

1. Definition. (Linear Combination.)

Let $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$ be vectors in \mathbb{R}^m .

Let \mathbf{v} be a vector in \mathbb{R}^m .

We say \mathbf{v} is a linear combination of $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$ if the statement (\dagger) holds:

(\dagger) There exist some real numbers $\alpha_1, \alpha_2, \dots, \alpha_n$ such that $\mathbf{v} = \alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \dots + \alpha_n\mathbf{u}_n$.

The expression $\alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \dots + \alpha_n\mathbf{u}_n$ on its own is called the linear combination of the vectors $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$ with respect to the scalars $\alpha_1, \alpha_2, \dots, \alpha_n$.

2. Lemma (A). ('Dictionary' between linear combinations and matrix-vector products.)

Let A be an $(m \times n)$ -matrix, and \mathbf{t} be a vector in \mathbb{R}^n .

Suppose that for each $j = 1, 2, \dots, n$, the j -th column of A is \mathbf{a}_j and the j -th entry of \mathbf{t}

is t_j . (So $A = [\mathbf{a}_1 | \mathbf{a}_2 | \dots | \mathbf{a}_n]$ and $\mathbf{t} = \begin{bmatrix} t_1 \\ t_2 \\ \vdots \\ t_n \end{bmatrix}$.)

Then $A\mathbf{t} = t_1\mathbf{a}_1 + t_2\mathbf{a}_2 + \dots + t_n\mathbf{a}_n$.

3. Proof of Lemma (A).

For each i, j , we denote the (i, j) -th entry of A by a_{ij} .

- The i -th entry of $A\mathbf{t}$ is given by $\sum_{j=1}^n a_{ij}t_j = t_1a_{i1} + t_2a_{i2} + \cdots + t_na_{in}$.
- For each j , the i -th entry of \mathbf{a}_j (which is the j -th column of A) is a_{ij} .

Then the i -th entry of $t_1\mathbf{a}_1 + t_2\mathbf{a}_2 + \cdots + t_n\mathbf{a}_n$ is $t_1a_{i1} + t_2a_{i2} + \cdots + t_na_{in}$.

The corresponding entries of $A\mathbf{t}$, $t_1\mathbf{a}_1 + t_2\mathbf{a}_2 + \cdots + t_n\mathbf{a}_n$ agree with each other.

Hence $A\mathbf{t} = t_1\mathbf{a}_1 + t_2\mathbf{a}_2 + \cdots + t_n\mathbf{a}_n$ indeed.

Remark.

Lemma (A) looks innocent, but it will serve as a useful tool in various situations.

4. Simple concrete examples.

$$(a) \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{bmatrix} = 1 \cdot \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + 2 \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + 3 \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} + 4 \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} + 5 \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

So $\begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{bmatrix}$ is the linear combination of $\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$, $\begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$, $\begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$, $\begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$, $\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$ with respect to the scalars 1, 2, 3, 4, 5.

A manifestation of the same relation is the equality below:

$$\begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{bmatrix} = \left[\begin{array}{c|c|c|c|c} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{array} \right] \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{bmatrix}.$$

$$(b) \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{bmatrix} = 1 \cdot \begin{bmatrix} 1 \\ 3 \\ 5 \\ 7 \\ 9 \end{bmatrix} + (-1) \begin{bmatrix} 0 \\ 2 \\ 4 \\ 6 \\ 8 \end{bmatrix} + 2 \begin{bmatrix} 0 \\ 0 \\ 1 \\ 3 \\ 5 \end{bmatrix} + (-1) \begin{bmatrix} 0 \\ 0 \\ 3 \\ 6 \end{bmatrix} + 0 \cdot \begin{bmatrix} 0 \\ 0 \\ 0 \\ 2 \end{bmatrix}.$$

So $\begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{bmatrix}$ is also a linear combination of $\begin{bmatrix} 1 \\ 3 \\ 5 \\ 7 \\ 9 \end{bmatrix}$, $\begin{bmatrix} 0 \\ 2 \\ 4 \\ 6 \\ 8 \end{bmatrix}$, $\begin{bmatrix} 0 \\ 0 \\ 1 \\ 3 \\ 5 \end{bmatrix}$, $\begin{bmatrix} 0 \\ 0 \\ 3 \\ 6 \end{bmatrix}$, $\begin{bmatrix} 0 \\ 0 \\ 0 \\ 2 \end{bmatrix}$ with respect to the scalars $1, -1, 2, -1, 0$.

A manifestation of the same relation is the equality below:

$$\begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{bmatrix} = \left[\begin{array}{c|c|c|c|c} 1 & 0 & 0 & 0 & 0 \\ 3 & 2 & 0 & 0 & 0 \\ 5 & 4 & 1 & 0 & 0 \\ 7 & 6 & 3 & 3 & 0 \\ 9 & 8 & 5 & 6 & 2 \end{array} \right] \begin{bmatrix} 1 \\ -1 \\ 2 \\ -1 \\ 0 \end{bmatrix}.$$

5. Theorem (1).

Let $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$ be vectors in \mathbb{R}^m .

The statements below are true:

- (a) The zero vector $\mathbf{0}$ in \mathbb{R}^m is a linear combination of $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$.
- (b) The sum of any two linear combinations of $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$ is a linear combination of $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$.
- (c) Every scalar multiple of any linear combination of $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$ is a linear combination of $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$.

6. Proof of Theorem (1).

Let $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$ be vectors in \mathbb{R}^m .

- (a) [Ask: Can we name some appropriate real numbers $\alpha_1, \alpha_2, \dots, \alpha_n$ for which the equality $\mathbf{0} = \alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \dots + \alpha_n\mathbf{u}_n$ holds?]

We have $\mathbf{0} = 0 \cdot \mathbf{u}_1 + 0 \cdot \mathbf{u}_2 + \dots + 0 \cdot \mathbf{u}_n$.

Then by definition, $\mathbf{0}$ is a linear combination of $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$.

(b) Suppose \mathbf{v} , \mathbf{w} are linear combinations of $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$.

Then, by definition, there exist some real numbers $\beta_1, \beta_2, \dots, \beta_n$ such that

$$\mathbf{v} = \beta_1 \mathbf{u}_1 + \beta_2 \mathbf{u}_2 + \dots + \beta_n \mathbf{u}_n.$$

Also, there exist some real numbers $\gamma_1, \gamma_2, \dots, \gamma_n$ such that

$$\mathbf{w} = \gamma_1 \mathbf{u}_1 + \gamma_2 \mathbf{u}_2 + \dots + \gamma_n \mathbf{u}_n.$$

[Ask: Can we name some appropriate real numbers $\alpha_1, \alpha_2, \dots, \alpha_n$ for which the equality $\mathbf{v} + \mathbf{w} = \alpha_1 \mathbf{u}_1 + \alpha_2 \mathbf{u}_2 + \dots + \alpha_n \mathbf{u}_n$ holds?]

Note that

$$\mathbf{v} + \mathbf{w} = \dots = (\beta_1 + \gamma_1) \mathbf{u}_1 + (\beta_2 + \gamma_2) \mathbf{u}_2 + \dots + (\beta_n + \gamma_n) \mathbf{u}_n,$$

and $\beta_1 + \gamma_1, \beta_2 + \gamma_2, \dots, \beta_n + \gamma_n$ are real numbers.

Then by definition, $\mathbf{v} + \mathbf{w}$ is a linear combination of $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$.

(c) Exercise.

7. Theorem (B).

Let $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$ be vectors in \mathbb{R}^m .

Every linear combination of (finitely many) linear combinations of $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$ is a linear combination of $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$.

Remark.

In fact, Theorem (B) is saying the same thing as Statement (b) and Statement (c) in Theorem (1) combined.

8. Proof of Theorem (B).

[This argument carries the same essence of the argument for Statement (b) and Statement (c) in Theorem (1).]

Let $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$ be vectors in \mathbb{R}^m .

Pick any $\mathbf{x} \in \mathbb{R}^m$.

Suppose \mathbf{x} is a linear combination of (finitely many) vectors in \mathbb{R}^m , say, $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_p$, which are linear combinations of $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$.

[Reminder: We want to see why \mathbf{x} is a linear combination of $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$.]

By definition, \mathbf{x} is a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_p$.

Then there exist some $\alpha_1, \alpha_2, \dots, \alpha_p \in \mathbb{R}$ such that $\mathbf{x} = \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \dots + \alpha_p\mathbf{v}_p$.

[Ask: Can we link up the \mathbf{u}_j 's with the \mathbf{v}_i 's so as to see that \mathbf{x} is a linear combination of $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$?]

By assumption, for each $j = 1, 2, \dots, p$, there exist some $\beta_{1j}, \beta_{2j}, \dots, \beta_{nj} \in \mathbb{R}$ such that $\mathbf{v}_j = \beta_{1j}\mathbf{u}_1 + \beta_{2j}\mathbf{u}_2 + \dots + \beta_{nj}\mathbf{u}_n$.

Then

$$\begin{aligned}\mathbf{x} &= \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \dots + \alpha_p\mathbf{v}_p \\ &= \alpha_1(\beta_{11}\mathbf{u}_1 + \beta_{21}\mathbf{u}_2 + \dots + \beta_{n1}\mathbf{u}_n) + \alpha_2(\beta_{12}\mathbf{u}_1 + \beta_{22}\mathbf{u}_2 + \dots + \beta_{n2}\mathbf{u}_n) \\ &\quad + \dots + \alpha_p(\beta_{1p}\mathbf{u}_1 + \beta_{2p}\mathbf{u}_2 + \dots + \beta_{np}\mathbf{u}_n) \\ &= (\beta_{11}\alpha_1 + \beta_{12}\alpha_2 + \dots + \beta_{1p}\alpha_p)\mathbf{u}_1 + (\beta_{21}\alpha_1 + \beta_{22}\alpha_2 + \dots + \beta_{2p}\alpha_p)\mathbf{u}_2 \\ &\quad + \dots + (\beta_{n1}\alpha_1 + \beta_{n2}\alpha_2 + \dots + \beta_{np}\alpha_p)\mathbf{u}_n\end{aligned}$$

Note $(\beta_{k1}\alpha_1 + \beta_{k2}\alpha_2 + \dots + \beta_{kp}\alpha_p)$ is a real number for each $j = 1, 2, \dots, n$.

Then \mathbf{x} is a linear combination of $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$.

9. **Alternative argument for Theorem (B).**

By applying mathematical induction, and by consciously applying Theorem (1), we can verify the statement

‘For any positive integer s , if $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_s$ are linear combinations of $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$ and $\alpha_1, \alpha_2, \dots, \alpha_s$ are real numbers then $\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \dots + \alpha_s\mathbf{v}_s$ is a linear combination of $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$.’

Such an argument is also legitimate.

10. Lemma (2).

Let $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n, \mathbf{v}$ be vectors in \mathbb{R}^m and $\alpha_1, \alpha_2, \dots, \alpha_n$ be real numbers.

Suppose A is an $(m \times m)$ -square matrix, and \mathbf{v} is a linear combination of $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$ and the respective scalars $\alpha_1, \alpha_2, \dots, \alpha_n$.

Then $A\mathbf{v}$ is a linear combination of the vectors $A\mathbf{u}_1, A\mathbf{u}_2, \dots, A\mathbf{u}_n$ and the respective scalars $\alpha_1, \alpha_2, \dots, \alpha_n$.

11. Proof of Lemma (2).

Let $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n, \mathbf{v}$ be vectors in \mathbb{R}^m and $\alpha_1, \alpha_2, \dots, \alpha_n$ be real numbers.

Suppose A is an $(m \times m)$ -square matrix, and \mathbf{v} is a linear combination of $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$ and the respective scalars $\alpha_1, \alpha_2, \dots, \alpha_n$.

By assumption, $\mathbf{v} = \alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \dots + \alpha_n\mathbf{u}_n$.

Then $A\mathbf{v} = A(\alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \dots + \alpha_n\mathbf{u}_n) = \alpha_1A\mathbf{u}_1 + \alpha_2A\mathbf{u}_2 + \dots + \alpha_nA\mathbf{u}_n$.

Hence $A\mathbf{v}$ is a linear combination of $A\mathbf{u}_1, A\mathbf{u}_2, \dots, A\mathbf{u}_n$ and the respective scalars $\alpha_1, \alpha_2, \dots, \alpha_n$.

12. Lemma (3). (A ‘partial converse’ of Lemma (2).)

Let $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n, \mathbf{v}$ be vectors in \mathbb{R}^m and $\alpha_1, \alpha_2, \dots, \alpha_n$ be real numbers.

Suppose A is a non-singular $(m \times m)$ -square matrix, and $A\mathbf{v}$ is a linear combination of the vectors $A\mathbf{u}_1, A\mathbf{u}_2, \dots, A\mathbf{u}_n$ and the respective scalars $\alpha_1, \alpha_2, \dots, \alpha_n$.

Then \mathbf{v} is a linear combination of the vectors $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$ and the respective scalars $\alpha_1, \alpha_2, \dots, \alpha_n$.

13. Proof of Lemma (3).

Let $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n, \mathbf{v}$ be vectors in \mathbb{R}^m and $\alpha_1, \alpha_2, \dots, \alpha_n$ be real numbers.

Suppose A is a non-singular $(m \times m)$ -square matrix, and $A\mathbf{v}$ is a linear combination of the vectors $A\mathbf{u}_1, A\mathbf{u}_2, \dots, A\mathbf{u}_n$ and the respective scalars $\alpha_1, \alpha_2, \dots, \alpha_n$.

By assumption, $A\mathbf{v} = \alpha_1 A\mathbf{u}_1 + \alpha_2 A\mathbf{u}_2 + \dots + \alpha_n A\mathbf{u}_n$.

Since A is non-singular, A is invertible. Therefore

$$\begin{aligned}\mathbf{v} &= I_m \mathbf{v} = (A^{-1}A)\mathbf{v} = A^{-1}(A\mathbf{v}) \\ &= A^{-1}(\alpha_1 A\mathbf{u}_1 + \alpha_2 A\mathbf{u}_2 + \dots + \alpha_n A\mathbf{u}_n) \\ &= \alpha_1 A^{-1}(A\mathbf{u}_1) + \alpha_2 A^{-1}(A\mathbf{u}_2) + \dots + \alpha_n A^{-1}(A\mathbf{u}_n) \\ &= \alpha_1 (A^{-1}A)\mathbf{u}_1 + \alpha_2 (A^{-1}A)\mathbf{u}_2 + \dots + \alpha_n (A^{-1}A)\mathbf{u}_n \\ &= \alpha_1 I_m \mathbf{u}_1 + \alpha_2 I_m \mathbf{u}_2 + \dots + \alpha_n I_m \mathbf{u}_n \\ &= \alpha_1 \mathbf{u}_1 + \alpha_2 \mathbf{u}_2 + \dots + \alpha_n \mathbf{u}_n\end{aligned}$$

Then \mathbf{v} is a linear combination of the vectors $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$ and the respective scalars $\alpha_1, \alpha_2, \dots, \alpha_n$.

14. We combine Lemma (2) and Lemma (3) to obtain Theorem (C) below:

Theorem (C).

Let $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n, \mathbf{v}$ be vectors in \mathbb{R}^m and $\alpha_1, \alpha_2, \dots, \alpha_n$ be real numbers.

Suppose A is a non-singular $(m \times m)$ -square matrix. Then the statements below are logically equivalent:

- (a) *\mathbf{v} is a linear combination of the vectors $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$ and the respective scalars $\alpha_1, \alpha_2, \dots, \alpha_n$.*
- (b) *$A\mathbf{v}$ is a linear combination of the vectors $A\mathbf{u}_1, A\mathbf{u}_2, \dots, A\mathbf{u}_n$ and the respective scalars $\alpha_1, \alpha_2, \dots, \alpha_n$.*

Remark.

In plain words, this result is saying that

linear relations amongst vectors (though not necessarily the individual vectors themselves) are preserved upon multiplication of the same non-singular square matrix from the left to the vectors.

When we think in terms of row operations, this result is saying that

linear relations amongst vectors (though not necessarily the individual vectors themselves) are preserved upon the application of the same sequence of row operations to the vectors.

15. **Theorem (4).** (Generalization of Lemma (2) and Lemma (3).)

Let $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n, \mathbf{v}$ be vectors in \mathbb{R}^m . Let A be a $(p \times m)$ -matrix.

(a) Suppose \mathbf{v} is a linear combination of $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$.

Then $A\mathbf{v}$ is a linear combination of the vectors $A\mathbf{u}_1, A\mathbf{u}_2, \dots, A\mathbf{u}_n$.

(b) Suppose $\mathcal{N}(A) = \{\mathbf{0}\}$, and $A\mathbf{v}$ is a linear combination of the vectors $A\mathbf{u}_1, A\mathbf{u}_2, \dots, A\mathbf{u}_n$.

Then \mathbf{v} is a linear combination of the vectors $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$.

Proof of Theorem (4). Exercise.