4.4 Basis for subspaces of column vectors.

0. *Assumed background.*

   - Whatever has been covered in Topics 1-3, especially:—
     - 1.5 *Linear combinations.*
     - 1.6 *Linear dependence and linear independence.*
     - 3.3 *Various necessary and sufficient conditions for invertibility.*
   - 4.1 *Sets of matrices and sets of vectors.*
   - 4.2 *Set equality (for sets of matrices and sets of vectors).*
   - 4.3 *Subspaces of column vectors.*

   Abstract. We introduce:—

   - the notions of basis for an arbitrary subspace of $\mathbb{R}^n$,
   - a re-formulation for the notion of invertibility of square matrices in terms of basis,
   - the 'Change-of-basis' Theorem for arbitrary subspaces of $\mathbb{R}^n$.

   In the *appendix*, we provide the proof of the 'Change-of-basis' Theorem.

1. **Definition. (Basis for a subspace of $\mathbb{R}^n$ over the reals.)**

   Let $\mathcal{V}$ be a subspace of $\mathbb{R}^n$ over the reals.

   (1) *(Suppose $\mathcal{V}$ is the zero subspace of $\mathbb{R}^n$ over the reals.)*
   We declare that **the empty set is the basis for** *(or* **of***)* $\{\mathbf{0}_n\}$ **over the reals**.

   (2) *Suppose $\mathcal{V}$ is not the zero subspace of $\mathbb{R}^n$ over the reals.*
   Suppose $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_p \in \mathcal{V}$.
   Then we say that $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_p$ **constitute a basis for** *(or* **of***)* $\mathcal{V}$ **over the reals** *if and only if both of the statements* (BL), (BS) *below hold:—*

   (BL) $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_p$ *are linearly independent over the reals.*

   (BS) *Every column vector belonging to $\mathcal{V}$ is a linear combination of $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_p$ over the reals.*

2. **Comments on the use of set language in the definition above.**

   In the context of this definition, it happens that the statement (BS) is equivalent to the statement (BS') that reads:—

   (BS') $\mathcal{V} = \mathsf{Span}\ (\{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_p\})$.

   For this reason, some people may choose to replace the condition (BS) by the condition (BS') in the entire statement of the definition for the notion of basis.
   This is an approach usually taken in a more advanced course in *linear algebra*, in which set language is used more extensively.

   The purpose and the advantage in such an approach are to allow linear algebra to be done on more general mathematical objects (not just matrices and vectors), in which a key concept to be introduced very soon, namely *dimension*, needs not to be restricted to be 'finite numbers'.

   (There is also the very small purpose of taking care of the zero subspace in a more natural way.)

   In this course, we choose not to use set language unless we must. (And there are not many situations we will have to take care of the zero subspace.)

3. The conditions (BL), (BS) in the statement of the definition for the notion of basis can be condensed into one condition, of the form of an existence-and-uniqueness statement. This is made precise in the result below.

   Theorem (1). (Re-formulation for the notion of basis.)

   Let $\mathcal{V}$ be a subspace of $\mathbb{R}^n$ over the reals.

   Suppose $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_p \in \mathcal{V}$.

   Then the statements below are logically equivalent:—

   (1) $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_p$ *constitute a basis for $\mathcal{V}$ over the reals.*

   (2) *For any $\mathbf{x} \in \mathcal{V}$, there exist some unique $\alpha_1, \alpha_2, \cdots, \alpha_p \in \mathbb{R}$ such that $\mathbf{x} = \alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \cdots + \alpha_p\mathbf{u}_p$.*

Proof of Theorem (1). Exercise. (This is a tedious word game in definitions and logic.)

4. **Theorem (2). (Standard base for $\mathbb{R}^n$.)**

Fix any positive integer n. Suppose $\mathbf{e}_k^{(n)} = E_{k,1}^{n,1}$ for each $k = 1, 2, \cdots, n$.

(So $\mathbf{e}_k^{(n)}$ is the column vector with n real entries whose k-th entry is 1 and whose every other entry is 0. .)

Then the n column vectors $\mathbf{e}_1^{(n)}, \mathbf{e}_2^{(n)}, \cdots, \mathbf{e}_n^{(n)}$ constitute a basis for $\mathbb{R}^n$.

Remark on terminology. $\mathbf{e}_1^{(n)}, \mathbf{e}_2^{(n)}, \cdots, \mathbf{e}_n^{(n)}$ are collectively called the **standard base for $\mathbb{R}^n$ (over the reals)**.

5. **Proof of Theorem (2).**

Fix any positive integer n. Suppose $\mathbf{e}_k^{(n)} = E_{k,1}^{n,1}$ for each $k = 1, 2, \cdots, n$.

(a) We verify that $\mathbf{e}_1^{(n)}, \mathbf{e}_2^{(n)}, \cdots, \mathbf{e}_n^{(n)}$ are linearly independent over the reals:—

- Pick any $\alpha_1, \alpha_2, \cdots, \alpha_n \in \mathbb{R}$.
  Suppose $\alpha_1 \mathbf{e}_1^{(n)} + \alpha_2 \mathbf{e}_2^{(n)} + \cdots + \alpha_n \mathbf{e}_n^{(n)} = \mathbf{0}_n$.
  We have $\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix} = \alpha_1 \mathbf{e}_1^{(n)} + \alpha_2 \mathbf{e}_2^{(n)} + \cdots + \alpha_n \mathbf{e}_n^{(n)} = \mathbf{0}_n$. Then $\alpha_1 = \alpha_2 = \cdots = \alpha_n = 0$.

(b) We verify that every column vector belonging to $\mathbb{R}^n$ is a linear combination of $\mathbf{e}_1^{(n)}, \mathbf{e}_2^{(n)}, \cdots, \mathbf{e}_n^{(n)}$ over the reals:

- Pick any $\mathbf{x} \in \mathbb{R}^n$.
  For each $j = 1, 2, \cdots, n$, denote the j-th entry of $\mathbf{x}$ by x_j.
  Then $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x_1 \mathbf{e}_1^{(n)} + x_2 \mathbf{e}_2^{(n)} + \cdots + x_n \mathbf{e}_n^{(n)}$.

6. **Bases of $\mathbb{R}^n$ over the reals.**

For the moment we focus on $\mathbb{R}^n$ as a subspace of $\mathbb{R}^n$ itself, and think of the bases of $\mathbb{R}^n$ over the reals. (More general subspaces of $\mathbb{R}^n$ are considered later.)

We make some observations with the various re-formulations of the notion of invertibility.

Given any n column vectors, say, $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n$ with n real entries, the statements (I), (II), (III) below are logically equivalent:—

(I) *The square matrix $[\ \mathbf{u}_1 \mid \mathbf{u}_2 \mid \cdots \mid \mathbf{u}_n\ ]$ is invertible.*

(II) *$\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n$ are linearly independent over the reals.*

(III) *Every column vector with n real entries is a linear combination of $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n$ are linearly independent over the reals.*

Because of the logical equivalence between the statement (II) and the statement (III), when either of them holds, the other will also hold. Then, as a consequence the statement (IV) will hold as well:—

(IV) *$\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n$ constitutes a basis for $\mathbb{R}^n$ over the reals.*

When the statement (IV) holds, the statements (II), (III) will certainly hold by the definition of the notion of basis.

7. The above consideration suggests the result below, which may be regarded as an 'extension' of the 'dictionary' on the various re-formulations for the notion of invertibility.

Theorem (3). (Re-formulations for the notion of invertibility, incorporating the point of view of basis.)

Suppose A is an $(n \times n)$-square matrix with real entries. Then:—

(a) The statements (1)-(6) are logically equivalent:—

(1) *A is invertible.*

(2) *The homogeneous system $\mathcal{LS}(A, \mathbf{0}_n)$ has no non-trivial solution.*

(3) *For any $\mathbf{b} \in \mathbb{R}^n$, the system $\mathcal{LS}(A, \mathbf{b})$ is consistent.*

(4) *The columns of A are linearly independent over the reals.*

(5) *Every column vector with n real entries is a linear combination of the columns of A over the real.*

(6) *The columns of A constitute a basis for $\mathbb{R}^n$ over the reals.*

(b) *The statements (1*)-(6*) are logically equivalent to each other, and also are also logically equivalent to each of the statements (1)-(6):—*

(1*) *A^t is invertible.*

(2*) *The homogeneous system $\mathcal{LS}(A^t, \mathbf{0}_n)$ has no non-trivial solution.*

(3*) *For any $\mathbf{b} \in \mathbb{R}^n$, the system $\mathcal{LS}(A^t, \mathbf{b})$ is consistent.*

(4*) *The rows of A are linearly independent over the reals.*

(5*) *Every row vector with n real entries is a linear combination of the rows of A over the reals.*

(6*) *The columns of A^t constitute a basis for $\mathbb{R}^n$ over the reals.*
(Or equivalently: *the respective transposes of the rows of A constitute a basis for $\mathbb{R}^n$ over the reals.*)

(c) *Now further suppose any one of the statements (1)-(6), (1*)-(6*) holds. (So all hold.)*
For each $j = 1, 2, \cdots, n$, denote the j-th column of A by $\mathbf{a}_j$.
Suppose $\mathbf{x} \in \mathbb{R}^n$. For each $k = 1, 2, \cdots, n$, denote the k-th entry of $A^{-1}\mathbf{x}$ by $\alpha_k(\mathbf{x})$.
Then $\mathbf{x} = \alpha_1(\mathbf{x})\ \mathbf{a}_1 + \alpha_2(\mathbf{x})\ \mathbf{a}_2 + \cdots + \alpha_n(\mathbf{x})\ \mathbf{a}_n$.

Remark. The argument for part (a) of the conclusion is outlined in the passages above the statement of Theorem (3). That for part (b) follows immediately from the logical equivalence between the statements (1), (1*). We only have to verify part (c) of the conclusion.

8. **Argument for part (c) of the conclusion in Theorem (3).**

Suppose A is an $(n \times n)$-square matrix with real entries.

Suppose any one of the statements (1)-(6), (1*)-(6*) holds. (So all hold.)

For each $j = 1, 2, \cdots, n$, denote the j-th column of A by $\mathbf{a}_j$.

Suppose $\mathbf{x} \in \mathbb{R}^n$. For each $k = 1, 2, \cdots, n$, denote the k-th entry of $A^{-1}\mathbf{x}$ by $\alpha_k(\mathbf{x})$.

Then

$$\mathbf{x} = I_n\mathbf{x} = (AA^{-1})\mathbf{x} = A(A^{-1}\mathbf{x}) = [\ \mathbf{a}_1 \mid \mathbf{a}_2 \mid \cdots \mid \mathbf{a}_n\ ] \begin{bmatrix} \alpha_1(\mathbf{x}) \\ \alpha_2(\mathbf{x}) \\ \vdots \\ \alpha_n(\mathbf{x}) \end{bmatrix} = \alpha_1(\mathbf{x})\ \mathbf{a}_1 + \alpha_2(\mathbf{x})\ \mathbf{a}_2 + \cdots + \alpha_n(\mathbf{x})\ \mathbf{a}_n.$$

9. **Example (1). (Display of various bases of $\mathbb{R}^n$, and illustration of the content of Theorem (3).)**

(For simplicity of presentation, we work with $\mathbb{R}^4$ here. But the idea can be generalized to $\mathbb{R}^n$ for each n.)

Apart from the standard base for $\mathbb{R}^4$, we display three other bases for $\mathbb{R}^4$ (which are not especially important), and justify our claims with slightly different methods, as suggested by the various re-formulations for the notion for invertibility.

(a) *Define* $\mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}$, $\mathbf{u}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}$, $\mathbf{u}_4 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$.

i. We verify that $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4$ are linearly independent:—

- Pick any $\alpha_1, \alpha_2, \alpha_3, \alpha_4 \in \mathbb{R}$. Suppose $\alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \alpha_3\mathbf{u}_3 + \alpha_4\mathbf{u}_4 = \mathbf{0}_4$.
  Then $\begin{bmatrix} \alpha_1 \\ \alpha_1 + \alpha_2 \\ \alpha_1 + \alpha_2 + \alpha_3 \\ \alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 \end{bmatrix} = \alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \alpha_3\mathbf{u}_3 + \alpha_4\mathbf{u}_4 = \mathbf{0}_4$.
  By comparing the respective entries, we deduce $\alpha_1 = 0$, $\alpha_2 = 0$, $\alpha_3 = 0$, $\alpha_4 = 0$ in succession.

ii. We verify that every column vector belonging to $\mathbb{R}^4$ is a linear combination of $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4$:

- Pick any $\mathbf{x} \in \mathbb{R}^4$. For each $j = 1, 2, 3, 4$, denote the j-th entry of $\mathbf{x}$ by x_j.
  We have

$$\begin{aligned} \mathbf{x} &= x_1\mathbf{e}_1^{(4)} + x_2\mathbf{e}_2^{(4)} + x_3\mathbf{e}_3^{(4)} + x_4\mathbf{e}_4^{(4)} \\ &= x_1(\mathbf{u}_1 - \mathbf{u}_2) + x_2(\mathbf{u}_2 - \mathbf{u}_3) + x_3(\mathbf{u}_3 - \mathbf{u}_4) + x_4\mathbf{u}_4 \\ &= x_1\mathbf{u}_1 + (x_2 - x_1)\mathbf{u}_2 + (x_3 - x_2)\mathbf{u}_3 + (x_4 - x_3)\mathbf{u}_4 \end{aligned}$$

iii. It follows that $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4$ constitute a basis for $\mathbb{R}^4$ over the reals.

From the calculations above, we see that for any $\mathbf{x} \in \mathbb{R}^4$, if $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$ then

$$\mathbf{x} = x_1\mathbf{u}_1 + (x_2 - x_1)\mathbf{u}_2 + (x_3 - x_2)\mathbf{u}_3 + (x_4 - x_3)\mathbf{u}_4.$$

This expression of such an $\mathbf{x}$ as a linear combination of $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4$ is uniquely determined.

(b) *Define* $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 2 \\ 3 \end{bmatrix}$, $\mathbf{v}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 2 \end{bmatrix}$, $\mathbf{v}_4 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$.

Write $V = [\ \mathbf{v}_1 \mid \mathbf{v}_2 \mid \mathbf{v}_3 \mid \mathbf{v}_4\ ]$.

i. We verify that $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4$ are linearly independent:—

- We obtain the sequence of row operations below joining V to some row-echelon form $V^\sharp$:—

$$V = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 3 & 2 & 1 & 0 \\ 4 & 3 & 2 & 1 \end{bmatrix} \xrightarrow{-2R_1+R_2} \xrightarrow{-3R_1+R_3} \xrightarrow{-4R_1+R_4}$$

$$\xrightarrow{-2R_2+R_3} \xrightarrow{-3R_2+R_4} \xrightarrow{-2R_3+R_4} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = V^\sharp.$$

$V^\sharp$ is in fact the reduced row-echelon form I_4, whose columns are all pivot columns.
It follows that $\mathcal{LS}(V,\ \mathbf{0}_4)$ has no non-trivial solution.

ii. We verify that every column vector belonging to $\mathbb{R}^4$ is a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4$:

- Pick any $\mathbf{x} \in \mathbb{R}^4$. [We show that the system $\mathcal{LS}(V,\ \mathbf{x})$ is consistent.]
Write $C_\mathbf{x} = [\ V \mid \mathbf{x}\ ]$.
We obtain the sequence of row operations below joining $C_\mathbf{x}$ to some row-echelon form $C_\mathbf{x}^\sharp$ below:—

$$C_\mathbf{x} = \left[\begin{array}{cccc|c} 1 & 0 & 0 & 0 & x_1 \\ 2 & 1 & 0 & 0 & x_2 \\ 3 & 2 & 1 & 0 & x_3 \\ 4 & 3 & 2 & 1 & x_4 \end{array}\right] \xrightarrow{-2R_1+R_2} \xrightarrow{-3R_1+R_3} \xrightarrow{-4R_1+R_4}$$

$$\xrightarrow{-2R_2+R_3} \xrightarrow{-3R_2+R_4} \xrightarrow{-2R_3+R_4} \left[\begin{array}{cccc|c} 1 & 0 & 0 & 0 & x_1 \\ 0 & 1 & 0 & 0 & -2x_1 + x_2 \\ 0 & 0 & 1 & 0 & x_1 - 2x_2 + x_3 \\ 0 & 0 & 0 & 1 & x_2 - 2x_3 + x_4 \end{array}\right] = C_\mathbf{x}^\sharp.$$

The last column of $C_\mathbf{x}^\sharp$ is a free column. It follows that $\mathcal{LS}(V,\ \mathbf{x})$ is consistent.
Hence $\mathbf{x}$ is a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4$.

iii. It follows that $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4$ constitute a basis for $\mathbb{R}^4$ over the reals.

From the calculations above, we see that for any $\mathbf{x} \in \mathbb{R}^4$, if $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$ then

$$\mathbf{x} = x_1\mathbf{v}_1 + (-2x_1 + x_2)\mathbf{v}_2 + (x_1 - 2x_2 + x_3)\mathbf{v}_3 + (x_2 - 2x_3 + x_4)\mathbf{v}_4.$$

This expression of such an $\mathbf{x}$ as a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4$ is uniquely determined.

(c) *Define* $\mathbf{w}_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}$, $\mathbf{w}_2 = \begin{bmatrix} 2 \\ 3 \\ 4 \\ 1 \end{bmatrix}$, $\mathbf{w}_3 = \begin{bmatrix} 3 \\ 4 \\ 1 \\ 2 \end{bmatrix}$, $\mathbf{w}_4 = \begin{bmatrix} 4 \\ 1 \\ 2 \\ 3 \end{bmatrix}$.

Write $W = [\ \mathbf{w}_1 \mid \mathbf{w}_2 \mid \mathbf{w}_3 \mid \mathbf{w}_4\ ]$.
Note that W is a square matrix. We verify that W is invertible:—

- Apply a sequence of row operations below to join W to some row-echelon form $W^\sharp$:—

$$W = \begin{bmatrix} 1 & 4 & 3 & 2 \\ 2 & 1 & 4 & 3 \\ 3 & 2 & 1 & 4 \\ 4 & 3 & 2 & 1 \end{bmatrix} \xrightarrow{-2R_1+R_2} \xrightarrow{-3R_1+R_2} \xrightarrow{-4R_1+R_2} \begin{bmatrix} 1 & 4 & 3 & 2 \\ 0 & -7 & -2 & -1 \\ 0 & -10 & -8 & -2 \\ 0 & -13 & -10 & -7 \end{bmatrix}$$

$$\xrightarrow{-2R_2+R_4} \xrightarrow{R_2 \leftrightarrow R_4} \begin{bmatrix} 1 & 4 & 3 & 2 \\ 0 & 1 & -6 & -5 \\ 0 & -7 & -2 & -1 \\ 0 & -10 & -8 & -2 \end{bmatrix} \xrightarrow{7R_2+R_3} \xrightarrow{10R_2+R_4} \xrightarrow{-\frac{1}{4}R_3} \xrightarrow{-\frac{1}{4}R_4} \begin{bmatrix} 1 & 4 & 3 & 2 \\ 0 & 1 & -6 & -5 \\ 0 & 0 & 11 & 9 \\ 0 & 0 & 17 & 13 \end{bmatrix}$$

$$\xrightarrow{-1R_3+R_4} \xrightarrow{-2R_4+R_3} \xrightarrow{6R_3+R_4} \begin{bmatrix} 1 & 4 & 3 & 2 \\ 0 & 1 & -6 & -5 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 10 \end{bmatrix} = W^\sharp$$

  Note that every column of $W^\sharp$ is a pivot column. Hence W is invertible.

Since W is invertible, the column vectors $\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3, \mathbf{w}_4$ constitute a basis for $\mathbb{R}^4$.

By direct computation (such as re-running the sequence of row operations above, starting with $[\ W \mid I_4\ ]$), we find that $W^{-1} = \dfrac{1}{40}\begin{bmatrix} -9 & 1 & 1 & 11 \\ 1 & 1 & 11 & -9 \\ 1 & 11 & -9 & 1 \\ 11 & -9 & 1 & 1 \end{bmatrix}$.

For any $\mathbf{x} \in \mathbb{R}^4$, if $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$ then

$$\mathbf{x} = W(W^{-1}\mathbf{x}) = \frac{-9x_1+x_2+x_3+11x_4}{40}\mathbf{w}_1 + \frac{x_1+x_2+11x_3-9x_4}{40}\mathbf{w}_2 + \frac{x_1+11x_2-9x_3+x_4}{40}\mathbf{w}_3 + \frac{11x_1-9x_2+x_3+x_4}{40}\mathbf{w}_4.$$

This expression of such an $\mathbf{x}$ as a linear combination of $\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3, \mathbf{w}_4$ is uniquely determined.

10. As seen in Example (1), there are many distinct bases for $\mathbb{R}^n$ over the reals. In the light of this, there are two natural questions to ask, for any two arbitrarily given bases of $\mathbb{R}^n$ over the reals:—

    (1) *How are the two bases compared with each other?*

    (2) *How are the expressions of the same column vector with n real entries as linear combinations of column vectors in the respective bases compared with each other?*

    We answer the two questions above in a slightly more general context, in the form of the result immediately below about bases for a general subspace of $\mathbb{R}^n$ over the reals.

    Its proof is provided in the *appendix*.

11. **Theorem (4). ('Change-of-basis' Theorem.)**

    Let $\mathcal{W}$ be a subspace of $\mathbb{R}^n$ over the reals.

    Suppose $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_p$ constitute a basis for $\mathcal{W}$ over the reals. Also suppose $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_p$ constitute a basis for $\mathcal{W}$ over the reals.
    Define $U = [\ \mathbf{u}_1 \mid \mathbf{u}_2 \mid \cdots \mid \mathbf{u}_p\ ]$, $V = [\ \mathbf{v}_1 \mid \mathbf{v}_2 \mid \cdots \mid \mathbf{v}_p\ ]$.
    Then the statements below hold:—

    (a) *There exists some unique invertible $(p \times p)$-square matrix S with real entries such that $U = VS$.*

    (b) *Let $\mathbf{x} \in \mathcal{W}$, and $\alpha_1, \alpha_2, \cdots, \alpha_p, \beta_1, \beta_2, \cdots, \beta_p \in \mathbb{R}$.*
    Suppose $\mathbf{x} = \alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \cdots + \alpha_p\mathbf{u}_p$ and $\mathbf{x} = \beta_1\mathbf{v}_1 + \beta_2\mathbf{v}_2 + \cdots + \beta_p\mathbf{v}_p$.

    Then $\begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{bmatrix} = S \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_p \end{bmatrix}$, *for the same invertible $(p \times p)$-square matrix S above.*

    Remark on terminologies. The matrix S, which relates the matrices U, V via the equality $U = VS$, is called the **change-of-basis matrix, relating the (ordered) basis $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_p$ for $\mathcal{W}$ to the (ordered) basis $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_p$ for $\mathcal{W}$.**

    Further remark. Because of the invertibility of S, the equality $V = US^{-1}$ holds. Hence the invertible matrix S^{-1} is the change-of-basis matrix relating the (ordered) basis $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_p$ for $\mathcal{W}$ to the (ordered) basis $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_p$ for $\mathcal{W}$.

12. **Further comments on Theorem (4).**

(a) Suppose $\mathcal{W} = \mathbb{R}^n$. Further suppose $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n$ constitute a basis for $\mathbb{R}^n$ over the reals. Also suppose $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_p$ constitute a basis for $\mathbb{R}^n$ over the reals.

Then U, V are in fact invertible $(n \times n)$-square matrices.

The uniquely determined square matrix S described in part (a) of the conclusion of Theorem (4) is actually given by $S = V^{-1}U$.

(b) In Theorem (4) we impose the seemingly 'restrictive' assumption that the number of column vectors in one basis for $\mathcal{W}$, namely $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_p$, is the same as that of the number of column vectors in another basis for $\mathcal{W}$, namely, $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_p$.

You may wonder what may happen if we do not assume these two bases have the same number of column vectors.
This question will turn out a moot one.
It actually happens that any two bases for the same subspace of $\mathbb{R}^n$ over the reals have the same number of column vectors. (We will state this result more formally later on.)

13. **Example (2). (Illustration on the content of Theorem (4), in the special case '$\mathbf{W} = \mathbb{R}^n$ and $p = n$'.)**

We have the standard base $\mathbf{e}_1^{(4)}, \mathbf{e}_2^{(4)}, \mathbf{e}_3^{(4)}, \mathbf{e}_4^{(4)}$ for $\mathbb{R}^4$ over the reals.

Note that $I_4 = \left[\ \mathbf{e}_1^{(4)} \mid \mathbf{e}_2^{(4)} \mid \mathbf{e}_3^{(4)} \mid \mathbf{e}_4^{(4)} \ \right]$.

These are also bases for $\mathbb{R}^4$ over the reals:—

- $\mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \mathbf{u}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}, \mathbf{u}_4 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$.
- $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 2 \\ 3 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 2 \end{bmatrix}, \mathbf{v}_4 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$.
- $\mathbf{w}_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}, \mathbf{w}_2 = \begin{bmatrix} 2 \\ 3 \\ 4 \\ 1 \end{bmatrix}, \mathbf{w}_3 = \begin{bmatrix} 3 \\ 4 \\ 1 \\ 2 \end{bmatrix}, \mathbf{w}_4 = \begin{bmatrix} 4 \\ 1 \\ 2 \\ 3 \end{bmatrix}$.

Write $U = [\ \mathbf{u}_1 \mid \mathbf{u}_2 \mid \mathbf{u}_3 \mid \mathbf{u}_4\ ]$, $V = [\ \mathbf{v}_1 \mid \mathbf{v}_2 \mid \mathbf{v}_3 \mid \mathbf{v}_4\ ]$, $W = [\ \mathbf{w}_1 \mid \mathbf{w}_2 \mid \mathbf{w}_3 \mid \mathbf{w}_4\ ]$.

Note that U, V, W are invertible matrices.

(a) We can provide another interpretation for the equalities $I_4 = UU^{-1}$, $I_4 = VV^{-1}$, $I_4 = WW^{-1}$ according to Theorem (4):—

i. The matrix $U^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \end{bmatrix}$ is the change-of-basis matrix relating the (ordered) basis $\mathbf{e}_1^{(4)}, \mathbf{e}_2^{(4)}, \mathbf{e}_3^{(4)}, \mathbf{e}_4^{(4)}$ for $\mathbb{R}^4$ to the (ordered) basis $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4$ for $\mathbb{R}^4$.

ii. The matrix $V^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 \\ 1 & -2 & 1 & 0 \\ 0 & 1 & -2 & 1 \end{bmatrix}$ is the change-of-basis matrix relating the (ordered) basis $\mathbf{e}_1^{(4)}, \mathbf{e}_2^{(4)}, \mathbf{e}_3^{(4)}, \mathbf{e}_4^{(4)}$ for $\mathbb{R}^4$ to the (ordered) basis $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4$ for $\mathbb{R}^4$.

iii. The matrix $W^{-1} = \dfrac{1}{40}\begin{bmatrix} -9 & 1 & 1 & 11 \\ 1 & 1 & 11 & -9 \\ 1 & 11 & -9 & 1 \\ 11 & -9 & 1 & 1 \end{bmatrix}$ is the change-of-basis matrix relating the (ordered) basis $\mathbf{e}_1^{(4)}, \mathbf{e}_2^{(4)}, \mathbf{e}_3^{(4)}, \mathbf{e}_4^{(4)}$ for $\mathbb{R}^4$ to the (ordered) basis $\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3, \mathbf{w}_4$ for $\mathbb{R}^4$.

(b) Write $S = V^{-1}U$.

i. Note that $S = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \end{bmatrix}$. The equality $U = VS$ holds.

S is the change-of-basis matrix relating the (ordered) basis $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4$ for $\mathbb{R}^4$ to the (ordered) basis $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4$ for $\mathbb{R}^4$.

ii. Note that $S^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix}$. The equality $V = US^{-1}$ holds.

S^{-1} is the change-of-basis matrix relating the (ordered) basis $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4$ for $\mathbb{R}^4$ to the (ordered) basis $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4$ for $\mathbb{R}^4$.

(c) Write $Q = W^{-1}U$.

i. Note that $Q = \dfrac{1}{40}\begin{bmatrix} 4 & 13 & 12 & 11 \\ 4 & 3 & 2 & -9 \\ 4 & 3 & -8 & 1 \\ 4 & -7 & 2 & 1 \end{bmatrix}$. The equality $U = WQ$ holds.

Q is the change-of-basis matrix relating the (ordered) basis $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4$ for $\mathbb{R}^4$ to the (ordered) basis $\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3, \mathbf{w}_4$ for $\mathbb{R}^4$.

ii. Note that $Q^{-1} = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 1 & 1 & 1 & -3 \\ 1 & 1 & -3 & 1 \\ 1 & -3 & 1 & 1 \end{bmatrix}$. The equality $W = UQ^{-1}$ holds.

Q^{-1} is the change-of-basis matrix relating the (ordered) basis $\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3, \mathbf{w}_4$ for $\mathbb{R}^4$ to the (ordered) basis $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4$ for $\mathbb{R}^4$.

(d) Write $P = W^{-1}V$.

i. Note that $P = \dfrac{1}{40}\begin{bmatrix} 40 & 36 & 23 & 11 \\ 0 & -4 & -7 & -9 \\ 0 & -4 & -7 & 1 \\ 0 & -4 & 3 & 1 \end{bmatrix}$. The equality $V = WP$ holds.

P is the change-of-basis matrix relating the (ordered) basis $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4$ for $\mathbb{R}^4$ to the (ordered) basis $\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3, \mathbf{w}_4$ for $\mathbb{R}^4$.

ii. Note that $P^{-1} = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & -1 & -2 & -7 \\ 0 & 0 & -4 & 4 \\ 0 & -4 & 4 & 0 \end{bmatrix}$. The equality $V = WP^{-1}$ holds.

P^{-1} is the change-of-basis matrix relating the (ordered) basis $\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3, \mathbf{w}_4$ for $\mathbb{R}^4$ to the (ordered) basis $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4$ for $\mathbb{R}^4$.

14. The illustrations in Example (3) are less trivial than that in Example (2) in the sense that they are concerned with proper subspaces of $\mathbb{R}^n$, and that the interplay between the vector equalities and the matrix equalities captures the 'algebraic essesnce' in the argument for Theorem (4).

Example (3). (Illustrations for the content of Theorem (4), and the 'algebra' in its argument.)

(a) *Let* $\mathbf{u}_1 = \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}$, $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$, *and* $\mathcal{W} = \mathrm{Span}(\{\mathbf{u}_1, \mathbf{u}_2\})$.

Write $U = [\ \mathbf{u}_1 \mid \mathbf{u}_2\ ]$, $V = [\ \mathbf{v}_1 \mid \mathbf{v}_2\ ]$.

We can verify that $\mathbf{u}_1, \mathbf{u}_2$ are linearly independent over the reals. They constitute a basis for $\mathcal{W}$ over the reals. We can also verify that $\mathbf{v}_1, \mathbf{v}_2$ constitute a basis for $\mathcal{W}$ over the reals.

i. We have the pair of vector equalities below:—

$$\begin{cases} \mathbf{u}_1 = \mathbf{v}_1 + \mathbf{v}_2 \\ \mathbf{u}_2 = -\mathbf{v}_1 + \mathbf{v}_2 \end{cases}$$

This pair of vector equalities can be encoded into the matrix equality $U = VS$, in which $S = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$.

We can verify directly from the definition of invertibility that S is indeed invertible.

The matrix S is the change-of-basis matrix relating the (ordered) basis $\mathbf{u}_1, \mathbf{u}_2$ for $\mathcal{W}$ to the (ordered) basis $\mathbf{v}_1, \mathbf{v}_2$ for $\mathcal{W}$.

ii. We have the matrix equality $V = US^{-1}$, in which $S^{-1} = \begin{bmatrix} 1/2 & 1/2 \\ -1/2 & 1/2 \end{bmatrix}$.

This matrix equality in fact encodes this pair of vector equalities:—

$$\begin{cases} \mathbf{v}_1 = \dfrac{1}{2}\mathbf{u}_1 - \dfrac{1}{2}\mathbf{u}_2 \\ \\ \mathbf{v}_2 = \dfrac{1}{2}\mathbf{u}_1 + \dfrac{1}{2}\mathbf{u}_2 \end{cases}$$

The matrix S^{-1} is the change-of-basis matrix relating the (ordered) basis $\mathbf{v}_1, \mathbf{v}_2$ for $\mathcal{W}$ to the (ordered) basis $\mathbf{u}_1, \mathbf{u}_2$ for $\mathcal{W}$.

(b) *Let* $\mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} 0 \\ 1 \\ 2 \\ 3 \end{bmatrix}$, $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 2 \\ 3 \\ 4 \\ 5 \end{bmatrix}$, *and* $\mathcal{W} = \mathrm{Span}(\{\mathbf{u}_1, \mathbf{u}_2\})$.

Write $U = [\ \mathbf{u}_1 \mid \mathbf{u}_2\ ]$, $V = [\ \mathbf{v}_1 \mid \mathbf{v}_2\ ]$.

We can verify that $\mathbf{u}_1, \mathbf{u}_2$ are linearly independent over the reals. They constitute a basis for $\mathcal{W}$ over the reals.
We can also verify that $\mathbf{v}_1, \mathbf{v}_2$ constitute a basis for $\mathcal{W}$ over the reals.

i. We have the pair of vector equalities below:—

$$\left\{ \begin{array}{rcrcr} \mathbf{u}_1 & = & -\mathbf{v}_1 & + & \mathbf{v}_2 \\ \mathbf{u}_2 & = & 2\mathbf{v}_1 & - & \mathbf{v}_2 \end{array} \right.$$

This pair of vector equalities can be encoded into the matrix equality $U = VS$, in which $S = \begin{bmatrix} -1 & 2 \\ 1 & -1 \end{bmatrix}$.
We can verify directly from the definition of invertibility that S is indeed invertible.
The matrix S is the change-of-basis matrix relating the (ordered) basis $\mathbf{u}_1, \mathbf{u}_2$ for $\mathcal{W}$ to the (ordered) basis $\mathbf{v}_1, \mathbf{v}_2$ for $\mathcal{W}$.

ii. We have the matrix equality $V = US^{-1}$, in which $S^{-1} = \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix}$.
This matrix equality in fact encodes this pair of vector equalities:—

$$\left\{ \begin{array}{rcrcr} \mathbf{v}_1 & = & \mathbf{u}_1 & + & \mathbf{u}_2 \\ \mathbf{v}_2 & = & 2\mathbf{u}_1 & + & \mathbf{u}_2 \end{array} \right.$$

The matrix S is the change-of-basis matrix relating the (ordered) basis $\mathbf{v}_1, \mathbf{v}_2$ for $\mathcal{W}$ to the (ordered) basis $\mathbf{u}_1, \mathbf{u}_2$ for $\mathcal{W}$.

(c) *Let* $\mathbf{u}_1 = \begin{bmatrix} 2 \\ 2 \\ 2 \\ 1 \\ 1 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}$, $\mathbf{u}_3 = \begin{bmatrix} 0 \\ -2 \\ 0 \\ -1 \\ 1 \end{bmatrix}$, $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}$, $\mathbf{v}_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}$, *and* $\mathcal{W} = \mathrm{Span}(\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\})$.

Write $U = [\ \mathbf{u}_1 \mid \mathbf{u}_2 \mid \mathbf{u}_3\ ]$, $V = [\ \mathbf{v}_1 \mid \mathbf{v}_2 \mid \mathbf{v}_3\ ]$.

We can verify that $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ are linearly independent over the reals. They constitute a basis for $\mathcal{W}$ over the reals.
We can also verify that $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ constitute a basis for $\mathcal{W}$ over the reals.

i. We have the triple of vector equalities below:—

$$\left\{ \begin{array}{rcrcrcr} \mathbf{u}_1 & = & \mathbf{v}_1 & + & \mathbf{v}_2 & + & \mathbf{v}_3 \\ \mathbf{u}_2 & = & \mathbf{v}_1 & + & \mathbf{v}_2 & - & \mathbf{v}_3 \\ \mathbf{u}_3 & = & \mathbf{v}_1 & - & \mathbf{v}_2 & - & \mathbf{v}_3 \end{array} \right.$$

This pair of vector equalities can be encoded into the matrix equality $U = VS$, in which $S = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & -1 \\ 1 & -1 & -1 \end{bmatrix}$.
We can verify directly from the definition of invertibility that S is indeed invertible.
The matrix S is the change-of-basis matrix relating the (ordered) basis $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ for $\mathcal{W}$ to the (ordered) basis $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ for $\mathcal{W}$.

ii. We have the matrix equality $V = US^{-1}$, in which $S^{-1} = \begin{bmatrix} 1/2 & 0 & 1/2 \\ 0 & 1/2 & -1/2 \\ 1/2 & -1/2 & 0 \end{bmatrix}$.
This matrix equality in fact encodes this triple of vector equalities:—

$$\left\{ \begin{array}{rcrcrcr} \mathbf{v}_1 & = & \frac{1}{2}\mathbf{u}_1 & & & + & \frac{1}{2}\mathbf{u}_3 \\ \mathbf{v}_2 & = & & & \frac{1}{2}\mathbf{u}_2 & - & \frac{1}{2}\mathbf{u}_3 \\ \mathbf{v}_3 & = & \frac{1}{2}\mathbf{u}_1 & - & \frac{1}{2}\mathbf{u}_2 & & \end{array} \right.$$

The matrix S is the change-of-basis matrix relating the (ordered) basis $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ for $\mathcal{W}$ to the (ordered) basis $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ for $\mathcal{W}$.