MATH1030 Bases for subspaces of $\mathbb{R}^n$.

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

   Let V be a subspace of $\mathbb{R}^n$. Suppose V is not the zero subspace of $\mathbb{R}^n$.

   We declare that if V is the zero subspace of $\mathbb{R}^n$ then the empty set is the basis for V.

   From now on suppose V is not the zero subspace of $\mathbb{R}^n$.

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

   The vectors $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_p$ are said to constitute a basis for V (or the set $\{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_p\}$ is said to be a basis for V) 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.*

   (BS) *Every vector in V is a linear combination of $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_p$.*

   Remarks.

   (a) In the set-up of this definition, V is assumed to be a subspace of $\mathbb{R}^n$. Then it is trivially true that every linear combination of $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_p$ is a vector in V.

   For this reason, the statement (BS) holds if and only if $V = \mathsf{Span}\,(\{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_p\})$.

   In fact, some people will replace (BS) by

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

   in the definition for the notion of *basis* above.

   (b) In books where set language is used thoroughly, and '*span of general sets*' are defined, the 'declaration' that *the empty set is the basis for the zero subspace* can be incorporated naturally into the rest of the definition.

2. **Example of basis: Standard base for $\mathbb{R}^n$.**

   Fix any positive integer n.

   For each $k = 1, 2, \cdots, n$, denote by $\mathbf{e}_k^{(n)}$ the vector in $\mathbb{R}^n$ whose k-th entry is 1 and whose every other entry is 0.

   $$\text{(So } \mathbf{e}_k^{(n)} = E_{k,1}^{n,1} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.\text{)}$$

   The n vectors $\mathbf{e}_1^{(n)}, \mathbf{e}_2^{(n)}, \cdots, \mathbf{e}_n^{(n)}$ are collectively called the standard base for $\mathbb{R}^n$.

3. **Theorem (1).**

   Let V be a subspace of $\mathbb{R}^n$.

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

   The statements below are logically equivalent:

   ($\sharp$) $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_p$ *constitute a basis for V.*

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

   Remark. The 'existence-and-uniqueness statement'

   '*For any $\mathbf{x} \in V$, there exists some unique $\alpha_1, \alpha_2, \cdots, \alpha_p \in \mathbb{R}$ such that $\mathbf{x} = \alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \cdots + \alpha_p\mathbf{u}_p$*'

   is to be understood as a very terse presentation of the passage below:

   Both statements (E), (U) *are true:*

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

   (U) *For any $\mathbf{x} \in V$, for any $\beta_1, \beta_2, \cdots, \beta_p, \gamma_1, \gamma_2, \cdots, \gamma_p \in \mathbb{R}$, if $\mathbf{x} = \beta_1\mathbf{u}_1 + \beta_2\mathbf{u}_2 + \cdots + \beta_p\mathbf{u}_p$ and $\mathbf{x} = \gamma_1\mathbf{u}_1 + \gamma_2\mathbf{u}_2 + \cdots + \gamma_p\mathbf{u}_p$ then $\beta_1 = \gamma_1$, $\beta_2 = \gamma_2$, ..., $\beta_p = \gamma_p$.*

   Further remark. The significance of Theorem (1) is that it allows us to think of a subspace of $\mathbb{R}^n$, say, V, with a basis, say, $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_p$ as a copy of $\mathbb{R}^p$, by setting up a 'dictionary' between the subspace V of $\mathbb{R}^n$ and the subspace $\mathbb{R}^p$ of $\mathbb{R}^p$. This 'dictionary' is described below:

For each $\mathbf{x} \in V$*, we identify* $\mathbf{x}$ *as the vector* $\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_p \end{bmatrix}$ *exactly when the vector* $\mathbf{x}$ *is expressed as the uniquely determined linear combination* $\mathbf{x} = \alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \cdots + \alpha_p\mathbf{u}_p$.

Vector addition and scalar multiplication are preserved in the following sense:

- Suppose the vectors $\mathbf{x}, \mathbf{y}$ of V are 'identified' as the respective vectors $\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_p \end{bmatrix}, \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{bmatrix}$.

  Then the vector $\mathbf{x} + \mathbf{y}$ of V is 'identified' as the vector $\begin{bmatrix} \alpha_1 + \beta_1 \\ \alpha_2 + \beta_2 \\ \vdots \\ \alpha_p + \beta_p \end{bmatrix}$, which is in fact the sum of $\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_p \end{bmatrix}, \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{bmatrix}$.

  Moreover, for any real number γ, the vector $\gamma\mathbf{x}$ of V is 'identified' as the vector $\begin{bmatrix} \gamma\alpha_1 \\ \gamma\alpha_2 \\ \vdots \\ \gamma\alpha_p \end{bmatrix}$, which is in fact the scalar multiple $\gamma \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_p \end{bmatrix}$.

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

   Let V be a subspace in $\mathbb{R}^n$. Suppose Suppose $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_p$ be vectors in V.

   We want to verify that the statement ($\sharp$), ($\flat$) are logically equivalent:

   ($\sharp$) $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_p$ *constitute a basis for* V.

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

   - Suppose ($\sharp$) holds. [We want to verify that ($\flat$) holds.]
     Pick any $\mathbf{x} \in V$.
     * By (BS), $\mathbf{x}$ is a linear combination of $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_p$. Then there exist some $\alpha_1, \alpha_2, \cdots, \alpha_p \in \mathbb{R}$ such that $\mathbf{x} = \alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \cdots + \alpha_p\mathbf{u}_p$.
     * Pick any $\beta_1, \beta_2, \cdots, \beta_p \in \mathbb{R}$. Suppose $\mathbf{x} = \beta_1\mathbf{u}_1 + \beta_2\mathbf{u}_2 + \cdots + \beta_p\mathbf{u}_p$.
       Then $\alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \cdots + \alpha_p\mathbf{u}_p = \mathbf{x} = \beta_1\mathbf{u}_1 + \beta_2\mathbf{u}_2 + \cdots + \beta_p\mathbf{u}_p$.
       Therefore
       $$\begin{aligned} (\beta_1 - \alpha_1)\mathbf{u}_1 + (\beta_2 - \alpha)\mathbf{u}_2 + \cdots + (\beta_p - \alpha_p)\mathbf{u}_p &= (\beta_1\mathbf{u}_1 + \beta_2\mathbf{u}_2 + \cdots + \beta_p\mathbf{u}_p) - (\alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \cdots + \alpha_p\mathbf{u}_p) \\ &= \mathbf{x} - \mathbf{x} = \mathbf{0} \end{aligned}$$
       By (BL), $\beta_1 - \alpha_1 = \beta_2 - \alpha_2 = \cdots = \beta_p - \alpha_p = 0$.
       Hence $\alpha_1 = \beta_1$, $\alpha_2 = \beta_2$, ..., $\alpha_p = \beta_p$.

     Hence ($\flat$) holds.
   - Suppose ($\flat$) holds.
     that for any $\mathbf{x} \in V$, there exist some unique $\alpha_1, \alpha_2, \cdots, \alpha_p \in \mathbb{R}$ such that $\mathbf{x} = \alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \cdots + \alpha_p\mathbf{u}_p$.
     * Pick any $\mathbf{x} \in V$. Then by ($\flat$), there exist some $\alpha_1, \alpha_2, \cdots, \alpha_p \in \mathbb{R}$ such that $\mathbf{x} = \alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \cdots + \alpha_p\mathbf{u}_p$.
       Therefore (BS) holds.
     * Pick any $\beta_1, \beta_2, \cdots, \beta_p \in \mathbb{R}$. Suppose $\beta_1\mathbf{u}_1 + \beta_2\mathbf{u}_2 + \cdots + \beta_p\mathbf{u}_p = \mathbf{0}$.
       Note that $\mathbf{0} = 0 \cdot \mathbf{u}_1 + 0 \cdot \mathbf{u}_2 + \cdots + 0 \cdot \mathbf{u}_p$.
       Then by ($\flat$), we have $\beta_1 = \beta_2 = \cdots = \beta_p = 0$.
       Therefore (BL) holds.

     Hence ($\flat$) holds.

5. **Theorem (2). (Re-formulation of the notion of basis in terms of systems of equations.)**

   Let V *be a non-zero subspace of* $\mathbb{R}^n$.

   Suppose $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_p$ *be vectors in* V*, and* U *is the* $(n \times p)$*-matrix given by* $U = [\ \mathbf{u}_1 \mid \mathbf{u}_2 \mid \cdots \mid \mathbf{u}_p\ ]$.

   The statements below are logically equivalent:

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

(b) *Both statements* (BL1), (BS1) *are true:*

(BL1) *The homogeneous system* $\mathcal{LS}(U, \mathbf{0})$ *has no non-trivial solution.*

(BS1) *For any* $\mathbf{b} \in V$, *the system* $\mathcal{LS}(U, \mathbf{b})$ *is consistent.*

Remark. The re-formulation in terms of systems of equations is not something convenient to use in practice.

Proof of Theorem (2). This is a direct consequence of the application of the respective 'dictionaries' between linear combinations and systems of linear equations, and between linear dependence and systems of linear equations.

6. **'Dictionary' between non-singular $(n \times n)$-square matrices and basis for $\mathbb{R}^n$.**

Recall the result $(\star)$ from the handout *How to determine whether a given vector is the linear combination of some vectors*, and the result $(\star\star)$ from the handout *Linear dependence and linear independence*:

$(\star)$ *Suppose* $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n$ *are vectors in* $\mathbb{R}^n$, *and* U *is the* $(n \times n)$*-square matrix given by* $U = [\ \mathbf{u}_1 \mid \mathbf{u}_2 \mid \cdots \mid \mathbf{u}_n\ ]$. *Then the statements below are logically equivalent:*

(a) *Every vector in* $\mathbb{R}^n$ *is a linear combination of* $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n$.

(b) U *is non-singular.*

(c) U *is invertible.*

$(\star\star)$ *Suppose* $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n$ *are vectors in* $\mathbb{R}^n$, *and* U *is the* $(n \times n)$*-square matrix given by* $U = [\ \mathbf{u}_1 \mid \mathbf{u}_2 \mid \cdots \mid \mathbf{u}_n\ ]$. *Then the statements below are logically equivalent:*

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

(b) U *is non-singular.*

(c) U *is invertible.*

The results $(\star)$ and $(\star\star)$ to give Theorem (3) below.

Theorem (3).

Suppose $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n$ *are vectors in* $\mathbb{R}^n$, *and* U *is the* $(n \times n)$*-square matrix given by* $U = [\ \mathbf{u}_1 \mid \mathbf{u}_2 \mid \cdots \mid \mathbf{u}_n\ ]$.
Then the statements below are logically equivalent:

(a) U *is non-singular.*

(b) U *is invertible.*

(c) *Every vector in* $\mathbb{R}^n$ *is a linear combination of* $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n$.

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

(e) $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n$ *constitute a basis for* $\mathbb{R}^n$.

Remark. This result will be merged with Theorem (E) in the Handout *Existence and uniqueness of solutions for a system of linear equations whose coefficient matrix is a square matrix* later, alongside more re-formulations for the notion of *non-singularity*.

7. **Theorem (A).**

Suppose V *is a subspace of* $\mathbb{R}^n$. *Then every basis for* V *has at most* n *vectors.*

Proof of Theorem (A).

Suppose V is a subspace of $\mathbb{R}^n$.

- If V is the zero subspace of $\mathbb{R}^n$ then its only basis, namely the empty set, has no vectors in it.
- From now on suppose V is not the zero subspace of $\mathbb{R}^n$. Suppose $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_p$ constitute a basis for V.
  By definition, $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_p$ are vectors in $\mathbb{R}^n$, and they are linearly independent.
  Then $p \leq n$.

8. **Theorem (B).**

Any two bases for a subspace of $\mathbb{R}^n$ *have the same number of vectors.*

Proof of Theorem (B). Postponed. (This result is a consequence of the 'Replacement Theorem'.)

Remark. In the light of the validity of this result, it makes sense to talk about the *dimension of a subspace of* $\mathbb{R}^n$, which is introduced later.

9. **Theorem (C).**

   Suppose V is a non-zero subspace of $\mathbb{R}^n$. Then V has a basis which consists of at least one and at most n vectors in $\mathbb{R}^n$.

   Comment on the significance of Theorem (C).

   We have already known that:

   - the null space of a matrix with n columns is a subspace of $\mathbb{R}^n$, and the span of several vectors of $\mathbb{R}^n$ is a subspace of $\mathbb{R}^n$, and furthermore,
   - the null space of a matrix with n columns is the span of some vectors in $\mathbb{R}^n$, while the span of several vectors in $\mathbb{R}^n$ is the null space of some matrix with n columns.

   According to Theorem (C), a subspace in $\mathbb{R}^n$ is the span of some vectors in $\mathbb{R}^n$. It follows that it is also the null space of some matrix with n columns.

   So the notions of *subspace*, *null space*, *span*, *column space* are manifestations of the same mathematical concept.

10. **Preparation for the proof of Theorem (C).**

    As preparation for the proof of Theorem (C), recall the result $(*)$ below, from the handout *More on linear dependence and linear independence*:

    $(*)$ *Let $\mathbf{w}_1, \mathbf{w}_2, \cdots, \mathbf{w}_k, \mathbf{v}$ be vectors in $\mathbb{R}^n$.*
    Suppose $\mathbf{w}_1, \mathbf{w}_2, \cdots, \mathbf{w}_k$ are linearly independent.
    Then the statements below are logically equivalent:

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

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

    Also recall the result $(**)$ below, from the handout *Linear dependence and linear independence*:

    $(**)$ *Let $\mathbf{w}_1, \mathbf{w}_2, \cdots, \mathbf{w}_\ell$ be vectors in $\mathbb{R}^n$. Suppose $\mathbf{w}_1, \mathbf{w}_2, \cdots, \mathbf{w}_\ell$ are linearly independent. Then $\ell \leq n$.*

11. **Proof of Theorem (C).**

    Suppose V is a non-zero subspace of $\mathbb{R}^n$.

    By assumption there is some vector, say, $\mathbf{u}_1$, which is not the zero vector in V.

    $\mathbf{u}_1$ is linearly independent.

    If every vector in V is a linear combination of $\mathbf{u}_1$ then, $\mathbf{u}_1$ constitutes a basis for V.

    Suppose that not every vector in V is a linear combination of $\mathbf{u}_1$. Then there is some vector in V, say, $\mathbf{u}_2$, so that $\mathbf{u}_2$ is not a linear combination of $\mathbf{u}_1$.

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

    If every vector in V is a linear combination of $\mathbf{u}_1, \mathbf{u}_2$ then, $\mathbf{u}_1, \mathbf{u}_2$ constitute a basis for V.

    Suppose that not every vector in V is a linear combination of $\mathbf{u}_1, \mathbf{u}_2$. Then there is some vector in V, say, $\mathbf{u}_3$, so that $\mathbf{u}_3$ is not a linear combination of $\mathbf{u}_1, \mathbf{u}_2$.

    By $(*)$, $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ are linearly independent.

    By repeating the above construction for j times, we obtain, in succession, some vectors $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_j$ in V, which are linearly independent vectors in $\mathbb{R}^n$.

    By $(**)$, we have $j \leq n$. So there is the last time, say, the p-th time of the construction. We have obtained the vectors $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_p$ in V, which are linearly independent vectors in $\mathbb{R}^n$.

    It is then necessarily true that every vector in V is a linear combination of $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_p$. (Otherwise, we could repeat the construction for the $(p+1)$-th time. That would be a contradiction.)

    It follows that the p vectors $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_p$ constitute a basis for V.