Linear Algebra Notes

Brendan Burkhart

1 Row operations and echelon form

Definition 1.1. A *linear equation* is an equation of the form $a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$

Definition 1.2. A *system of linear equations* (or *linear system*) is a set of linear equations of the same variables (e.g. $x_1, x_2, \ldots, x_n$).

Definition 1.3. The *solution set* of a system of linear equations is the intersection of the solution set of each individual linear equation.

Systems of linear equations can be represented via matrices, where each column is a specific variable, each row is a linear equation, and the entries are the coefficients. Augmentations represent the constant term (denoted b in definition 1.1).

Example 1.1. The linear system

$$\begin{aligned} x + 2y - z &= -1 \\ 2x + 2y + z &= 1 \\ 3x + 5y - 2z &= -1 \end{aligned}$$

can be represented by the augmented matrix below.

$$\left[\begin{array}{ccc|c} 1 & 2 & -1 & -1 \\ 2 & 2 & 1 & 1 \\ 3 & 5 & -2 & -1 \end{array}\right]$$

Row operations are certain operations on matrices or systems of linear equations which can be used to simply or otherwise transform the system, while leaving the solution set unchanged.

Definition 1.4. Row operations:

1. Swap two rows/equations
2. Multiply a single row by a non-zero scalar
3. Add a scalar multiple of one row to another

Theorem 1.1. *If a linear system is obtained from another by one of the row operations, then the two systems have the same set of solutions.*

Proof.

1. As the intersection of sets is commutative, by Definition 1.3, the solution set of a linear system is unaffected by the order of the individual linear systems, therefore swapping rows/equations does not change the solution set.
2. Since multiplying a single row by a non-zero scalar is equivalent to multiplying both sides of the represented equation by that scalar, the solution set is clearly left unchanged.
3. Since adding a multiple of one row to another row is equivalent to adding equal quantities to both sides of the equation, no solutions are removed from the solution set. If the row $(sb_1 + a_1)x_1 + (sb_2 + a_2) + \cdots + (sb_n + a_n)x_n = sk_b + k_a$ is the result of adding s times row $b_1x_1 + \cdots + b_nx_n = k_b$ to row $a_1x_1 + \cdots + a_nx_n = k_a$, then $(sb_1 + a_1)x_1 + (sb_2 + a_2) + \cdots + (sb_n + a_n)x_n = s(b_1x_1 + \cdots + b_nx_n) + k_a$ which implies that $a_1x_1 + \cdots + a_nx_n = k_a$, so no information was lost and the solution set remains the same.

□

Definition 1.5. A matrix in (row) *echelon form* has all rows of zeroes at the bottom, and the leading term of each row is to the right of the row leading term above it.

Definition 1.6. In *reduced row echelon form*, a matrix in row echelon form additionally has leading terms which are all 1, and the entries above and below each leading term are zero.

Remark. The reduced row echelon form of a particular matrix is always unique, however the row echelon form is not. Additionally, it is always possible to transform a given matrix into either echelon form via the row operations.

Example 1.2.

$$\left[\begin{array}{ccc|c} 1 & 0 & 1 & 4 \\ 1 & -1 & 2 & 5 \\ 4 & -1 & 5 & 17 \end{array}\right] \xrightarrow[r_3 \to r_3 - 4r_1]{r_2 \to r_2 - r_1} \left[\begin{array}{ccc|c} 1 & 0 & 1 & 4 \\ 0 & -1 & 1 & 1 \\ 0 & -1 & 1 & 1 \end{array}\right] \xrightarrow{r_2 \to -r_2} \left[\begin{array}{ccc|c} 1 & 0 & 1 & 4 \\ 0 & 1 & -1 & -1 \\ 0 & -1 & 1 & 1 \end{array}\right]$$

$$\xrightarrow{r_3 \to r_3 + r_2} \left[\begin{array}{ccc|c} 1 & 0 & 1 & 4 \\ 0 & 1 & -1 & -1 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Since there is no leading term corresponding to z, z is a free variable. If we let z be zero, then the augmentation yields a particular solution to this linear system.

Solution set in vector form:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 4 \\ -1 \\ 0 \end{pmatrix} + z \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}$$

Solution set in parametric form:

$$\{(4 - z, -1 + z, z) \mid z \in \mathbb{R}\}$$

2 Vector Spaces

Definition 2.1. Let F be a field. A *vector space* over F consists of a set V together with two operations:

$$+ : V \times V \to V \text{ (vector addition)}$$

$$\cdot : F \times V \to V \text{ (scalar multiplication)}$$

which satisfy the following axioms for all $v, w, e \in V$ and $r, s \in F$:

1. $v + w = w + v$ (additive commutativity)
2. $(v + w) + u = v + (w + u)$ (additive associativity)
3. $\exists \vec{0} \in V$ s.t. $v + \vec{0} = v$ (additive identity)
4. $\exists z \in V$ s.t. $v + z = \vec{0}$ (additive inverse)
5. $(r + s) \cdot v = r \cdot v + s \cdot v$ (distributivity over scalar addition)
6. $r \cdot (v + w) = r \cdot v + r \cdot w$ (distributivity over vector addition)

7. $(rs) \cdot v = r \cdot (s \cdot v)$ (multiplicative associativity)

8. $1 \cdot v = v$ (multiplicative identity)

Remark. You can prove the following for all $v \in V$ and $r \in F$ from those axioms (see Lemma 2.1):

- The additive identity ($\vec{0}$) is unique.
- Additive inverses are unique.
- $0 \cdot v = \vec{0}$.
- $-1 \cdot v = -v$.
- $r \cdot \vec{0} = \vec{0}$.

Example 2.1. $\mathbb{R}^n$ is a vector space over $\mathbb{R}$ with $+$ and $\cdot$ defined as usual.

Example 2.2. Let F be any field. Then

$$F^n = \left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \,\middle|\, x_i \in F \text{ for } i = 1, \dots, n \right\}$$

is a vector space over F with $x + y$ defined as

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix}$$

and $r \cdot x$ defined as

$$r \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} rx_1 \\ \vdots \\ rx_n \end{pmatrix}.$$

Consider the field $\mathbb{F}_2 = \{0, 1\}$, and the vector space $\mathbb{F}_2^2$. This vector space has four elements: $\left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\}$. In general, $\mathbb{F}_2^n$ has 2^n elements. Note that since $0 + 0 = 0$ and $1 + 1 = 0$, every element of $\mathbb{F}_2^n$ is its own inverse. Furthermore, every element of *any* vector space V over $\mathbb{F}_2$ is its own inverse. This is because $v + v = 1 \cdot v + 1 \cdot v = (1 + 1) \cdot v = 0 \cdot v = \vec{0}$.

Consider the empty set. It cannot be a vector space, since a vector space requires the existence of an additive identity. However, the set $V = \{\star\}$ with $\star \cdot \star = \star$ and $r \cdot \star = \star$ is a vector space, called the **trivial** vector space. Note that $\star = \vec{0}$.

Example 2.3. $\mathbb{C}$ forms a vector space over $\mathbb{R}$, with element-wise addition and $r \cdot (a + bi) = (ra) + (rb)i$.

Remark. $\mathbb{C} = \mathbb{C}^1$ forms a vector space over $\mathbb{C}$, as does $\mathbb{R} = \mathbb{R}^1$ over $\mathbb{R}$.

Example 2.4. Let V be the set of 2×2 matrices over $\mathbb{R}$, denoted $M_{2\times 2}(R)$:

$$\left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \,\middle|\, a, b, c, d \in \mathbb{R} \right\}.$$

Define vector addition and scalar multiplication as follows.

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} + \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix} = \begin{pmatrix} a + a' & b + b' \\ c + c' & d + d' \end{pmatrix}$$

$$r \cdot \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} ra & rb \\ rc & rd \end{pmatrix}$$

Note that V is essentially equivalent to $\mathbb{R}^4$. V is a vector space over $\mathbb{R}$, and in general $M_{m\times n}(F)$ with entry-wise vector addition and scalar multiplication forms a vector space over F.

Example 2.5. Let F be any field, and $n \in \mathbb{N}$. Then define the set $P_n(F)$, the set of all polynomials with coefficients in F of degree at most n, as follows.

$$P_n(F) = \{a_0 + a_1 x_1 + \cdots + a_n x^n \,|\, a_i \in F \text{ for } i = 0, \ldots, n\}$$

$$(a_0 + \cdots + a_n x^n) + (b_0 + \cdots + b_n x^n) = (a_0 + b_0) + \cdots + (a_n + b_n) x^n$$
$$r(a_0 + \cdots + a_n x^n) = ra_0 + \cdots + ra_n x^n$$

$P_n(F)$ is a vector space over F. Similarly, $P(F)$ (the set of polynomials with coefficients in F of *any* degree) forms a vector space over F.

Example 2.6. Let S be any set, F any field, and $V = \{f : S \to F\}$. V forms a vector space over F with $(f + g)(s) = f(s) + g(s)$ and $(rf)(s) = r(f(s))$ for all $r \in F, s \in S$.

Notice that if $S = \{1, \ldots, n\}$, we get F^n, since $f \in V$ is a function from a coordinate index i to the value of that coordinate x_i. If $S = \mathbb{N}$ and $V = \{(a_1, a_2, \ldots) \,|\, a_i \in F \text{ for } i = 1, 2, \ldots\}$, we similarly get the set of all sequences in F.

Lemma 2.1. *Let V be a vector space over a field F. Then for all $v \in V$ and $r \in F$:*

1. *The additive identity is unique.*
2. *The additive inverse of v is unique.*
3. $0 \cdot v = \vec{0}$.

4. $(-1) \cdot v + v = \vec{0}$.

5. $r \cdot \vec{0} = \vec{0}$.

Proof.

1. Let i_1 and i_2 be additive identities of V. Then $i_1 = i_1 + i_2 = i_2$, and by transitivity $i_1 = i_2$. The additive identity is therefore unique.

2. Let w_1 and w_2 be additive inverses of v. Then $v + w_1 = \vec{0}$ and $v + w_2 = \vec{0}$, so $v + w_1 = v + w_2$. Then $w_1 + (v + w_1) = w_1 + (v + w_2)$, which implies that $(w_1 + v) + w_1 = (w_1 + v) + w_2$. Therefore, $w_1 = w_2$, so the additive inverse of v is unique and so can be denoted $-v$.

3. $0 \cdot v + 0 \cdot v = (0 + 0) \cdot v = 0 \cdot v = 0 \cdot v + \vec{0}$. Since we have additive inverses, we have cancellation, so $0 \cdot v = \vec{0}$.

4. $(-1) \cdot v + v = (-1) \cdot v + 1 \cdot v = (-1 + 1) \cdot v = 0 \cdot v = \vec{0}$.

5. $r \cdot \vec{0} = r \cdot (v + -v) = r \cdot (1 + -1) \cdot v = (r \cdot 0) \cdot v = 0 \cdot v = \vec{0}$.

$\square$

Definition 2.2. Let V be a vector space over F, and $W \subseteq V$. We say W is a *subspace* of V if it is itself a vector space over F with the operations defined as for V.

Example 2.7. Let $W = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \;\middle|\; x + y + z = 0 \right\}$. W is a vector subspace over $\mathbb{R}$ with the operations inherited from $\mathbb{R}^3$. W is a plane within $\mathbb{R}^3$, with normal $(1, 1, 1)$.

Proposition 2.2. *A subset W of V is a subspace if and only if*

1. W is non-empty.

2. W is closed under vector addition.

3. W is closed under scalar multiplication.

Proof.

$(\Longrightarrow)$ If W is a subspace, then by definition it is itself a vector space, so it must contain $\vec{0}$ and so it is non-empty. Additionally, (2) and (3) follow from the definition of binary operations.

$(\Longleftarrow)$ Since W is a subset of V with the same operations, all axioms of vector spaces follow except for the existence of the additive identity and additive inverses, and W being closed under vector addition and scalar multiplication (as these are defined as binary operations for V rather than W).

(2) and (3) then guarantee that W is closed under vector addition and scalar multiplication. Since W is closed under scalar multiplication, we know that $(0 \cdot w) \in W$ for any $w \in W$. Since $0 \cdot w = \vec{0}$, we know that $\vec{0} \in W$ so W has an additive identity. Similarly, since $(-1 \cdot w) \in W$ and $-1 \cdot w = -w$, every element in W must have an additive inverse. $\square$

Example 2.8. Let $V = \mathbb{R}^3$, $W = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \,\middle|\, x + y + z = 1 \right\}$. Then W cannot be a subspace of V, as $\vec{0} \notin W$.

Example 2.9. Let $V = \mathbb{R}^n$, and W be the set of solutions to a homogeneous system of equations (every equation is equal to zero). Then $\{\vec{0}\}$ is a solution to the system of linear equations, and since every solution is equal to zero, the set of solutions is closed under addition and multiplication, so W is a subspace of V.

Example 2.10. If V is any vector space over F, then both $\{\vec{0}\}$ and V are subspaces of V.

Definition 2.3. $\left\{\vec{0}\right\}$ is the *trivial* subspace.

3 Linear Combinations and Spans

Definition 3.1. Let V be a vector space over F, and $S \subseteq V$ with $S \neq \varnothing$. We say $v \in V$ is a *linear combination* of vectors in S if

$$v = c_1 s_1 + \cdots + c_n c_n \text{ for some } n > 1, c_i \in F, s_i \in S \text{for i} = 1, \ldots, n$$

Example 3.1. In $\mathbb{R}^3$, $\begin{pmatrix} 2 \\ 3 \\ 0 \end{pmatrix} = 2 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + 3 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$, so $\begin{pmatrix} 2 \\ 3 \\ 0 \end{pmatrix}$ is a linear combination of vectors in $\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$.

Definition 3.2. Let $S \subseteq V$. The *span* (also *linear span* or *linear hull*) of S is the set of all linear combinations of vectors in S.

$$\operatorname{span}(S) = \{c_1 s_1 + \cdots + c_n s_n \,|\, n \in \mathbb{N}, c_i \in F, s_i \in S\}$$

By convention, if $S = \varnothing$, we define $\operatorname{span}(S) = \vec{0}$.

Remark. If $S \subseteq T$, then $\operatorname{span}(S) \subseteq \operatorname{span}(T)$.

Lemma 3.1. *Let V be a vector space over F, and $s \subseteq V$. Then* $\operatorname{span}(S)$ *is a subspace of V.*

Proof. If $S = \varnothing$, then $\operatorname{span}(S) = \left\{\vec{0}\right\}$, so $\operatorname{span}(S)$ is the trivial subspace.

If $S \neq \varnothing$, then there exists some $s \in S$. Since $(1 \cdot s) \in \operatorname{span}(S)$, we know that $\operatorname{span}(S) \neq \varnothing$. Let $v, w \in \operatorname{span}(S)$, $r \in F$. We need to prove that $v + r \cdot w \in \operatorname{span}(S)$. Since $v \in \operatorname{span}(S)$, we know that $v = c_1 s_1 + \cdots + c_n s_n$ for some $n \geq 1$, $c_i \in F$, and $s_i \in S$. Similarly, $w = d_1 t_1 + \cdots + d_m t_m$ for some $m \geq 1$, $d_i \in F$, and $t_i \in S$. Then $v + rw = c_1 s_1 + \cdots + c_n s_n + r d_1 t_1 + \cdots + r d_m t_m$, so $v + r \cdot w \in \operatorname{span}(S)$. Therefore, by Lemma 2.2, $\operatorname{span}(S)$ is a subspace of V. $\square$

Remark. We also proved that $S \subseteq \operatorname{span}(S)$, and that $\operatorname{span}(S)$ is the smallest subspace of V that contains S, in the sense that if W is a subspace of V such that $S \subseteq W$, then $\operatorname{span}(S) \subseteq W$.

Example 3.2. Let v be a vector space over F. Given $S \subseteq V$, and $v \in V$, how do we know if $v \in \operatorname{span}(S)$?

Let $V = P_2(\mathbb{R})$, $S = \{x^2 + 3x - 2, 2x^2 + 5x - 3\}$, and $v = -x^2 - 4x + 4$. We are trying to determine $a, b \in F$ such that $-x^2 - 4x + 4 = a(x^2 + 3x - 2) + b(2x^2 + 5x - 3)$. This would imply that $a + 2b = -1$, $3a + 5b = -4$, and $-2a - 3b = 4$. This system of equations can be represented as a matrix, and solved for a and b.

$$\left[\begin{array}{rr|r} 1 & 2 & -1 \\ 3 & 5 & -4 \\ -2 & -3 & 4 \end{array}\right] \begin{array}{c} \xrightarrow{r_2 \to r_2 - 3r_1} \\ \xrightarrow{r_3 \to r_3 + 2r_1} \end{array} \left[\begin{array}{rr|r} 1 & 2 & -1 \\ 0 & -1 & -1 \\ 0 & 1 & 2 \end{array}\right]$$

This would imply that $0b = 1$, which is a contradiction, so no such a, b exist.

Lemma 3.2. *Let V be a vector space over F, $S \subseteq V$, and $v \in V$. Then* $\operatorname{span}(S \cup \{v\}) = \operatorname{span}(S)$ *if and only if* $v \in \operatorname{span}(S)$.

Proof.

($\Longrightarrow$) If $v \notin \operatorname{span}(S)$, then since $v \in \operatorname{span}(S \cup \{v\})$, we know that $\operatorname{span}(S) \neq \operatorname{span}(S \cup \{v\})$.

($\Longleftarrow$) Assume $v \in \operatorname{span}(S)$. Then $v = c_1 v_1 + \cdots + c_n s_n$ for $c_i \in F$ and $s_i \in S$. Since $S \subseteq S \cup \{v\}$, we have $\operatorname{span}(S) \subseteq \operatorname{span}(S \cup \{v\})$. Let $w \in \operatorname{span}(S \cup \{v\})$, so $w = d_1 t_1 + \cdots + d_m t_n + d_{m+1} v$. Thus, $w = d_1 t_1 + \cdots + d_m t_m + d_{m+1}(c_1 + \cdots + c_n)$, so $w \in \operatorname{span}(S)$ and $\operatorname{span}(S \cup \{v\}) \subseteq \operatorname{span}(S)$. Therefore, $\operatorname{span}(S) = \operatorname{span}(S \cup \{v\})$. $\square$