Matrix Algebra

Daniel N. Hauser

This note summarizes some of the background concepts relating to vector spaces and matricies.

Vectors and Vector Spaces

In this course, we mostly work in the vector space $\mathbb{R}^n$. That is, ordered n-tuples of real numbers $(v_1, v_2, \ldots v_n)$ along with two operations, addition and scalar multiplication. We can add two vectors by adding the components together, and we can "scale" a vector by multiplying it by a real number (called a scalar). We can think about vectors as a way to express length and direction, which is why in some classes you may have seen them drawn as arrows that start at the origin and end at the n-tuple.

A few important concepts

Definition. *A* linear subspace $S \subseteq \mathbb{R}^n$ *is a set such that*

1. $u, v \in S \Rightarrow u + v \in S$
2. $v \in S \Rightarrow av \in S$ *for every scalar* $a \in \mathbb{R}$

A subspace S is a set we can't "escape" through applications of our vector space operations. So this is a vector space that lives "inside" of $\mathbb{R}^n$ It is also convenient to think about when vectors can be represented using other vectors

Definition. *We say a set of non-zero vectors* $\{v^1, v^2 \ldots v^k\}$ *are* linearly dependent *if there exist scalars* $\{a_1, a_2, \ldots a_k\}$*, at least one non-zero such that*

$$\sum_{j=1}^{k} a_j v^j = 0$$

Otherwise they are said to be linearly independent

Any set of linearly independent vectors essentially describes a linear subspace (the set of vectors we can make using addition and scalar multiplication) and any linear subspace can be described in terms of a collection of linearly independent vectors.[1] Given any

[1] Any collection of more than n vectors must be linearly dependent. It is a system of n linear equations with $k > n$ unknowns.

collection of vectors $\{v^1, v^2, \ldots v^k\}$, this subspace they describe $S = \{u : u = \sum_{i=1}^{k} \alpha_i v^i,\ \alpha_i \in \mathbb{R}\}$ is called the *span*.

We often talk about the dimension of a space. Intuitively $\mathbb{R}^n$ has dimension n. This captures the number of "independent" directions we can move in in $\mathbb{R}^n$ (e.g. vertically and horizontally in $\mathbb{R}^2$). We can define a similar concept for an arbitrary linear space S.

Definition. *A vector space $S \subseteq \mathbb{R}^n$ has* dimension *k if there exists a collection of linearly independent vectors $V = \{v^1, v^2, \ldots v^k\}$ such that $S = span(V)$. V is then a* basis *of S.*

Every linear space has a basis and every basis has the same number of elements.[2]

A vector operation you probably remember is the dot product, or inner product (written as either $u \cdot v$ or $\langle u, v \rangle$). This is given by $u \cdot v = \sum_{i=1}^{n} u_i v_i$.[3] We can use this to compactly write linear equations

$$v_1 x_1 + v_2 x_2 + \ldots v_n x_n = c$$

as $v \cdot x = c$. If we have a system of multiple equations $v^j \cdot x = c_j$ we can write this as a matrix $Vx = c$, where x and c are column vectors and V is a matrix with rows equal to the v^j's.[4]

[2] These things both seem obvious, but turn out to require a bit of work to prove. If you want some practice with the concepts, proving these yourself may be a good idea.

[3] Geometrically $\mathbb{R}^2$, it's easy to see that $u \cdot v > 0$ if the angle between any two vectors is acute, $u \cdot v < 0$ if it's obtuse and $u \cdot v = 0$ if they are orthogonal. This gives us a natural notion of these concepts for $\mathbb{R}^n$, $n > 2$.

[4] The product of any two matricies AB is the matrix C where $(C)_{ij}$ is the dot product of the ith row of A with the jth column of B. Somewhat annoyingly, we can't multiply every matrix, this only makes sense if $A \in \mathbb{R}^{m \times n}$ and $B \in \mathbb{R}^{n \times k}$, which yields a matrix $C \in \mathbb{R}^{m \times k}$. Moreover, $AB \neq BA$, even if the dimensions are such that AB and BA are both well defined.

Matricies

Consider the equation

$$Vx = c.$$

You probably know some ways to solve this for a specific V and c.[5] The more important questions for our purposes are (i) does this have a solution, (ii) if so, how many. The theory from the previous section gives us an equivalent formulation for (i), does c lie in the subspace given by the span of the columns of V. More subtly, (ii) can be formulated as the question, what is the dimension of the subspace given by vectors that solve $Vx = 0$?[6] The former space is called the column space, the latter is called the kernel or null space. We can define the row space similarly. These are all vector spaces. Note that every vector in the kernel is orthogonal to every vector in the row space.

[5] Some you may remember.

1. Gaussian elimination
2. LU, QR, and/or Cholesky decomposition
3. Cramer's rule
4. linsolve(V,c)

[6] If x solves $Vx = c$ and y solves $Vy = 0$ then $V(x + y) = c$. Similarly if $Vx = c$ and $Vy = c$ then $V(x - y) = 0$. So a system of equations has at most one solution if and only if the kernel is 0 vector. Otherwise it has infinitely many solutions.

Theorem 1. *Given matrix $V \in \mathbb{R}^{m \times n}$, the dimension of the row space (called the* rank *of the matrix) and the dimension of the kernel add up to n.*

The rank of a matrix is thus the number of linearly independent rows. Equivalently, it can be show to be the number of linearly independent columns. A $m \times n$-matrix is said to be full rank if it has rank $\min\{m, n\}$, i.e. if it has the maximal possible rank. The previous theorem almost directly implies the "Fundamental Theorem of Linear Algebra"

Theorem 2. *Let U be the set of solutions to $V \cdot x = c$ where $V \in \mathbb{R}^{m \times n}$, $c \in \mathbb{R}^m$. If $U \neq \emptyset$ then U is then an affine subspace*[7] *and*

$$dim(U) = n - rank(V).$$

This, together with the observation that a solution exists iff c is the column space we can answer (i) and (ii) about $Vx = c$.

1. If $rank(V) = n$ then the system of equations has exactly one solution if the augmented matrix $[V|c]$ (V with the additional column c) satisfies $rank(V) = rank([V|c])$.[8]

2. If $rank(V) < n$ then this has infinitely many solutions if the augmented matrix $[V|c]$ satisfies $rank(V) = rank([V|c])$.

3. Otherwise it has no solution.

A special case is when $V \in \mathbb{R}^{n \times n}$ and has full rank. This is called a non-singular matrix and we can find a matrix V^{-1} where the ith column solves

$$Vx = e^i$$

where e^i is the vector that is 1 in the *ith* component and 0 in all others. This is called the inverse matrix as

$$VV^{-1} = V^{-1}V = I$$

We can then easily solve any system of linear equations as $x = V^{-1}c$ is the unique solution to $Vx = c$ by properties of matrix multiplication.

Finally, we can talk about linear functions, i.e. functions $f : \mathbb{R}^n \to \mathbb{R}^m$ where $f(ax + by) = af(x) + bf(y)$ for any scalars a and b and

[7] An affine subspace is essentially a vector space that doesn't have to include the origin, i.e. U is affine if there exists a vector $w \in \mathbb{R}^n$ and a linear subspace $S \subseteq \mathbb{R}^n$ s.t. $U = w + S$. We can define dimension in the obvious way.

[8] This augmented matrix thing is a fancier way to express the c living the column space condition. If $rank(V) = m$, then it holds for any c. Otherwise it does not hold for some cs.

any $x, y \in \mathbb{R}^n$. Any such function can be expressed as $f(x) = Ax$ for some matrix $A \in \mathbb{R}^{m \times n}$. The above observation gives us that if $rank(A) = m$ then this is injective, if $rank(A) = n$ then it is surjective.