Vector Spaces: Bases and Dimension

All material from Chapter 2 and 8 of Linear Algebra by Hoffman and Kunze.

Definitions: Let V be a vector space over F. A subset S of V is said to be **linearly independent** if for any finite set $\{\sigma_1, \ldots, \sigma_n\}$ of distinct elements of S, if

$$c_1\sigma_1 + \cdots + c_n\sigma_n = \vec{0}$$

implies

$$c_1 = c_2 = \cdots = c_n = 0.$$

$S \subset V$ is said to be **linearly dependent** if it is not linearly independent.

Proposition:

1. If $S \subset S'$ and S is linearly dependent, then S' is linearly dependent.
2. If $S \subset S'$ and S' is linearly independent, then S is linearly independent.
3. If $\vec{0} \in S$ the S is linearly dependent.
4. A set is linearly dependent if and only if some linear combination of its distinct elements equals the zero vector without all of the scalars equal to 0

Definition: A subset S of a vector space V is said to be a **basis** for V if S is linearly independent and the span of S is V. V is said to be **finite-dimensional** if it has a basis that is a finite set. If V is not finite-dimensional we say V is **infinite-dimensional**.

Important Example: Let F be any field, and let F^∞ be the set of infinite sequences of elements of F, that is, the set of all functions from $\{0, 1, \ldots,\}$ into F. We know this is a vector space over F. F^∞ is sometimes called the set of **formal power series** over F. Let

$$F[x] = \{\phi \in F^\infty : \phi_k \neq 0 \text{ for only finitely many } k\}$$

$F[x]$ is the **polynomials over** F. Finally, for each positive integer N let

$$F_N[x] = \{\phi \in F^\infty : \phi_k = 0 \text{ for } k > N\}.$$

$F_N[x]$ is the polynomials over F of **degree less than or equal to** N.

Define $\epsilon^{(j)} \in F^\infty$ by

$$\epsilon_k^{(j)} = \begin{cases} 0 & \text{if} \quad j \neq k \\ 1 & \text{if} \quad j = k \end{cases}$$

Both $F_N[x]$ and $F[x]$ are vector spaces over F.

The set $\{\epsilon^{(0)}, \cdots, \epsilon^{(N)}\}$ is a basis for $F_N[x]$. The set $\{\epsilon^{(0)}, \cdots, \epsilon^{(N)}, \ldots\}$ is a basis for $F[x]$, but is not a basis for F^∞ because linear combinations do not involve infinite sums. For this reason, the $\epsilon^{(j)}$ are called the **standard basis vectors** of $F[x]$.

(As vector spaces, we could consider $F_N[x]$ and F^{N+1} as the same, but they are not, as they describe functions with different domains. We will return to this when we consider vector space isomorphisms.)

Lemma: Let $A \in F^{n\times n}$. If $\det(A) = 0$ then $\{X \in F^{n\times 1} : AX = [0]\}$ contains at least two elements.

Demonstration: We prove this by induction on n. If $n = 1$ this is clear, for then $A = [0]$. So, suppose this is true for $n = k$. We need to show it is true for $n = k + 1$. We do know that $A\text{adj}(A) = [0]$, and that $A\vec{0} = \vec{0}$. We have to find some $X \neq \vec{0}$ with $AX = \vec{0}$. If some column of $\text{adj}(A) \neq \vec{0}$ we are done. If not, then $\det(A(1|j)) = 0$ for all $j \in \{1, \ldots, k + 1\}$.

With this in mind, let us introduce the following notation. First, let $U \in \{F^{(k+1)\times 1}$ have $U(1) = 1$ and $U(j) = 0$ for $j > 1$. Next, if $n > 1$ and $X \in F^{n\times 1}$, let $X^{(i)}$ be the element in F^{n-1} obtained from X by removing the element in row i of X. For example, if

$$X = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix},$$

then

$$\begin{aligned} X^{(1)} &= \begin{bmatrix} 2 \\ 3 \end{bmatrix} \\ X^{(2)} &= \begin{bmatrix} 1 \\ 3 \end{bmatrix} \\ X^{(3)} &= \begin{bmatrix} 1 \\ 2 \end{bmatrix} \end{aligned}$$

Returning to our argument, since $\det(A(1|i)) = 0$ for each $\in \{1, 2, \ldots, k+1\}$ there is a some $X_i \in F^{(k+1)\times 1}$ so that

- $X_i(i) = 0$ (the i^{th} component of X_i is 0);
- $X_i^{(i)} \neq \vec{0}$
- $A(1|i)X_i^{(i)} = \vec{0}$.
- $AX_i = x_i U$ for some $x_i \in F$.

To see this last assertion, note that if $j \neq 1$:

$$\begin{aligned} (AX_i)(j) &= \sum_{h=1}^{k+1} A(j,h)X_i(h) \\ &= \sum_{h=1, h\neq i}^{k+1} A(j,h)X_i(h) \\ &= \sum_{h=1, h\neq i}^{k+1} A(1|i)(j,h)X_i^{(i)}(h) \\ &= \left(A(1|i)X^{(i)}\right)(j) \\ &= 0. \end{aligned}$$

Now, if some $x_i = 0$, we are done. If $x_i \neq 0$ for all i we might as well assume that $AX_i = U$ for each $i \in \{1, \ldots, k+1\}$. Observe now that the X_i cannot all be the same, for then they would each equal $\vec{0}$. For example, if $k = 3$, we would have

$$\begin{bmatrix} 0 \\ w_2 \\ w_3 \\ w_4 \end{bmatrix} = \begin{bmatrix} x_1 \\ 0 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ 0 \\ y_4 \end{bmatrix} = \begin{bmatrix} z_1 \\ z_2 \\ z_3 \\ 0 \end{bmatrix}$$

So choose $X_i \neq X_j$ and put $X = X_i - X_j$. Then $AX = AX_i - AX_j = U - U = \vec{0}$.

Lemma: Let $m < n$ be positive integers, let F be a field, and let $A \in F^{m\times n}$. Then $\{X \in F^{n\times 1} : AX = [0]\}$ has at least two elements.

Note: This is the result that says that if there are more unknowns than equations in a homogeneous linear system, then the system has more than one solution.

Demonstration: We make another partitioning argument. We note that the equation $AX = \vec{0}$ is equivalent to an equation $BY + CZ = \vec{0}$ where

- B is the $m \times m$ matrix whose columns are the first m columns of A;
- $Y \in F^{m\times 1}$;
- C is the $m \times (n-m)$ matrix whose columns are the last $n-m$ columns of A;
- $Z \in F^{(n-m)\times 1}$.

If $\det(B) \neq 0$ choose $Z \neq \vec{0}$ and $Y = -B^{-1}CZ$. If we let the first m entries of X be those of Y and the last $n-m$ entries of X be Z then $X \neq \vec{0}$ and $AX = 0$. If $\det(B) = 0$, then choose a non-zero vector Y so that $BY = 0$ and let $Z = \vec{0}$.

Theorem 2.4: Suppose that the vector space V is spanned by the vectors $\beta_1, \beta_2, \ldots, \beta_m$. Then any linearly independent set in V has no more than m elements.

Demonstration: Suppose that $n > m$ and $\alpha_1, \ldots, \alpha_n$ are distinct elements of V. Then

$$\alpha_j = \sum_{i=1}^{m} A_{i,j}\beta_i$$

Let A be the matrix with $A(i,j) = A_{i,j}$. There is some $X \neq \vec{0}$ so that $AX = \vec{0}$. Put $x_j = X(j,1)$. Thenn

$$\begin{aligned} \sum_{j=1}^{n} x_j\alpha_j &= \sum_{j=1}^{n} x_j \left(\sum_{i=1}^{m} A_{i,j}\beta_i \right) \\ &= \sum_{i=1}^{m} \left(\sum_{j=1}^{n} A_{i,j}x_j \right) \beta_i \\ &= \sum_{i=1}^{m} (AX)(i,1)\beta_i = \vec{0} \end{aligned}$$

Corollary: If V is a finite dimensional vector space then any two bases of V have the same (finite) number of elements.

Definition: If V is a finite dimensional vector space, its **dimension** is the number of elements in every basis. We will denote the dimension of V by $\dim(V)$.

Corollary: Let V be a finite dimensional vector space with dimension n. Then

1. any subset of V which contains more than n elements is linearly dependent.
2. no subset of V which contains fewer than n elements can span V.

Lemma: Let S be a linearly independent subset of a vector space V. Suppose that $\beta \in V$ is not in the span of S. Then $S \cup \{\beta\}$ is linearly independent. In addition, if V is an inner product space and S is a finite orthonormal set, then there is a vector γ so that $S \cup \{\gamma\}$ is an orthonormal set.

Theorem 2.5: If W is a subspace of a finite dimensional vector space V every linearly independent subset of W is is finite and is part of a finite basis for V. If, in addition, V is an inner product space, then every orthonormal subset of W is part of an orthonormal basis of V.

Theorem 8.3: Let V be an inner product space and let $\beta_1, \ldots \beta_n$ be linearly independent elements of V. Then there are orthonormal vectors $\alpha_1, \ldots, \alpha_n$ that are a basis for the span of $\{\beta_1, \ldots, \beta_n\}$.

Corollary: Every finite dimensional inner product space has an orthonormal basis.

Corollary: If W is a proper subspace of a finite dimensional vector space V then W is finite dimensional and $\dim(W) < \dim(V)$.

Corollary: Every non-empty linearly independent subset of a finite dimensional vector space is part of a basis for that space.

Corollary: Let $A \in F^{n \times n}$. If the row vectors of A or the column vectors of A are linearly independent, then A is invertible. Conversely, if A is invertible, then its rows and columns are linearly independent.

Theorem 2.6: If W_1 and W_2 are finite-dimensional subspaces of V then $W_1 + W_2$ is finite dimensional and

$$\dim(W_1 + W_2) + \dim(W_1 \cap W_2) = \dim(W_1) + \dim(W_2).$$

Definition: If $V = W_1 + W_2$ where the W_i are subspaces of V and $W_1 \cap W_2 = \{\vec{0}\}$ we say that V is the **(interior) direct sum** of W_1 and W_2 and we will write $V = W_1 \oplus W_2$.

As we have seen in the homework if $V = W_1 \oplus W_2$ then for each $\nu \in V$ there are unique vectors $\omega_i \in W_i$ such that $\nu = \omega_1 + \omega_2$.

Definition: If S is a nonempty subset of an inner product space V and $\nu \in V$, we say that $\sigma \in S$ is a **best approximation** to ν by elements of S if

$$\|\nu - \sigma\| \leq \|\nu - \gamma\|$$

for every $\gamma \in S$.

Theorem 8.4: Let W be a subspace of the inner product space V and let $\nu \in V$.

1. The vector $\omega \in W$ is a best approximation to ν by elements of W if and only if $\nu - \omega \in W^\perp$.
2. If a best approximation to ν by elements of W exists, it is unique.
3. If W is finite dimensional with orthonormal basis $\{\omega_1, \ldots, \omega_k\}$ then the best approximation of ν by elements of W is

$$\sum_{j=1}^{k} (\nu|\omega_k)\omega_k$$

Demonstration: Suppose that $\nu - \omega \in W^\perp$ and $\beta \in W$. Then since $\beta - \omega \in W$ we have $(\nu - \omega) \perp (\omega - \beta)$ so the Pythagorean Theorem says

$$\|\nu - \beta\|^2 = \|(\nu - \omega) + (\omega - \beta)\|^2 = \|\nu - \omega\|^2 + \|\omega - \beta\|^2 \geq \|\nu - \omega\|^2$$

as we claimed.

On the other hand, suppose that $\nu - \omega$ is not in $W^\perp$. Then there is some $\alpha \in W$ so that $(\nu - \omega|\alpha) \neq 0$. In particular, $\alpha \neq \vec{0}$. Let

$$\rho = \mathrm{proj}_\alpha(\nu - \omega) = \frac{(\nu - \omega|\alpha)}{(\alpha|\alpha)}\alpha \in W.$$

Remember that for any pair of vectors $\alpha \neq \vec{0}$ and β that $\mathrm{proj}_\alpha(\beta) \perp (\beta - \mathrm{proj}_\alpha(\beta))$!

Note that $\rho \neq \vec{0}$ since $(\nu - \omega|\alpha) \neq 0$ and $\alpha \neq \vec{0}$. Now we have, via the Pythagorean theorem(!), that

$$\|\nu - \omega\|^2 = \|\rho\|^2 + \|\nu - \omega - \rho\|^2 > \|\nu - \omega - \rho\|^2.$$

Since $\omega + \rho \in W$ we have that $\omega + \rho$ is closer to ν than ω is. That means ω is not the best approximation to ν by elements of W. This proves the first assertion.

As for the second assertion, suppose that both ω and α are best approximations to ν in W. We will show $\|\omega - \alpha\|^2 = 0$. Remember that $\nu - \omega$ and $\nu - \alpha$ are both in $W^\perp$, so each is perpendicular to $\omega - \alpha \in W$. So,

$$\|\omega-\alpha\|^2 = (\omega-\alpha|\omega-\alpha) = ((\omega-\nu)+(\nu-\alpha)|\omega-\alpha) = (\omega-\nu|\omega-\alpha)+(\nu-\alpha|\omega-\alpha) = 0+0 = 0.$$

Finally, it is easy to check that

$$\nu - \sum_{j=1}^{k} (\nu|\omega_k)\omega_k$$

is perpendicular to each ω_k, so it is perpendicular to eac element of W. The first part of the theorem tells us it is the best approximation to ν by an element of W.

Bessel's Inequality: Suppose that $\{\sigma_1, \ldots, \sigma_k, \ldots\}$ is an orthonormal subset of an inner product space V and $\nu \in V$. Then for any positive integer N,

$$\sum_{k=1}^{N} |(\nu|\sigma_k)|^2 \leq \|\nu\|^2$$

and there is equality if and only if $\nu \in \mathrm{span}\{\sigma_1, \ldots, \sigma_N\}$. In particular, the infinite sequence

$$x_N = \sum_{k=1}^{N} |(\nu|\sigma_k)|^2$$

is convergent.

Demonstration: Fix N and put $W = \mathrm{span}\{\sigma_1, \ldots, \sigma_N\}$ and

$$\rho = \sum_{k=1}^{N} (\nu|\sigma_k)\sigma_k.$$

The Theorem tells us that $\rho \perp (\nu - \rho)$, so the Pythagorean theorem tells us that

$$\|\nu\|^2 = \|\rho\|^2 + \|\nu - \rho\|^2 \geq \|\rho\|^2 = \sum_{k=1}^{N} |(\nu|\sigma_k)|^2.$$

The only way to have equality here is for $\nu = \rho \in W$.

The last assertion is that an increasing sequence that is bounded above is convergent.

Definition: If a best linear approximation to $\nu \in V$ by elements of W exists, it is called the **orthogonal projection of ν on to W**. If every element of V has an orthogonal projection onto W, then we have a function from V to W and this function is called the **orthogonal projection of V onto W**.

Corollary: If there is an orthogonal projection from V onto a subspace W, it must be unique. If we denote this orthonal projection by E then the function $F(\alpha) = \alpha - E(\alpha)$ is the orthogonal projection of V onto $W^{\perp}$. If W is finite dimensional then there is an orthogonal projection from V onto W.

Theorem 8.5: Let W be a finite dimensional subspace of an inner product space V, and let E be the orthogonal projection of V onto W. The $E \circ E = E$, E is onto W, $E(\omega) = 0$ if $\omega \in W^{\perp}$ and $V = W \oplus W^{\perp}$.

Theorem: Let $A \in F^{n \times n}$, $X \in F^{n \times 1}$, and $Y \in F^{n \times 1}$. Each of the following statements implies all of the others.

1. $\det(A) \neq 0$.
2. A is invertible.
3. $AX = \vec{0}$ implies $X = \vec{0}$.
4. $AX = Y$ has a solution for each Y.
5. There is some $B \in F^{n \times n}$ so that $AB = I$.
6. There is some $C \in F^{n \times n}$ so that $CA = I$.
7. The dimension of the column space of A is n.
8. The dimension of the row space of A is n.
9. If we define $T : F^{n \times 1} \to F^{n \times 1}$ by $T(X) = AX$ then T is one-to-one.
10. If we define $T : F^{n \times 1} \to F^{n \times 1}$ by $T(X) = AX$ then T is onto.