Math 416 Lecture Note: Week 4

Daesung Kim

1 Bases and dimensions: definitions

Let V be a vector space over $\mathbb{R}$. It is interesting to select a few vectors in V that describe the whole space V. In other words, we want to find a subset $S \subseteq V$ that spans V. This means that every vector in V can be written as a linear combination of S. The goal of this section is to characterize an "optimal" spanning set S. If a spanning set S is optimal, one expects that the linear combination of a vector will be "unique". This leads to linear independence.

Definition 1.1. Let V be a vector space over $\mathbb{R}$. A basis β for V is a linearly independent subset of V that generates V. We say that the vectors of β form a basis for V.

Example 1.2. Let $V = \mathbb{R}^n$ and $e_i \in V$ be such that the i-th entry is 1 and others are zero. Then $E = \{e_1, \cdots, e_n\}$ is a basis for V. This basis is called the standard basis for $\mathbb{R}^n$. To see $\mathrm{Span}(E) = \mathbb{R}^n$, pick an arbitrary vector $x \in \mathbb{R}^n$ and describe it as a linear combination of E. It is easy to see that

$$x = (x_1, \cdots, x_n) = x_1 e_1 + \cdots + x_n e_n.$$

To prove E is linearly independent, consider a linear equation

$$a_1 e_1 + \cdots + a_n e_n = 0$$

and see if it has the only trivial solution ($a_1 = \cdots = a_n = 0$).

Example 1.3. Let $V = \mathcal{P}_n(\mathbb{R})$ be the set of all real polynomials $p(x)$ with $\deg(p) \le n$. Then, $\{1, x, x^2, \cdots, x^n\}$ is a basis for V.

Theorem 1.4. *Let V be a vector space over $\mathbb{R}$ and $\beta = \{v_1, \cdots, v_n\}$ a subset of V. Then, β is a basis for V if and only if $v \in V$ can be uniquely expressed as a linear combination of $v_1, \cdots, v_n$.*

Proof. ($\Rightarrow$): Let $v \in V$. Since β is a basis for V, $\mathrm{Span}(\beta) = V$ and v is a linear combination of $v_1, \cdots, v_n$. Suppose there are two linear combinations

$$v = a_1 v_1 + \cdots + a_n v_n = b_1 v_1 + \cdots + b_n v_n$$

for some $a_i, b_i \in \mathbb{R}$. It then follows that

$$(a_1 - b_1)v_1 + \cdots + (a_n - b_n)v_n = 0.$$

Since β is linearly independent, we have $(a_i - b_i) = 0$ for all i (see Problem 2 in HW 3).

($\Leftarrow$): Since every vector in V is a linear combination of β, $\mathrm{Span}(\beta) = V$. Let

$$a_1 v_1 + \cdots + a_n v_n = 0.$$

Since 0 has a unique linear combination representation and $0 = 0v_1 + \cdots + 0v_n$, we have $a_1 = a_2 = \cdots = a_n = 0$, which implies that β is linearly independent. Thus, β is a basis for V. $\square$

Notation 1.5. For a finite set G, we use the notation $|G|$ to denote the number of elements in G.

Theorem 1.6. *Let V be a vector space over $\mathbb{R}$, S a finite subset of V, and* $\operatorname{Span}(S) = V$*. Then, there exists a subset β of S that is a basis for V.*

Proof. If S is linearly independent, then $S \subseteq S$ is a basis for V. (Note: If $S = \varnothing$ or $S = \{v\}$ with $v \neq 0$, then S is linearly independent.) If $S = \{0\}$, then $V = \{0\}$. Let $\beta = \varnothing$. Suppose $n \geq 2$ and $S = \{v_1, v_2, \cdots, v_n\}$ is linearly dependent. Then, there exist a_i, not all zero, such that

$$a_1 v_1 + a_2 v_2 + \cdots + a_n v_n = 0.$$

After relabeling and manipulating, we get

$$v_n = a_1 v_1 + a_2 v_2 + \cdots + a_{n-1} v_{n-1}.$$

By the previous lemma, we have $\operatorname{Span}(\{v_1, v_2, \cdots, v_{n-1}\}) = V$. Repeat this procedure finitely many times until we get a linearly independent subset of S. □

Definition 1.7. The dimension of a vector space V is the number of vectors in a basis for V. The dimension is denoted by $\dim(V)$.

Question 1.8. Is the definition well-defined? Since the definition depends on the choice of a basis, we need to justify that the dimension is consistent regardless of the choice of a basis.

2 Bases and dimensions: Replacement theorem

The goal of this section is to show that the definition of dimension is well-defined. Indeed, we prove more stronger result, which is called Replacement theorem.

Theorem 2.1 (Replacement Theorem)**.** *Let V be a vector space over $\mathbb{R}$, G a subset of V with $|G| = n$ and $\operatorname{Span}(G) = V$. Let L be a linearly independent subset of V with $|L| = m$. Then,*

(i) *$m \leq n$, and*

(ii) *there exists a subset $H \subseteq G$ such that $|H| = n - m$ and $\operatorname{Span}(L \cup H) = V$.*

Proof. We use an induction on m. If $m = 0$, then $L = \varnothing$ and $0 \leq n$. We choose $H = G$ then $\operatorname{Span}(L \cup H) = V$. Let $m \geq 1$ and assume that the conclusions are true for $k < m$. Let $L = \{w_1, \cdots, w_m\}$. We apply the induction hypothesis to $L' = \{w_1, \cdots, w_{m-1}\}$ so that $m - 1 \leq n$ and there exists a subset $H' = \{h_1, \cdots, h_{n-m+1}\}$ of G such that $|H'| = n - m + 1$ and $\operatorname{Span}(L' \cup H') = V$. (Note that L' is linearly independent.) Then, w_m can be written as a linear combination of $L' \cup H'$

$$w_m = a_1 w_1 + \cdots + a_{m-1} w_{m-1} + b_1 h_1 + \cdots b_{n-m+1} h_{n-m+1}.$$

There exists a nonzero b_i since $\{w_1, \cdots, w_m\}$ is linearly independent. If $|H'| = 0$, then w_m is a linear combination of $w_1, \cdots, w_{m-1}$, which is a contradiction. Thus, $m \leq n$. Define $H = H' \setminus \{h_i\}$. Since

$$\operatorname{Span}(L' \cup H') = \operatorname{Span}(L \cup H') = \operatorname{Span}((L \cup H) \cup \{h_i\})$$

and $h_i \in \operatorname{Span}(L \cup H)$, it follows from the lemma that $\operatorname{Span}(L \cup H) = V$. □

Corollary 2.2. *Let V be a vector space over $\mathbb{R}$ having a finite basis. If β and γ are bases for V, then $|\beta| = |\gamma|$.*

Proof. Suppose $|\beta| \neq |\gamma|$. Without loss of generality, we assume $n = |\beta| < |\gamma| = m$. Let S be a subset of γ with $|S| = n + 1$, then S is linearly independent and β generates V. By Replacement theorem, $n + 1 \leq n$, a contradiction. □

Definition 2.3. A vector space is called finite-dimensional if $\dim(V) < \infty$ and infinite-dimensional otherwise.

Example 2.4. $\dim(\{0\}) = 0$, $\dim(\mathbb{R}^n) = n$, $\dim(\mathcal{M}_{m \times n}(\mathbb{R})) = mn$, and $\dim(\mathcal{P}_n(\mathbb{R})) = n + 1$.

Corollary 2.5. *Let V be a vector space over $\mathbb{R}$ with $\dim(V) = n$ and S a subset of V.*

(i) *If $\operatorname{Span}(S) = V$, then $|S| \geq n$.*

(ii) *If $\operatorname{Span}(S) = V$ and $|S| = n$, then S is a basis for V.*

(iii) *If S is linearly independent and $|S| = n$, then S is a basis for V.*

(iv) *If S is linearly independent, then there exists a basis $\widetilde{S}$ for V such that $S \subseteq \widetilde{S}$.*

Proof. Let β be a basis for V, then $|\beta| = n$.

(i) By Theorem 1.6, there exists a subset S' of S that is a basis for V. Thus, $n = |S'| \leq |S|$.

(ii) If $n = |S'| \leq |S| = n$, then $S' = S$ so that S is a basis for V.

(iii) By Theorem 2.1, there exists a subset H of β such that $|H| = n - n = 0$ and $S \cup H$ generates V. Since $H = \varnothing$, the result follows.

(iv) By Theorem 2.1, there exists a subset H of β such that $S \cup H$ generates V and $|H| = n - |S|$. Since $|S \cup H| \leq |S| + |H| = n$, (i) implies $|S \cup H| = n$. Let $\widetilde{S} = S \cup H$, then the result follows from (ii).

□

Theorem 2.6 (The dimension of subspaces)**.** *Let V be a finite dimensional vector space over $\mathbb{R}$ and W a subspace of V. Then, $\dim(W) \leq \dim(V)$. Moreover, if $\dim(W) = \dim(V)$, then $V = W$.*

Proof. Let $\dim(V) = n$ and β a basis for V. Since β spans W, there exists a subset $\gamma \subset \beta$ that is a basis for W by [FIS, Theorem 1.9]. Therefore, $\dim(W) = |\gamma| \leq |\beta| = \dim(V)$.

Suppose $\dim(W) = \dim(V) = n$. Since γ is linearly independent, [FIS, Corollary 2, p. 47] implies that γ is a basis for V. In particular, we have $W = \mathrm{Span}(\gamma) = V$. □

Corollary 2.7 (Basis Extension)**.** *Let V be a finite dimensional vector space over $\mathbb{R}$ and W a subspace of V. If β is a basis for W, then there exists a basis γ for V such that $\beta \subseteq \gamma$.*

Proof. By the corollary (iv) above. □

Example 2.8. Let $V = \mathcal{M}_{2\times 2}(\mathbb{R})$ be the set of all (2×2) matrices with real entries. Consider a subspace

$$W = \{A \in \mathcal{M}_{2\times 2}(\mathbb{R}) : A^t = A\} \leq V.$$

Every $S \in W$ can be written as

$$\begin{pmatrix} a & b \\ b & d \end{pmatrix}.$$

Let

$$\beta = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}$$

Then, one can see that β is a basis for W and the dimension of W is 3. We already know that the dimension of V is 4. Note that

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \notin W = \mathrm{Span}(\beta).$$

By [FIS, Theorem 1.7], $\beta \cup \{A\}$ is linearly independent. Since $|\beta \cup \{A\}| = 4 = \dim(V)$, it is actually a basis for V.

Corollary 2.9 (Change of Bases)**.** *Let V be a vector space over $\mathbb{R}$ with $\dim(V) = n \in \mathbb{N}$ and S a basis for V. Let $A = (A_{ij}) \in \mathcal{M}_{n\times n}(\mathbb{R})$ and define $T = \{w_1, \cdots, w_n\}$ where*

$$w_j = A_{1j}v_1 + \cdots + A_{nj}v_n = \sum_{i=1}^{n} A_{ij}v_i$$

for each $j = 1, 2, \cdots, n$. If the linear system associated to $(A, 0)$ has exactly one (trivial) solution, then T is a basis for V.

Proof. By Problem in Homework 3, T is linearly independent. Let $j \neq k$, suppose $w_j = w_k$. Then,

$$0 = w_j - w_k = \sum_{i=1}^{n} (A_{ij} - A_{ik})v_i.$$

Since S is linearly independent, $A_{ij} = A_{ik}$ for all $i = 1, 2, \cdots, n$. This is a contradiction. Thus, $|T| = n$. By Corollary 2.5, T is a basis for V. □

3 Bases and dimensions: Linear systems

Definition 3.1. A system of linear equations

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

is called *homogeneous* if $b_i = 0$ for all $i = 1, 2, \cdots, m$.

Definition 3.2. Let $A \in \mathcal{M}_{m\times n}(\mathbb{R})$. The null space of A is the set of all solutions to a homogenous linear system $LS(A, 0)$, denoted by $\mathcal{N}(A)$.

Remark 3.3. In Problem in Homework 2, it is shown that $\mathcal{N}(A)$ is a subspace of $\mathbb{R}^n$.

Question 3.4. Find the basis and the dimension of $\mathcal{N}(A)$.

Example 3.5. Consider (3×4) matrix

$$\begin{pmatrix} 1 & 0 & 8 & 0 \\ 0 & 1 & -3 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

and the linear system $LS(A, 0)$

$$\begin{cases} x_1 + 8x_3 = 0 \\ x_2 - 3x_3 = 0 \\ x_4 = 0. \end{cases}$$

The solution set is $\{t(-8, 3, 1, 0) : t \in \mathbb{R}\}$, which is a line passing through the origin and $(-8, 3, 1, 0)$. The basis is $\{(-8, 3, 1, 0)\}$ (when $t = 1$) and the dimension is 1.

Theorem 3.6. *Let $A \in \mathcal{M}_{m\times n}(\mathbb{R})$. If a RREF of A has p pivot columns, then the dimension of $\mathcal{N}(A)$ is $n - p$.*

Proof. Let R be a RREF of A and $R = (C_1, C_2, \cdots, C_n)$ where C_i is the i-th row of R for each i. Suppose that $C_{j(1)}, C_{j(2)}, \cdots, C_{j(n-p)}$ are non-pivot columns of R. Let e_i be the standard basis for $\mathbb{R}^n$. Let $x = (x_1, x_2, \cdots, x_n) \in \mathcal{N}(A)$. If C_l is a pivot column, then x_l is a linear combination of $x_{j(k)}$ for $j(k) \geq l$ because the matrix R is in reduced row-echelon form. Thus,

$$\begin{aligned} x &= x_1e_1 + x_2e_2 + \cdots + x_ne_n \\ &= x_{j(1)}v_{j(1)} + x_{j(2)}v_{j(2)} + \cdots + x_{j(n-p)}v_{j(n-p)} \end{aligned}$$

for some $v_{j(1)}, \cdots, v_{j(n-p)} \in \mathbb{R}^n$. Since x_l is a linear combination of $x_{j(k)}$ for $j(k) \geq l$, for each $k = 1, 2, \cdots, n - p$, $v_{j(k)}$ looks like

$$v_{j(k)} = (*, \cdots, *, 1, 0 \cdots, 0)$$

where 1 appears exactly at the $j(k)$-entry of $v_{j(k)}$ and 0 after that. Let $\beta = \{v_{j(k)} : k = 1, 2, \cdots, n - p\}$. We claim that β is a basis for $\mathcal{N}(A)$. By the observation above, β spans $\mathcal{N}(A)$. We need to show that β is linearly independent. Let

$$a_{j(1)}v_{j(1)} + a_{j(2)}v_{j(2)} + \cdots + a_{j(n-p)}v_{j(n-p)} = 0$$

If we look at the $j(n - p)$-th entry of $v_{j(k)}$, then $v_{j(n-p)}$ only has a nonzero entry, which means that $a_{j(n-p)} = 0$. Thus, we have

$$a_{j(1)}v_{j(1)} + a_{j(2)}v_{j(2)} + \cdots + a_{j(n-p-1)}v_{j(n-p-1)} = 0.$$

If we look at the $j(n - p - 1)$-th entry of $v_{j(k)}$ for $k \leq n - p - 1$, then $v_{j(n-p-1)}$ only has a nonzero entry so that $a_{j(n-p-1)} = 0$. Repeating this procedure, we get $a_{j(1)} = \cdots = a_{j(n-p)} = 0$. $\square$

Definition 3.7. The row space of a matrix $A \in \mathcal{M}_{m\times n}(\mathbb{R})$ is the space spanned by its rows in $\mathbb{R}^n$, denoted by $\mathcal{R}(A)$.

Example 3.8. Consider

$$A = \begin{pmatrix} 1 & 0 & -2 & 7 \\ 0 & 1 & 4 & -4 \\ 0 & 0 & 1 & -2 \\ 0 & 0 & 1 & -2 \end{pmatrix}$$

The set of rows spans $\mathcal{R}(A)$ by definition but is not linearly independent (because the 3rd and 4th columns are same). It is row-equivalent to

$$P = \begin{pmatrix} 1 & 0 & 0 & 3 \\ 0 & 1 & 0 & 4 \\ 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Then the set of all nonzero rows are linearly independent (Exercise).

Theorem 3.9. *Let $A, B \in \mathcal{M}_{m\times n}(\mathbb{R})$. If $A \sim B$, then $\mathcal{R}(A) = \mathcal{R}(B)$.*

Proof. It is enough to assume that B is obtained from A by a single row operation. Let A_i be the i-th row of A and write $A = (A_1, A_2, \cdots, A_m)^t$. Then $\mathcal{R}(A) = \mathrm{Span}(\{A_1, A_2, \cdots, A_m\})$. It is trivial that the first two operations does not change $\mathcal{R}(A)$. Thus, it suffices to show that for $S = \{A_1, A_2, \cdots, A_m\}$ and $T = \{A_1 + A_2, A_2, \cdots, A_m\}$,

$$\mathrm{Span}(S) = \mathrm{Span}(T).$$

Since $A_1 + A_2 \in \mathrm{Span}(S)$, $T \subset \mathrm{Span}(S)$. Since $\mathrm{Span}(T)$ is the smallest subspace containing T, $\mathrm{Span}(T) \subset \mathrm{Span}(S)$. Similarly, it follows from

$$(A_1 + A_2) - A_2 = A_1 \in \mathrm{Span}(T)$$

that $\mathrm{Span}(S) \subset \mathrm{Span}(T)$. $\square$

Theorem 3.10. *Let $A \in \mathcal{M}_{m\times n}(\mathbb{R})$ and P a RREF of A. Then, the set of nonzero rows of P forms a basis for $\mathcal{R}(A)$.*

Proof. Let $P = (R_1, R_2, \cdots, R_s, 0, \cdots, 0)^t$ where $s < m$ and R_i is a nonzero row of P for each $i = 1, 2, \cdots, s$. By definition, we have $\mathrm{Span}(\{R_1, R_2, \cdots, R_s\}) = \mathcal{R}(P) = \mathcal{R}(A)$. It is enough to show that the nonzero rows of P is linearly independent. Consider a system of linear equations

$$x_1 R_1 + \cdots + x_s R_s = 0.$$

Since P is in RREF, R_1 looks like

$$R_1 = (0, \cdots, 0, 1, *, \cdots, *).$$

Suppose the leading 1 of R_1 appears at the i-th entry. Then the i-th entries of all other rows $R_2, \cdots, R_s$ are zero, which means that $x_1 = 0$. In the same reason, we see that x_i are all zero. Thus, it is linearly independent. $\square$

References

[FIS] Freidberg, Insel, and Spence, *Linear Algebra*, 4th edition, 2002.

[Bee] Beezer, *A First Course in Linear Algebra*, Version 3.5, 2015.

Department of Mathematics, University of Illinois at Urbana-Champaign
E-mail address: daesungk@illinois.edu