Linear Algebra: Vector Spaces

Arjun Vardhan

†
Created: 7th July 2022
Last updated: 27th July 2022

1 Vector Spaces

- A vector space V is a set, along with two operations, vector addition and scalar multiplication, satisfying the following properties:
  1. $v + w = w + v$ for all $v, w \in V$.
  2. $(u + v) + w = u + (v + w)$ for all $u, v, w \in V$.
  3. There exists a zero vector $0 \in V$ such that $v + 0 = v$ for all $v \in V$.
  4. For every $v \in V$ there exists $w \in V$ such that $v + w = 0$. Usually denoted $-v$.
  5. $1 \cdot v = v$ for all $v \in V$.
  6. $(\alpha\beta) \cdot v = \alpha \cdot (\beta \cdot v)$, for all $v \in V$, for all scalars α, β.
  7. $\alpha \cdot (u + v) = \alpha \cdot u + \alpha \cdot v$, for all $u, v \in V$, for all scalars α.
  8. $(\alpha + \beta) \cdot v = \alpha \cdot v + \beta \cdot v$, for all $v \in V$, for all scalars α, β.
- **The zero vector is unique.** *Proof:* Let $0, 0' \in V$ such that $0 + v = v$ and $0' + v = v$ for all $v \in V$. Then $0 = 0 + 0' = 0'$. ■
- **Additive inverses are unique.** *Proof:* Let $v \in V$. Let $u, w \in V$ such that $v + w = v + u = 0$. Then $v + w + w = v + u + w \implies (v + w) + w = (v + w) + u \implies w = u$. ■
- $0 = 0 \cdot v$ **for all** $v \in V$**.** *Proof:* $0 \cdot v = (0 + 0) \cdot v = 0 \cdot v + 0 \cdot v \implies 0 = 0 \cdot v$. ■
- $-v = (-1) \cdot v$ **for all** $v \in V$**.** *Proof:* $(-1) \cdot v + v = (-1) \cdot v + 1 \cdot v = (-1 + 1) \cdot v = 0$. Thus $(-1) \cdot v = -v$. ■
- The scalars are always from a field, usually $\mathbb{R}$ or $\mathbb{C}$.
- An $m \times n$ matrix is a rectangular array with m rows and n columns. Entries of a matrix are denoted a_{ij} or $(A)_{i,j}$, where i is the row and j is the column.
- Given a matrix A, its transpose A^T is the matrix formed by transforming the rows of A into columns. Formally, $(A)_{i,j} = (A^T)_{j,i}$.

2 Linear Combinations and Bases

- Let V be a vector space, and let $v_1, v_2, ..., v_p \in V$ be a collection of vectors. A linear combination of $v_1, v_2, ..., v_p$ is a sum of the form $\alpha_1 v_1 + \alpha_2 v_2 + ... + \alpha_p v_p = \sum_{k=1}^{p} \alpha_k v_k$.
- A system of vectors $v_1, v_2, ..., v_p \in V$ is called a basis for V if any vector $v \in V$ admits a unique representation as a linear combination of $v_1, v_2, ..., v_p$, i.e, $v = \alpha_1 v_1 + \alpha_2 v_2 + ... + \alpha_p v_p = \sum_{k=1}^{p} \alpha_k v_k$. The scalars $\alpha_1, \alpha_2, ..., \alpha_p$ are called the coordinates of v with respect to this basis.

- Consider the vector space $\mathbb{F}^n$, where $\mathbb{F}$ is either $\mathbb{R}$ or $\mathbb{C}$. Let $e_1, e_2, ..., e_n \in \mathbb{F}^n$, where e_k is the vector whose entries are all 0 except the kth entry, which is 1. Clearly, any vector in $\mathbb{F}^n$ can be expressed uniquely as a linear combination of $e_1, e_2, ..., e_n$. This system is called the standard basis in $\mathbb{F}^n$.

- A system of vectors $v_1, v_2, ..., v_p \in V$ is called a generating system or spanning system if any vector in V can be represented as a linear combination of them.

- A linear combination $\alpha_1 v_1 + \alpha_2 v_2 + ... + \alpha_p v_p$ is called trivial if $a_k = 0$ for all k.

- A system of vectors $v_1, v_2, ..., v_p \in V$ is called linearly independent if the trivial linear combination is the only linear combination of $v_1, v_2, ..., v_p$ that equals 0.

- If a system is not linearly independent, it is said to be linearly dependent. I.e, if there exists a nontrivial linear combination of $v_1, v_2, ..., v_p$ that equals 0.

- **A system of vectors $v_1, v_2, ..., v_p \in V$ is linearly dependent if and only if one of the vectors v_k can be represented as a linear combination of the others.** *Proof:* Suppose the system $v_1, v_2, ..., v_p$ is linearly dependent. Then there exist scalars $\alpha_1, \alpha_2, ..., \alpha_p$, such that $\sum_{m=1}^{p} \alpha_m v_m = 0$ with $\sum_{m=1}^{p} |\alpha_m| \neq 0$. Let k be an index such that $\alpha_k \neq 0$. Then $\alpha_k v_k = -\sum_{m=1, m\neq k}^{p} \alpha_m v_m \implies v_k = \frac{1}{\alpha_k} \sum_{m=1, m\neq k}^{p} \alpha_m v_m$. Conversely, suppose $v_k = \sum_{m=1, m\neq k}^{p} \beta_m v_m$. Then $0 = v_k - \sum_{m=1, m\neq k}^{p} \beta_m v_m$, which is a nontrivial linear combination. ■

- **A basis is linearly independent.** *Proof:* The trivial linear combination is equal to 0, but as each representation is unique, that is the only linear combination of the basis elements that equals 0. ■

- Thus every basis is linearly independent and a generating system.

- **If a system of vectors $v_1, v_2, ..., v_p$ is a linearly independent generating system, then it is a basis.** *Proof:* Let $v \in V$. Then $v = \alpha_1 v_1 + \alpha_2 v_2 + ... + \alpha_p v_p$ for some scalars $\alpha_1, \alpha_2, ..., \alpha_p$. Suppose v has another representation as a linear combination of $v_1, v_2, ..., v_p$, i.e, $v = a_1 v_1 + a_2 v_2 + ... + a_p v_p$. Then $0 = v - v = (\alpha_1 - a_1)v_1 + (\alpha_2 - a_2)v_2 + ... + (\alpha_p - a_p)v_p$. As the system is linearly independent, this is the trivial linear combination and so $\alpha_k - a_k = 0$ for all k. Thus the representation is unique and $v_1, v_2, ..., v_p$ constitute a basis for V. ■

- **Any finite generating system contains a basis.** *Proof:* Suppose $v_1, v_2, ..., v_p \in V$ is a generating system. If it is linearly independent, we are done, so suppose that it isn't. Then there exists a vector v_k that can be represented as a linear combination of the others. Therefore any linear combination of $v_1, v_2, ..., v_p$ can be represented as a linear combination of those vectors without v_k. So we can delete v_k from the system and it will still be a generating system. If this new system is linearly independent, we are done. Otherwise this procedure will be repeated until we obtain a linearly independent system. The process must terminate because otherwise we would end up with an empty set. ■

3 Linear Transformations

- Let V, W be vector spaces over the same field $\mathbb{F}$. A function $T : V \to W$ is called a linear transformation if $T(u+v) = T(u) + T(v)$, $T(\alpha \cdot v) = \alpha \cdot T(v)$ for all $u, v \in V$, all $\alpha \in \mathbb{F}$.

- Let $T : \mathbb{F}^n \to \mathbb{F}^m$ be a linear transformation. Let $e_1, e_2, ..., e_n$ be the standard basis in $\mathbb{F}^n$. Let $a_1, a_2, ...a_n \in \mathbb{F}^m$ such that $a_k = T(e_k)$ for all k. Let $x \in \mathbb{F}^n$, $x = (x_1, x_2, ..., x_n)$. Then $x = x_1 e_1 + x_2 e_2 + ... + x_n e_n$. Thus $T(x) = T(x_1 e_1 + x_2 e_2 + ... x_n e_n) = x_1 a_1 + x_2 a_2 + ... + x_n a_n$. Therefore knowing how a linear transformation T acts on the standard basis is enough to calculate $T(x)$ for all $x \in V$.

- If we join the vectors $a_1, a_2, ..., a_n$ left to right, we obtain the matrix representation of the linear transformation, generally denoted $[T]$. If we multiply this matrix with a vector $x \in V$, we obtain $T(x)$.
- In fact it is not necessary to consider the standard basis. The action of a linear transformation on any generating set is enough to describe it completely.
- **Let $T : V \to W$ be a linear transformation. Then $T(0) = 0$.** *Proof:* $T(0) = T(0+0) = T(0) + T(0) \implies T(0) = 0$. ■

4 Vector Space of Linear Transformations

- Let $T_1, T_2 : V \to W$ be linear transformations and let α be a scalar. Define $T_1 + T_2(v) = T_1(v) + T_2(v)$ and $\alpha T(v) = \alpha \cdot (T(v))$. It is easy to check that $T_1 + T_2$ and αT are linear transformations themselves.
- For fixed vector spaces V and W, let the set of all linear transformations from V to W be denoted by $L(V, W)$. It can easily be verified that $L(V, W)$ is a vector space with respect to the operations defined above.

5 Matrix Multiplication

- The product of two matrices A and B is defined only if A is an $m \times n$ matrix and B is an $n \times r$ matrix.
- Matrix multiplication is defined as such: $(AB)_{i,j} = (A)_i \cdot (B)_j$. That is, $(AB)_{i,j}$ is the dot product of the ith row of A and the jth column of B.
- Let A, B, C be matrices, and suppose that the following products are defined. Then:
  1. $A(BC) = (AB)C$.
  2. $A(B + C) = AB + AC$.
  3. $A(\alpha B) = (\alpha A)B = \alpha(AB)$.

  Matrix multiplication is not, in general, commutative.
- **$(AB)^T = B^T A^T$.** *Proof:* Let $(AB)^T = C$, and $B^T A^T = D$. Let A be $m \times n$, and let B be $n \times r$. Then C is $r \times m$ and D is $r \times m$, so their dimensions match. $(C)_{i,j} = (AB)_{j,i} = \sum_{k=1}^{n} (A)_{j,k}(B)_{k,i}$. Now $(D)_{i,j}$ is the dot product of the ith row of B^T with the jth column of A^T, which is nothing but the ith column of B and the jth row of A. So $(D)_{i,j} = \sum_{k=1}^{n} (A)_{j,k}(B)_{k,i} = (C)_{i,j}$. Thus $C = D$. ■
- For a square $(n \times n)$ matrix A, its trace, denoted trace A, is the sum of all its diagonal entries, i.e, trace $A = \sum_{k=1}^{n} (A)_{k,k}$.
- **Let A and B be matrices of size $m \times n$ and $n \times m$ respectively. Then, trace $(AB) =$ trace (BA).** *Proof:* trace $(AB) = \sum_{k=1}^{n} (A)_{1,k}(B)_{k,1} + \sum_{k=1}^{n} (A)_{2,k}(B)_{k,2} + ... + \sum_{k=1}^{n} (A)_{m,k}(B)_{k,m}$. trace $(BA) = \sum_{k=1}^{m} (B)_{1,k}(A)_{k,1} + \sum_{k=1}^{m} (B)_{2,k}(A)_{k,2} + ... + \sum_{k=1}^{m} (B)_{n,k}(A)_{k,n} = \sum_{k=2}^{m} (B)_{1,k}(A)_{k,1} + \sum_{k=2}^{m} (B)_{2,k}(A)_{k,2} + ... + \sum_{k=2}^{m} (B)_{n,k}(A)_{k,n} + \sum_{k=1}^{n} (A)_{1,k}(B)_{k,1} = \sum_{k=3}^{m} (B)_{1,k}(A)_{k,1} + \sum_{k=3}^{m} (B)_{2,k}(A)_{k,2} +$

$... + \sum_{k=3}^{m}(B)_{n,k}(A)_{k,n} + \sum_{k=1}^{n}(A)_{1,k}(B)_{k,1} + \sum_{k=1}^{n}(A)_{2,k}(B)_{k,2}$. Thus we can conclude that trace $(BA) = \sum_{k=1}^{n}(A)_{1,k}(B)_{k,1} + \sum_{k=1}^{n}(A)_{2,k}(B)_{k,2} + ... + \sum_{k=1}^{n}(A)_{m,k}(B)_{k,m} =$ trace (AB). ■

6 Invertible Transformations and Isomorphisms

- Let V be a vector space. Then $I : V \to V$; $I(v) = v$ for all $v \in V$ is called the identity transformation.

- If $I : \mathbb{F}^n \to \mathbb{F}^n$ is the identity transformation in $\mathbb{F}^n$, its matrix is the identity matrix of size n, denoted I_n ($n \times n$ matrix with 1 in the main diagonal entries and 0 everywhere else).

- Let $T : V \to W$ be a linear transformation. We say that T is left invertible if there exists $A : W \to V$ such that $A(T(v)) = v$ for all $v \in V$. T is right invertible if there exists $B : W \to V$ such that $T(B(v)) = v$ for all $v \in V$. I.e, $A \circ T = I_V$ and $T \circ B = I_W$, where I is the identity transformation in V and W respectively. If T is both left and right invertible, it is said to be invertible.

- **If $T : V \to W$ is invertible, then its left and right inverses A and B are the same and unique.** *Proof:* $AT = I = TB$. Then $ATB = (AT)B = IB = B$, but $ATB = A(TB) = AI = A$. Thus $B = A$. As for uniqueness, suppose there exists A_1 such that $A_1T = I$. Then repeating the above reasoning we get $A_1 = B = A$. ■

- **Corollary: A linear transformation $T : V \to W$ is invertible if and only if there exists a unique linear transformation $T^{-1} : W \to V$ such that $T \circ T^{-1} = I_W$ and $T^{-1} \circ T = I_V$.**

- A matrix is invertible if its corresponding linear transformation is invertible.

- An invertible matrix must always be square ($n \times n$).

- **If linear transformations A and B are invertible, then if the product AB is defined, it is invertible and $(AB)^{-1} = B^{-1}A^{-1}$.** *Proof:* $(AB)(B^{-1}A^{-1}) = AIA^{-1} = I$. And $(B^{-1}A^{-1})AB = B^{-1}IB = I$. ■

- **If a matrix A is inveritible, then A^T is also invertible and $(A^T)^{-1} = (A^{-1})^T$.** *Proof:* Since $(AB)^T = B^TA^T$, it follows that $(A^{-1})^TA^T = (AA^{-1})^T = I^T = I$. Similarly, $A^T(A^{-1})^T = (A^{-1}A)^T = I$. ■

- An invertible linear transformation is called an isomorphism. Two vector spaces V and W are isomorphic if an isomorphism between them exists, denoted $V \cong W$.

- **Let $T : V \to W$ be an isomorphism, and let $v_1, v_2, ..., v_n$ be linearly independent in V. Then the system $T(v_1), T(v_2), ..., T(v_n)$ is linearly independent in W.** *Proof:* Suppose there exist scalars $\alpha_1, \alpha_2, ..., \alpha_n$, not all 0, such that $\alpha_1T(v_1) + \alpha_2T(v_2) + ... + \alpha_nT(v_n) = 0$. Then $T(\alpha_1v_1 + \alpha_2v_2 + ... + \alpha_nv_n) = 0$. As T is an isomorphism, $T(0) = 0$, and thus $\alpha_1v_1 + \alpha_2v_2 + ... + \alpha_nv_n = 0$, which is a contradiction. So $T(v_1), T(v_2), ..., T(v_n)$ are linearly independent. ■

- **Let $T : V \to W$ be an isomorphism, and let $v_1, v_2, ..., v_n$ be a generating system in V. Then $T(v_1), T(v_2), ..., T(v_n)$ is a generating system in W.** *Proof:* Let $w \in W$. As T is an isomorphism, there exists $v \in V$ such that $T(v) = w$. There exist scalars $\alpha_1, \alpha_2, ..., \alpha_n$ such that $\alpha_1v_1 + \alpha_2v_2 + ... + \alpha_nv_n = v$. Thus $T(\alpha_1v_1 + \alpha_2v_2 + ... + \alpha_nv_n) = T(v) \implies \alpha_1T(v_1) + \alpha_2T(v_2) + ... + \alpha_nT(v_n) = w$. Therefore $T(v_1), T(v_2), ..., T(v_n)$ is a generating system. ■

- **Corollary: Let $T : V \to W$ be an isomorphism, and let $v_1, v_2, ..., v_n$ be a basis in V. Then $T(v_1), T(v_2), ..., T(v_n)$ is a basis in W.**

- **Let $T : V \to W$ be a linear transformation, and let $v_1, v_2, ..., v_n$ and $w_1, w_2, ..., w_n$ be bases in V and W respectively. If $T(v_k) = w_k$ for all k, then T is an isomorphism.** *Proof:* Let $T^{-1} : W \to V$, $T^{-1}(w_k) = v_k$ for all k. So T is invertible and thus an isomorphism. ■

- **Let $A : X \to Y$ be a linear transformation. Then A is invertible if and only if for all $b \in Y$, the equation $Ax = b$ has a unique solution in X.** *Proof:* Suppose A is invertible. Then $x = A^{-1}b$ is a solution for the equation. Suppose that for some $x_1 \in X$, $Ax_1 = b$. Then $A^{-1}Ax_1 = A^{-1}b \implies x_1 = A^{-1}b = x$. Conversely, suppose the equation has a unique solution in X for all $b \in Y$. Let $T(b)$ be this unique solution for each $b \in Y$. Let $b_1, b_2 \in Y$ and let α_1, α_2 be scalars. Let $x_1 = T(b_1)$ and let $x_2 = T(b_2)$. Then $A(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_1 Ax_1 + \alpha_2 Ax_2 = \alpha_1 b_1 + \alpha_2 b_2$, and thus $T(\alpha_1 b_1 + \alpha_2 b_2) = \alpha_1 T(b_1) + \alpha_2 T(b_2)$. So T is a linear transformation. Now let $x \in X$ and let $y = Ax$. Then $TAx = Ty = x$, so $TA = I$. Similarly let $y \in Y$ and let $x = Ty$. Then $ATy = Ax = y$, so $AT = I$. Therefore $T = A^{-1}$. $\blacksquare$

- **Corollary: An $n \times n$ matrix is invertible if and only if its columns form a basis in $\mathbb{F}^n$.**

7 Subspaces

- A subspace of a vector space is a nonempty subset $V_0 \subseteq V$ such that $u, v \in V_0 \implies v + u \in V_0$, $v \in V_0 \implies \alpha \cdot v \in V_0$ for all scalars α. A subspace is a vector space itself, with the same operations as the parent space.

- Two trivial subspaces of any vector space V are V itself, and $\{0\}$, the set containing only the zero vector.

- **Let U be a subspace of V. Then $0 \in U$.** *Proof:* As U is nonempty, let $u \in U$. Then $-1 \cdot u = -u \in U \implies u + (-u) = 0 \in U$. $\blacksquare$

- Every linear transformation $T : V \to W$ has two associated subspaces:

  1. The null space, or kernel of T, denoted $\text{Ker}(T) = \{v \in V : T(v) = 0\}$.
  2. The range of T, denoted $\text{Ran}(T) = \{w \in W : \text{there exists } v \in V,\ T(v) = w\}$.

- **Let $T : V \to W$ be a linear transformation. Then $\text{Ker}(T)$ is a subspace of V.** *Proof:* $0 \in \text{Ker}(T)$, so it is nonempty. Suppose $u, v \in \text{Ker}(T)$. Then $T(u) = T(v) = 0 \implies T(u) + T(v) = T(u + v) = 0 \implies u + v \in \text{Ker}(T)$. Let α be a scalar. Then $T(\alpha \cdot v) = \alpha \cdot T(v) = \alpha \cdot 0 = 0$. Thus $\alpha \cdot v \in \text{Ker}(T)$. $\blacksquare$

- **Let $T : V \to W$ be a linear transformation. Then $\text{Ran}(T)$ is a subspace of W.** *Proof:* $0 \in \text{Ran}(T)$, so it is nonempty. Suppose $u, v \in \text{Ran}(T)$. Then there exist $a, b \in V$ such that $T(a) = u$ and $T(b) = v$. As $u + v = T(a) + T(b) = T(a + b)$, we have $u + v \in \text{Ran}(T)$. Also $\alpha \cdot v = \alpha \cdot T(b) = T(\alpha \cdot b)$, so $\alpha \cdot v \in \text{Ran}(T)$. $\blacksquare$

- Given a system of vectors $v_1, v_2, ..., v_n \in V$, their linear span, denoted $\text{span}(v_1, v_2, ..., v_n)$, is the set of all linear combinations of $v_1, v_2, ..., v_n$.

- **Let $v_1, v_2, ..., v_n \in V$. Then $\text{span}(v_1, v_2, ..., v_n)$ is a subspace of V.** *Proof:* The trivial linear combination, i.e $0 \in \text{span}(v_1, v_2, ..., v_n)$, so it is nonempty. Let $u, w \in \text{span}(v_1, v_2, ..., v_n)$. Then $u = \alpha_1 v_1 + ... + \alpha_n v_n$ and $w = \beta_1 v_1 + ... + \beta_n v_n$. So $(\alpha_1 + \beta_1)v_1 + ... + (\alpha_n + \beta_n)v_n = u + w \in \text{span}(v_1, v_2, ..., v_n)$. Let γ be a scalar. Then $(\gamma\alpha_1)v_1 + ... + (\gamma\alpha_n)v_n = \gamma \cdot u \in \text{span}(v_1, v_2, ..., v_n)$. $\blacksquare$

- **Let X and Y be subspaces of V. Then $X \cap Y$ is a subspace of V.** *Proof:* As $0 \in X$ and $0 \in Y$, $0 \in X \cap Y$ and thus it is nonempty. Let $u, v \in X \cap Y$. As X and Y are both subspaces themselves, $u + v$ and $\alpha \cdot v$ belong to both of them. So $u + v, \alpha \cdot v \in X \cap Y$. $\blacksquare$

- Let $X, Y \subset V$. Then $X + Y = \{v \in V : v = x + y, x \in X, y \in Y\}$. That is, $X + Y$ is the set of all vectors in V that can be expressed as the sum of a vector from X and a vector from Y.

- **$X + Y$ is a subspace of V if X and Y are subspaces themselves.** *Proof:* $0 = 0 + 0$ so $0 \in X + Y$. Suppose $u, v \in X + Y$. Then $u = x_1 + y_1$ and $v = x_2 + y_2$ where $x_1, x_2 \in X$ and $y_1, y_2 \in Y$. Then $(x_1 + x_2) + (y_1 + y_2) = u + v \in X + Y$. Also, $\alpha \cdot x_1 + \alpha \cdot y_1 = \alpha \cdot u \in X + Y$. $\blacksquare$