MATH1030 Theoretical results on linear dependence and linear independence.

1. Recall the definition for the respective notions of *linear dependence* and *linear independence*.

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

   We say that $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n$ are linearly dependent if and only if the statement (LD) holds:

   (LD) There exist some $\alpha_1, \alpha_2, \cdots, \alpha_n \in \mathbb{R}$ such that $\alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \cdots + \alpha_n\mathbf{u}_n = \mathbf{0}$ and $\alpha_1, \alpha_2, \cdots, \alpha_n$ are not all zero.

   The equality $\alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \cdots + \alpha_n\mathbf{u}_n = \mathbf{0}$ in which $\alpha_1, \alpha_2, \cdots, \alpha_n$ are not all zero is called a non-trivial linear relation of $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n$.

2. The notions of linear dependence and linear combinations are linked up in the result below:

   Theorem (H).

   Let $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n$ be vectors in $\mathbb{R}^m$. The statements below are logically equivalent:

   (a) *$\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n$ are linearly dependent.*

   (b) *At least one vector amongst $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n$ is a linear combination of the others.*

3. **Proof of Theorem (H).**

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

   - Suppose $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n$ are linearly dependent.
     Then there exist some $\alpha_1, \alpha_2, \cdots, \alpha_n \in \mathbb{R}$ such that $\alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \cdots + \alpha_n\mathbf{u}_n = \mathbf{0}$ and $\alpha_1, \alpha_2, \cdots, \alpha_n$ are not all zero.
     Without loss of generality, assume $\alpha_1 \neq 0$.
     Then $\mathbf{u}_1 = (-\frac{\alpha_2}{\alpha_1})\mathbf{u}_2 + (-\frac{\alpha_3}{\alpha_1})\mathbf{u}_3 + \cdots + (-\frac{\alpha_n}{\alpha_1})\mathbf{u}_n$.
     Hence $\mathbf{u}_1$ is a linear combination of $\mathbf{u}_2, \mathbf{u}_3, \cdots, \mathbf{u}_n$.
   - Suppose at least one vector amongst $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n$ is a linear combination of the others.
     Without loss of generality, assume $\mathbf{u}_1$ is a linear combination of $\mathbf{u}_2, \mathbf{u}_3, \cdots, \mathbf{u}_n$.
     Then there exist some $\beta_2, \beta_3, \cdots, \beta_n \in \mathbb{R}$ such that $\mathbf{u}_1 = \beta_2\mathbf{u}_2 + \beta_3\mathbf{u}_3 + \cdots + \beta_n\mathbf{u}_n$.
     Therefore $1 \cdot \mathbf{u}_1 + (-\beta_2)\mathbf{u}_2 + (-\beta_3)\mathbf{u}_3 + \cdots + (-\beta_n)\mathbf{u}_n = \mathbf{0}$. This is a non-trivial linear relation of $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n$.
     Hence $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n$ are linearly dependent.

4. **Illustration of the idea in Theorem (H).**

   Write $\mathbf{u}_1 = \begin{bmatrix}1\\2\\3\\4\\5\end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix}1\\3\\5\\7\\9\end{bmatrix}$, $\mathbf{u}_3 = \begin{bmatrix}0\\1\\2\\3\\4\end{bmatrix}$, $\mathbf{u}_4 = \begin{bmatrix}0\\0\\1\\3\\5\end{bmatrix}$, $\mathbf{u}_5 = \begin{bmatrix}0\\0\\0\\1\\2\end{bmatrix}$.

   We have

   $$1 \cdot \mathbf{u}_1 + (-1)\mathbf{u}_2 + 2\mathbf{u}_3 + (-2)\mathbf{u}_4 + 3\mathbf{u}_5 = \mathbf{0}_5.$$

   We also have

   $$\begin{aligned}
   \mathbf{u}_1 &= 1 \cdot \mathbf{u}_2 - 2\mathbf{u}_3 + 2\mathbf{u}_4 - 3\mathbf{u}_5 \\
   \mathbf{u}_2 &= 1 \cdot \mathbf{u}_1 + 2\mathbf{u}_3 + (-2)\mathbf{u}_4 + 3\mathbf{u}_5 \\
   \mathbf{u}_3 &= -\frac{1}{2}\mathbf{u}_1 + \frac{1}{2}\mathbf{u}_2 + 1 \cdot \mathbf{u}_4 - \frac{3}{2}\mathbf{u}_5 \\
   \mathbf{u}_4 &= \frac{1}{2}\mathbf{u}_1 - \frac{1}{2}\mathbf{u}_2 + 1 \cdot \mathbf{u}_3 + \frac{3}{2}\mathbf{u}_5 \\
   \mathbf{u}_5 &= -\frac{1}{3}\mathbf{u}_1 + \frac{1}{3}\mathbf{u}_2 - \frac{2}{3}\mathbf{u}_3 + \frac{2}{3}\mathbf{u}_4
   \end{aligned}$$

   Theorem (H) says that each of these six equalities is valid exactly because of the validity of each other.

5. By logic, Corollary (1) to Theorem (H) holds immediately as a re-formulation of Theorem (H).

   The notion of linear independence can be understood through this re-formulation of Theorem (H): it corresponds to our heuristic understanding of the word *independence* in daily language.

   Corollary (1) to Theorem (H).

   Let $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n$ be vectors in $\mathbb{R}^m$. The statements below are logically equivalent:

(a) $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n$ *are linearly independent.*

(b) *None of the vector amongst* $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n$ *is a linear combination of the others.*

6. Corollary (2) to Theorem (H) is an immediate consequence of Theorem (H).

   Corollary (2) to Theorem (H).

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

   Suppose $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n$ *are linearly independent.*

   Then the statements below are logically equivalent:

   (a) $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n, \mathbf{v}$ *are linearly dependent.*

   (b) $\mathbf{v}$ *is a linear combination of* $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n$.

7. **Proof of Corollary (2) to Theorem (H).**

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

   Suppose $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n$ are linearly independent.

   - Suppose $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n, \mathbf{v}$ are linearly dependent.
     Then there exist some $\alpha_1, \alpha_2, \cdots, \alpha_n, \beta \in \mathbb{R}$ such that $\alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \cdots + \alpha_n\mathbf{u}_n + \beta\mathbf{v} = \mathbf{0}$ and $\alpha_1, \alpha_2, \cdots, \alpha_n, \beta$ are not all zero.
     We verify that $\beta \neq 0$:
     * Suppose it were true that $\beta = 0$. Then we would have $\alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \cdots + \alpha_n\mathbf{u}_n = \mathbf{0}$ and $\alpha_1, \alpha_2, \cdots, \alpha_n$ are not all zero. Therefore $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n$ would be linearly dependent. Contradiction arises.

     As $\beta \neq 0$, we have $\mathbf{v} = (-\frac{\alpha_1}{\beta})\mathbf{u}_1 + (-\frac{\alpha_2}{\beta})\mathbf{u}_2 + \cdots + (-\frac{\alpha_n}{\beta})\mathbf{u}_n + \beta\mathbf{v}$.

     Therefore $\mathbf{v}$ is a linear combination of $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n$.
   - Suppose $\mathbf{v}$ is a linear combination of $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n$.
     Then there exist some $\gamma_1, \gamma_2, \cdots, \gamma_n \in \mathbb{R}$ such that $\mathbf{v} = \gamma_1\mathbf{u}_1 + \gamma_2\mathbf{u}_2 + \cdots + \gamma_n\mathbf{u}_n$.
     Therefore $\gamma_1\mathbf{u}_1 + \gamma_2\mathbf{u}_2 + \cdots + \gamma_n\mathbf{u}_n - 1 \cdot \mathbf{v} = \mathbf{0}$.
     Hence $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n, \mathbf{v}$ are linearly dependent.

8. By logic, Corollary (3) to Theorem (H) holds immediately as a re-formulation of Corollary (2) to Theorem (H).

   Corollary (3) to Theorem (H).

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

   Suppose $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n$ *are linearly independent.*

   Then the statements below are logically equivalent:

   (a) $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n, \mathbf{v}$ *are linearly independent.*

   (b) $\mathbf{v}$ *is not a linear combination of* $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n$.

9. We are going to give a useful re-formulations for the notion of *linear independence* which is useful in the study of theoretical questions.

   We start by recalling the re-formulations for the notion of *linearly independence* below:

   Let $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n$ *be vectors in* $\mathbb{R}^m$. *Define the* $(m \times n)$*-matrix* U *by* $U = [\ \mathbf{u}_1 \mid \mathbf{u}_2 \mid \cdots \mid \mathbf{u}_n\ ]$.
   The statements below are logically equivalent:

   ($\sim\diamondsuit$) $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n$ *are linearly independent.*

   ($\sim\clubsuit$) *The only solution of homogeneous system of linear equations* $\mathcal{LS}(U, \mathbf{0}_m)$ *is the trivial solution.*

   ($\sim\heartsuit$) $\mathcal{N}(U) = \{\mathbf{0}_n\}$.

   ($\sim\spadesuit$) *For any* $\mathbf{t} \in \mathbb{R}^n$, *if* $\mathbf{t} \in \mathcal{N}(U)$ *then* $\mathbf{t} = \mathbf{0}_n$.

   With the help of the 'dictionary' between linear combinations and matrix-vector products applied on the statement ($\sim\spadesuit$) above, we obtain another re-formulation, in the form of Lemma (I), for the notion of linear independence, which is useful for theoretical discussions (or in calculations in which the vectors involved are not given in 'concrete' terms).

10. **Lemma (I).**

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

    The statements below are logically equivalent to each other:

(LI_0) $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n$ *are linearly independent.*

(LI) *For any* $\beta_1, \beta_2, \cdots, \beta_n \in \mathbb{R}$, *if* $\beta_1\mathbf{u}_1 + \beta_2\mathbf{u}_2 + \cdots + \beta_n\mathbf{u}_n = \mathbf{0}_m$ *then* $\beta_1 = \beta_2 = \cdots = \beta_n = 0$.

Remark. In fact, in many standard textbooks, people simply refers to the statement (LI) as the definition for '$\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n$ are linearly independent vectors in $\mathbb{R}^m$'.

11. **Theorem (1).**

Let $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_k$ *be vectors in* $\mathbb{R}^m$.

Suppose $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_k$ *are linearly independent.*

Then for each $j = 1, 2, \cdots, k$, *the vectors* $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_j$ *are linearly independent.*

Remark. In plain words, Theorem (1) says:

Any 'portion' of a collection of linearly independent vectors is itself a collection of linearly independent vectors.

12. **Proof of Theorem (1).**

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

Suppose $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_k$ are linearly independent.

Suppose $j = 1, 2, \cdots, k$. Pick any $\alpha_1, \alpha_2, \cdots, \alpha_j \in \mathbb{R}$.

Suppose $\alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \cdots + \alpha_j\mathbf{u}_j = \mathbf{0}$.

Then $\alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \cdots + \alpha_j\mathbf{u}_j + 0 \cdot \mathbf{u}_{j+1} + 0 \cdot \mathbf{u}_{j+2} + \cdots + 0 \cdot \mathbf{u}_k = \mathbf{0}$.

By assumption $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_k$ are linearly independent. Then $\alpha_1 = \alpha_2 = \cdots = \alpha_j = 0$.

It follows that $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_j$ are linearly independent.

13. Theorem (1) is logically equivalent to the statement below, which is called the 'contra-positive re-formulation' of Theorem (1).

Theorem (2).

Let $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n, \mathbf{w}_1, \mathbf{w}_2, \cdots, \mathbf{w}_p$ *be vectors in* $\mathbb{R}^m$.

Suppose $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n$ *are linearly dependent.*

Then $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n, \mathbf{w}_1, \mathbf{w}_2, \cdots, \mathbf{w}_p$ *are linearly dependent.*

Remark. In plain words, Theorem (2) says:

A collection of vectors is definitely linearly dependent when soem portion of it is, on its own, a collection of linearly dependent vectors.

14. **Lemma (3).**

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

Suppose A *is an* $(m \times m)$*-square matrix, and* $A\mathbf{u}_1, A\mathbf{u}_2, \cdots, A\mathbf{u}_n$ *are linearly independent.*

Then $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n$ *are linearly independent.*

15. **Proof of Lemma (3).**

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

Suppose A is an $(m \times m)$-square matrix, and $A\mathbf{u}_1, A\mathbf{u}_2, \cdots, A\mathbf{u}_n$ are linearly independent..

[We want to verify that $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n$ are linearly independent. This amounts to verifying the statement 'for any $\beta_1, \beta_2, \cdots, \beta_n \in \mathbb{R}$, if $\beta_1\mathbf{u}_1 + \beta_2\mathbf{u}_2 + \cdots + \beta_n\mathbf{u}_n = \mathbf{0}_m$ then $\beta_1 = \beta_2 = \cdots = \beta_n = 0$']

Pick any $\beta_1, \beta_2, \cdots, \beta_n \in \mathbb{R}$. Suppose $\beta_1\mathbf{u}_1 + \beta_2\mathbf{u}_2 + \cdots + \beta_n\mathbf{u}_n = \mathbf{0}_m$.

Then

$$\begin{aligned} \beta_1 A\mathbf{u}_1 + \beta A\mathbf{u}_2 + \cdots + \beta_n A\mathbf{u}_n &= A(\beta_1\mathbf{u}_1 + \beta_2\mathbf{u}_2 + \cdots + \beta_n\mathbf{u}_n) \\ &= A\mathbf{0}_m = \mathbf{0}_m \end{aligned}$$

Then, by assumption, $\beta_1 = \beta_2 = \cdots = \beta_n = 0$.

Therefore $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n$ linearly independent.

16. **Lemma (4). (A 'partial converse' of Lemma (3).)**

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

Suppose A *is a non-singular* $(m \times m)$*-square matrix, and* $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n$ *are linearly independent.*

Then $A\mathbf{u}_1, A\mathbf{u}_2, \cdots, A\mathbf{u}_n$ *are linearly independent.*

17. **Proof of Lemma (4).**

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

Suppose A is a non-singular $(m \times m)$-square matrix, and $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n$ are linearly independent.

[We want to verify that $A\mathbf{u}_1, A\mathbf{u}_2, \cdots, A\mathbf{u}_n$ are linearly independent. This amounts to verifying the statement 'for any $\beta_1, \beta_2, \cdots, \beta_n \in \mathbb{R}$, if $\beta_1 A\mathbf{u}_1 + \beta_2 A\mathbf{u}_2 + \cdots + \beta_n A\mathbf{u}_n = \mathbf{0}_m$ then $\beta_1 = \beta_2 = \cdots = \beta_n = 0$']

Pick any $\beta_1, \beta_2, \cdots, \beta_n \in \mathbb{R}$. Suppose $\beta_1 A\mathbf{u}_1 + \beta_2 A\mathbf{u}_2 + \cdots + \beta_n A\mathbf{u}_n = \mathbf{0}_m$.

Then

$$\begin{aligned} \beta_1\mathbf{u}_1 + \beta\mathbf{u}_2 + \cdots + \beta_n\mathbf{u}_n &= \beta_1 I_m\mathbf{u}_1 + \beta I_m\mathbf{u}_2 + \cdots + \beta_n I_m\mathbf{u}_n \\ &= \beta_1(A^{-1}A)\mathbf{u}_1 + \beta(A^{-1}A)\mathbf{u}_2 + \cdots + \beta_n(A^{-1}A)\mathbf{u}_n \\ &= \beta_1 A^{-1}(A\mathbf{u}_1) + \beta A^{-1}(A\mathbf{u}_2) + \cdots + \beta_n A^{-1}(A\mathbf{u}_n) \\ &= A^{-1}(\beta_1 A\mathbf{u}_1 + \beta_2 A\mathbf{u}_2 + \cdots + \beta_n A\mathbf{u}_n) \\ &= A^{-1}\mathbf{0}_m = \mathbf{0}_m \end{aligned}$$

By assumption, $\beta_1 = \beta_2 = \cdots = \beta_n = 0$.

Therefore $A\mathbf{u}_1, A\mathbf{u}_2, \cdots, A\mathbf{u}_n$ linearly independent.

18. We combine Lemma (3) and Lemma (4) to obtain Theorem (J) below:

Theorem (J).

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

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

(a) *$\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n$ are linearly independent.*

(b) *$A\mathbf{u}_1, A\mathbf{u}_2, \cdots, A\mathbf{u}_n$ are linearly independent.*

Remark. In plain words, this result is saying that

linear independence for a collection of 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 independence for a collection of vectors (though not necessarily the individual vectors themselves) are preserved upon the application of the same sequence of row operations to the vectors.

19. **Theorem (5). (Generalization of Lemma (3) and Lemma (4).)**

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

(a) *Suppose $A\mathbf{u}_1, A\mathbf{u}_2, \cdots, A\mathbf{u}_n$.*
Then $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n$ are linearly independent.

(b) *Suppose $\mathcal{N}(A) = \{\mathbf{0}\}$, and $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n$ are linearly independent.*
Then $A\mathbf{u}_1, A\mathbf{u}_2, \cdots, A\mathbf{u}_n$ are linearly independent.

Proof of Theorem (5). Exercise.