank A, is the number of nonzero rows in the reduced row echelon form of A. The **nullity** of A, denoted by nullity A, is $n-$rank A.

- The rank of a matrix equals the number of pivot columns in the matrix.
- The nullity of a matrix equals the number of nonpivot columns in the matrix.

Thus, in a matrix in reduced row echelon form with rank k, the standard vectors $\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_k$ must appear, in order, among the pivot columns of the matrix.

If an $n \times n$ matrix has rank n, then its reduced row echelon form is I_n.

$$[\boldsymbol{e}_1 \quad \boldsymbol{e}_2 \quad \cdots \quad \boldsymbol{e}_n] = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} = I_n$$

Let $[R\ \boldsymbol{c}]$ be the reduced row echelon form of $[A\ \boldsymbol{b}]$. Then

- The system $R\boldsymbol{x}=\boldsymbol{c}$ is equivalent to $A\boldsymbol{x}=\boldsymbol{b}$.
- Each basic variable of the system $A\boldsymbol{x}=\boldsymbol{b}$ corresponds to the leading entry of exactly one nonzero row of $[R\ \boldsymbol{c}]$.
- If n is the number of columns of A, the number of free variables of $A\boldsymbol{x}=\boldsymbol{b}$ equals n minus the number of basic variables.

If $A\boldsymbol{x}=\boldsymbol{b}$ is the matrix form of a consistent system of linear equations, then

a) the number of basic variables in a general solution of the system equals the rank of A;
b) the number of free variables in a general solution of the system equals the nullity of A.

- Thus a consistent system of linear equations has a unique solution iff the nullity of its coefficient matrix equals 0.
- Equivalently, a consistent system of linear equations has infinitely many solutions iff the nullity of its coefficient matrix is positive.

(Numbers of *redundant/nonredundant* equations $\leftrightarrow$ rank $[A\ \boldsymbol{b}]$?)

An example:

A system of linear equations:

$$\begin{aligned} x_1 \quad + \quad x_2 + \quad x_3 &= 1 \\ x_1 \qquad\qquad + 3x_3 &= -2 + s \\ x_1 - \ x_2 \quad + \ rx_3 &= 3 \end{aligned}$$

- Its augmented matrix and row echelon form:

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 0 & 3 & -2+s \\ 1 & -1 & r & 3 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & -1 & 2 & -3+s \\ 0 & 0 & r-5 & 8-2s \end{bmatrix}$$

a) If $r = 5$ and $s \neq 4$, the original system is inconsistent.

b) If $r = 5$ and $s = 4$, the original system is consistent. Its general solution has a free variable. (infinitely many solutions)

c) If $r \neq 5$, the original system has a unique solution. (consistent, rank A=3)

Theorem 1.5 (Test for consistency)

The following conditions are equivalent:

a) The matrix equation $A\boldsymbol{x}=\boldsymbol{b}$ is consistent.

b) The vector $\boldsymbol{b}$ is a linear combination of the columns of A.

c) The reduced row echelon form of the augmented matrix $[A\ \boldsymbol{b}]$ has no row of the form $[0\ 0\ \cdots\ 0\ d]$, where $d \neq 0$.

PROOF in outline:

1. (a) iff (b), by the definition of Matrix-vector product:

$$\exists \boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_v \end{bmatrix} \text{ such that } A\boldsymbol{x} = \boldsymbol{b} \text{ iff } x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 + \cdots + x_n\boldsymbol{a}_n = \boldsymbol{b}.$$

2. (a) iff (c) , by the following two sub-proofs
   - ✓ (a) → (c) , if (a) then (c), through proving $\overline{\text{(c)}} \rightarrow \overline{\text{(a)}}$ instead.
   - ✓ (c) → (a).

1.6 The span of a set of vectors

Definition For a nonempty set $S=\{\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_k\}$ of vectors in R^n, define the **span** of S to be the set of all linear combinations of $\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_k$ in R^n. This set is denoted by Span S or Span$\{\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_k\}$.

- The span of $\{\boldsymbol{u}\}$ consists of all multiple of $\boldsymbol{u}$.
- The span of $\{\boldsymbol{0}\}$ is $\{\boldsymbol{0}\}$.
- Span $\{\boldsymbol{e}_1, \boldsymbol{e}_2\} = xy$_plane.

Let $S=\{\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_k\}$ be a set of vectors from R^n, and let A be a matrix whose columns are $\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_k$. Then a vector $\boldsymbol{v}$ from R^n is in the span S iff the equation $A\boldsymbol{x}=\boldsymbol{v}$ is consistent.

$\boldsymbol{v} \in$ span S iff $x_1\boldsymbol{u}_1 + x_2\boldsymbol{u}_2 + \cdots + x_k\boldsymbol{u}_k = \boldsymbol{v}$, i.e. $A\boldsymbol{x} = \boldsymbol{v}$, is consistent.

Examples:

a) Is $\boldsymbol{v} = \begin{bmatrix} 2 \\ 1 \\ 3 \\ -1 \end{bmatrix}$ in the span of $S = \left\{ \begin{bmatrix} 1 \\ 2 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ -2 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 8 \\ -1 \\ 5 \end{bmatrix} \right\}$?

$\Rightarrow$ Is $A\boldsymbol{x} = \boldsymbol{v}$ consistent?

$$[A\ \boldsymbol{v}] = \begin{bmatrix} 1 & -1 & 1 & 2 \\ 2 & 1 & 8 & 1 \\ 1 & -2 & -1 & 3 \\ 1 & 1 & 5 & -1 \end{bmatrix} \to [R\ \boldsymbol{c}] = \begin{bmatrix} 1 & 0 & 3 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \Longleftarrow \text{No.}$$

b) Let $S = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -2 \\ -1 \end{bmatrix} \right\}$. Show that Span $S = \mathcal{R}^3$.

$\Rightarrow$ Show that $A\boldsymbol{x} = \boldsymbol{v}$ is consistent for every $\boldsymbol{v} \in \mathcal{R}^3$.

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & -2 \\ 0 & 0 & 1 & -1 \end{bmatrix} \to R = \begin{bmatrix} 1 & 0 & 0 & 3 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -1 \end{bmatrix} \Longleftarrow \text{Yes, by Theorem 1.5}$$

- Previous problems:
  Start with a subset S of R^n and then describe V=Span S.

- The opposite problem:
  Start with a set V and then find a set of vectors S for which Span S=V.
  - ✓ If V is a set of vectors from R^n and Span S=V, then S is a **generating set** for V or S **generates** V.

Theorem 1.6

The following statements about an $m \times n$ matrix A are equivalent:

a) The span of the columns of A is R^m.

b) The equation $A\boldsymbol{x}=\boldsymbol{b}$ has at least one solution (that is, $A\boldsymbol{x}=\boldsymbol{b}$ is consistent) for each $\boldsymbol{b}$ in R^m.

c) The rank of A is m, the number of rows of A.

d) The reduced row echelon form of A has no zero rows.

e) There is a pivot position in each row of A.

- a) and b) are equivalent by Theorem 1.5
- c) , d), and e) are equivalent. ….
- Then, prove b) and c) are equivalent ?

  $\overline{c)} \to \overline{b)}$ by … $[R \quad \boldsymbol{e}_m] \to [A \quad d]$

  $c) \to b)$ by … $[R \quad \boldsymbol{c}] \to [A \quad \boldsymbol{b}]$

$$A = \begin{array}{c c} & \leftarrow \quad n \quad \rightarrow \\ m\left\{ \begin{array}{c} \\ \\ \\ \\ \end{array} \right. & \begin{bmatrix} * & * & \cdots & * \\ * & * & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ * & * & \cdots & * \end{bmatrix} \end{array}$$

<Errata in the proof of Theorem 1.6, on page 71 : …. R must have no ~~non~~zero rows. …. R has no ~~non~~zero rows >

Theorem 1.7

Let $S=\{\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_k\}$ be a set of vectors from R^n, and let $\boldsymbol{v}$ be a vector in R^n. Then Span $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_k, \boldsymbol{v}\}$ = Span $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_k\}$ if and only if $\boldsymbol{v}$ belongs to the span of S.

Proof:

Let $S_A = \text{Span}\{\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_k, \boldsymbol{v}\}$ and $S_B = \text{Span}\{\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_k\}$.

- Prove $\boldsymbol{v} \in S_B \rightarrow S_A = S_B$

  Suppose that $\boldsymbol{v} \in S_B$. Then $\boldsymbol{v} = \sum_{i=1}^{k} a_i \boldsymbol{u}_i$ for some scalars $a_1, a_2, \cdots, a_k$.

  ✓ If $\boldsymbol{w} \in S_A$, then $\boldsymbol{w} = \sum_{i=1}^{k} c_i \boldsymbol{u}_i + b\boldsymbol{v}$ for some scalars $c_1, c_2, \cdots, c_k, b$. ....

  $\boldsymbol{w} = \sum_{i=1}^{k} c_i \boldsymbol{u}_i + b \sum_{i=1}^{k} a_i \boldsymbol{u}_i = \sum_{i=1}^{k} (c_i + b a_i) \boldsymbol{u}_i$. Thus $\boldsymbol{w} \in S_B$. $\Rightarrow S_A \subset S_B$.

  ✓ Any vector $\boldsymbol{q} = \sum_{i=1}^{k} d_i \boldsymbol{u}_i \in S_B$ can be written as $\boldsymbol{q} = \sum_{i=1}^{k} d_i \boldsymbol{u}_i + 0\boldsymbol{v}$.

  Then $\boldsymbol{q} \in S_A$. $\Rightarrow S_B \subset S_A$.

  It follows that $S_A = S_B$.

- Prove $\boldsymbol{v} \notin S_B \rightarrow S_A \neq S_B$, instead of proving $S_A = S_B \rightarrow \boldsymbol{v} \in S_B$.

  ✓ Suppose that $\boldsymbol{v} \notin S_B$.

  Since $\boldsymbol{v} = \sum_{i=1}^{k} 0\boldsymbol{u}_i + 1\,\boldsymbol{v}$, $\boldsymbol{v} \in S_A$.

  Therefore $S_A \neq S_B$.

1.7 Linear dependence and linear independence

Definitions A set of k vectors $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_k\}$ in R^n is called **linearly dependent** if there exists scalars $c_1, c_2, \cdots, c_k$, not all 0, such that

$$c_1 \boldsymbol{u}_1 + c_2 \boldsymbol{u}_2 + \cdots + c_k \boldsymbol{u}_k = \boldsymbol{0}.$$

In this case, we also say that **the vectors $\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_k$ are linearly dependent**.

A set of k vectors $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_k\}$ is called **linearly independent** if the only scalars $c_1, c_2, \cdots, c_k$ such that

$$c_1 \boldsymbol{u}_1 + c_2 \boldsymbol{u}_2 + \cdots + c_k \boldsymbol{u}_k = \boldsymbol{0}$$

are $c_1 = c_2 = \cdots = c_k = 0$. In this case, we also say that **the vectors $\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_k$ are linearly independent**.

An example:

Is the set $S = \left\{ \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 4 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \right\}$ linearly dependent or linearly independent?

$\Rightarrow$ Determine whether $A\boldsymbol{x} = \mathbf{0}$ has a nonzero solution, where

$$[A\ \mathbf{0}] = \begin{bmatrix} 1 & 1 & 1 & 1 & 0 \\ 2 & 0 & 4 & 2 & 0 \\ 1 & 1 & 1 & 3 & 0 \end{bmatrix} \rightarrow [R\ \mathrm{c}] = \begin{bmatrix} 1 & 0 & 2 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}.$$

The general solution is $\begin{aligned} x_1 &= -2x_3 \\ x_2 &= \quad\ x_3 \\ x_3 & \quad\ \text{free} \\ x_4 &= \quad\ 0. \end{aligned}$ or $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = x_3 \begin{bmatrix} -2 \\ 1 \\ 1 \\ 0 \end{bmatrix}.$

Thus S is a linearly dependent subset of $\mathcal{R}^3$.

- Any finite subset S= $\{\boldsymbol{0}, \boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_k\}$ of R^n that contains the zero vector is linearly dependent.

The set $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_k\}$ is linearly dependent iff there exists a nonzero solution of $A\boldsymbol{x}=\boldsymbol{0}$, where $A=[\boldsymbol{u}_1 \;\; \boldsymbol{u}_2 \;\; \cdots \;\; \boldsymbol{u}_k]$.

Theorem 1.8

The following statements about an $m \times n$ matrix A are equivalent:

a) The columns of A are linearly independent.
b) The equation $A\mathbf{x}=\boldsymbol{b}$ has at most one solution for each $\boldsymbol{b}$ in R^m.
c) The nullity of A is zero.
d) The rank of A is n, the number of columns of A.
e) The columns of the reduced row echelon form of A are distinct standard vectors in R^m.
f) The only solution of $A\mathbf{x}=\boldsymbol{0}$ is $\boldsymbol{0}$.
g) There is a pivot position in each column of A.

a↔f↔g

b→c→d→e→f→b

Prove f)→b) by proving $\overline{b)} \to \overline{f)}$.

$\overline{b)}$: Let $A\mathbf{x}=\boldsymbol{b}$ and $A\mathbf{y}=\boldsymbol{b}$ where $\mathbf{x}\neq \mathbf{y}$.

Then $A(\mathbf{x}-\mathbf{y})=\boldsymbol{b}-\boldsymbol{b}=\boldsymbol{0}$ but $\mathbf{x}\neq \mathbf{y}$. $\to \overline{f)}$

$$A = m\left\{ \begin{array}{c} \leftarrow \quad n \quad \rightarrow \\ \begin{bmatrix} * & * & \cdots & * \\ * & * & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ * & * & \cdots & * \end{bmatrix} \end{array} \right.$$

The equation $A\mathbf{x}=\mathbf{b}$ is called **homogeneous** if $\mathbf{b}=\mathbf{0}$.

- A homogeneous equation must be consistent.
  - ✓ $\mathbf{0}$ is a solution.
  - ✓ whether $\mathbf{0}$ is the only solution ?
- *A homogeneous system of linear equations with more variables than equations has infinitely many solutions.*
- The number of solutions of $A\mathbf{x}=0$ determines the linear dependence or independence of the columns of A

A **vector form** of the general solution of $A\mathbf{x}=0$:

- The solution of $A\mathbf{x}=0$ is expressed as a linear combination of vectors in which the coefficients are the *free* variables in the general solution.
  - ✓ *The vectors that appear in the vector form are linearly independent.*

Theorem 1.9

Vectors $\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_k$ in R^n are linearly dependent iff $\boldsymbol{u}_1=\boldsymbol{0}$ or there exists an $i\geq 2$ such that $\boldsymbol{u}_i$ is a linear combination of the preceding vectors $\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_{i-1}$.

Proof:

Suppose that $\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_k$ are linearly dependent. $\Rightarrow$

Then

- ✓ $\{\boldsymbol{u}_1 = \boldsymbol{0}\}$ or
- ✓ $\{\boldsymbol{u}_1 \neq \boldsymbol{0}$ and $\exists\, c_1, c_2, \cdots, c_k$, not all zero, such that

$c_1\boldsymbol{u}_1 + c_2\boldsymbol{u}_2 + \cdots + c_k\boldsymbol{u}_k = 0.\}$

Let i be the largest index such that $c_i \neq 0$. ……

$$\boldsymbol{u}_i = \frac{-c_1}{c_i}\boldsymbol{u}_1 - \frac{c_2}{c_i}\boldsymbol{u}_2 - \cdots - \frac{c_{i-1}}{c_i}\boldsymbol{u}_{i-1}$$

Thus $\boldsymbol{u}_i$ is a linear combination of $\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_{i-1}$.

? $\Leftarrow$?

Properties of linearly dependent and independent sets

1. A set consisting of a single nonzero vector is linearly independent, but $\{\boldsymbol{0}\}$ is linearly dependent. (by Theorem 1.9)
2. A set of two vectors $\{\boldsymbol{u}_1, \boldsymbol{u}_2\}$ is linearly dependent iff $\boldsymbol{u}_1=0$ or $\boldsymbol{u}_2$ is in the span of $\{\boldsymbol{u}_1\}$. Hence, a set of two vectors is linearly dependent iff one of the vectors is a multiple of the others. (by Theorem 1.9)
3. Let $S=\{\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_k\}$ be a linearly independent subset of R^n and $\boldsymbol{v}$ in R^n. Then $\boldsymbol{v}$ does not belong to the span of S iff $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_k, \boldsymbol{v}\}$ is linearly independent. (Proof by Theorem 1.9)
4. Every subset of R^n containing more than n vectors must be linearly dependent.
5. If S is a subset of R^n and no vector can be removed from S without changing its span, then S is linearly independent. (by Theorem 1.9)

Examples: Linearly independent or linearly dependent?

$$S_1 = \left\{ \begin{bmatrix} 3 \\ -1 \\ 7 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -2 \\ 5 \\ 1 \end{bmatrix} \right\}, S_2 = \left\{ \begin{bmatrix} -3 \\ 7 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 9 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 2 \end{bmatrix} \right\},$$

$$S_3 = \left\{ \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 3 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 4 \\ -2 \\ 3 \end{bmatrix} \right\}$$

Let A be an $m \times n$ matrix with reduced row echelon form R. Properties listed in the same row of the following table are equivalent.

The rank of A	**The number of solutions of Ax$=b$**	**The columns of A**	**The reduced row echelon form R of A**
rank $A=m$	$A\mathbf{x}=\mathbf{b}$ has at least one solution for every $\mathbf{b}$ in R^m.	The columns of A are a generating set for R^m.	Every row of R contains a pivot position.
rank $A=n$	$A\mathbf{x}=\mathbf{b}$ has at most one solution for every $\mathbf{b}$ in R^m.	The columns of A are linearly independent.	Every column of R contains a pivot position.