

# Chapter 1

## Vector Spaces

### 1.1 The Fundamental theorem of Subspaces

Let  $U$  be a subspace in  $\mathbb{R}^n$  with a spanning set of  $\vec{x}_1 \dots \vec{x}_m$ . If  $\vec{y}_1 \dots \vec{y}_k$  is linearly independent,  $k \leq m$

The proof of this is built on the replacement lemma: If we are given a set of vectors  $\vec{U}_1 \rightarrow \vec{U}_l$  and  $v = a_1 \vec{U}_1 \dots a_l \vec{U}_l$  with  $a_1 \neq 0$ , then the span of  $u_1 \rightarrow u_l$  is equal to the span of  $V, U_2 \dots U_l$ . We can collect terms and factor out to prove this.

Definition of Equality of  $U$  and  $W$ : For every  $\vec{x}$ ,  $\vec{x}$  is in subspace  $U$  iff  $\vec{x}$  is in  $W$ , ie.  $U \subset W$ . Also, for  $x$  to be in a set, it must fulfil the properties of the set.

#### 1.1.1 Bases and Dimension

A **Basis** is a set that first spans a given subspace and second is linearly independent.

The **Invariance Thm** suggests that if there are two bases of  $U$ , then they have the same number of vectors.

The **Dimension** of any subspace except  $\vec{0}$  is the number of vectors in any of its bases.

### 1.2 Orthogonalization

(This is continued as of November 14th's lecture) Matrix Template

$$\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$$

**Normalizing** an orthogonal set  $x_1 \dots x_k$  means to use  $\frac{x_1}{|x_1|} \dots \frac{x_k}{|x_k|}$

**Pythagoreas in  $\mathbb{R}^n$**  if  $\vec{x}_1 \dots \vec{x}_n$  is orthogonal, then  $|\vec{x}_1 + \vec{x}_k|^2 = |\vec{x}_1|^2 \dots + |\vec{x}_k|^2$  (Prove this for homework, as a dot product against itself.)

#### 1.2.1 The Expansion Theorem

If  $x_1$  to  $x_k$  is an orthogonal basis of a subspace  $U$  then for any  $\vec{x}$  in the subspace, we can write  $\vec{x}$  is equal to the

$$\vec{x} = \sum_{i=1}^k \frac{\vec{f}_i \cdot \vec{x}}{|\vec{f}_i|^2}$$

is

$$\sum_{i=1}^k \text{proj}_{\vec{f}_i} \vec{x}$$

Proof:  $\vec{x} = \sum_i a_i \vec{f}_i$

For any  $j$ ,  $\vec{x} \cdot \vec{f}_j = \sum_i a_i \vec{f}_i \cdot \vec{f}_j$ , which will give  $a_j(\vec{f}_j \cdot \vec{f}_j)$  which is the magnitude squared. We can

use this for all j.

$$\left\{ \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ -1 \end{bmatrix} \right\} \text{ is orthogonal.}$$

Write  $\begin{bmatrix} 5 \\ -1 \\ 3 \end{bmatrix}$  as a linear combination.

Prior to this, we had to assign parameters and solve, but now we can project  $\begin{bmatrix} 5 \\ -1 \\ 3 \end{bmatrix}$  onto each of our three orthogonal vectors, and sum it. The fractional

$$\text{term from the projection } \frac{\begin{bmatrix} 5 \\ -1 \\ 3 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}}{\left| \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} \right|^2} \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} + \dots$$

plus the rest of the orthogonal set.

$$= \frac{1}{3} \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} + \dots$$

## 1.3 5.4 - newish Subspaces

Given an  $m \times n$  matrix A

- $\text{row}(A)$  = span of the rows of A. This is its own object.
- $\text{col}(A)$  = span of the cols of A. = the set of everything of the form where c is col  $a_1 \vec{c}_1 + \dots + a_n \vec{c}_n$  —  $a_1 \rightarrow a_n$  are real  
=  $\{A\vec{x} | \vec{x} \in \mathbb{R}^n\}$  which is actually the image of A!

If A is equiv to B by any number of elementary row operations, their row spaces are also equivalent. ( $\text{row}(A) = \text{row}(B)$ ). The proof is done in three parts, and we need to prove that none of the three row ops change anything.

$$\text{Let } A = \begin{bmatrix} r_1 \\ r_2 \\ \dots \\ r_n \end{bmatrix}, \text{row}(A) = \text{span}(\vec{r}_1 + \vec{r}_2 + \dots \vec{r}_m)$$

- Rearranging them just changes the order that you get somewhere
- Multiplying by an arbitrary scalar can be swallowed by the coefficient from span.
- Addition of a scalar multiple of anything can be factored out and summed somewhere else.
- The proof of this is left as an exercise to the reader.

$\therefore |A \sim R$  with R in row echelon form, then  $\text{row}(A) = \text{row}(R)$ .

From yesterday, the nonzero rows of R are independent, therefore they form a basis for the row space of R, therefore they form a basis for the row space of A, so their number is constant.

Therefore, the number of nonzero rows in R, which is equal to the number of nonleading variables is a constant. This is called the **rank** of the matrix. So, the rank exists!

So the number of parameters in  $A\vec{x} = \vec{0}$  is constant, so  $n - \text{rank}(A) = \dim(\text{null}(A))$ .

### 1.3.1 Useful Algorithms

Given vectors  $\{\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n\}$ , find a basis for the vector space they span.

#### Method I

write each vector  $\vec{x}_i$  as a row in a, and reduce it to REF. Then, take the nonzero rows as your basis. This yields a basis easily, but the basis will not be a subset of your initial vectors. However, you may take the rows in A corresponding to the rows in R that are nonzero, given that you have made no swaps or reversed them in A while doing them in R.

#### Method II

Write the vectors as columns of A, and reduce to REF. The basis is the collection of columns in A, that correspond to the columns with leading terms (ones) in R.

To get a basis of the null space, write your vectors as columns in a matrix, and row reduce. Then, write your solutions as parameters. The basic solutions are the basis of the null space.

### 1.3.2 Dimension Theorem, Rank + Nullity

If  $A$  is  $m \times n$  with rank  $r$ , the dimension of the null space of  $A$  plus the dimension of the image space of  $A$  is equal to  $n$ , the number of columns. This is because the image space corresponds to the leading variables and null to the nonleading variables, and each column must be one or another.

This lets us know the possible dimensions of both spaces, given just the size of the matrix. For example, for a  $3 \times 5$   $A$ ,  $\dim(\text{null}(A)) + \dim(\text{im}(A)) = 5$ . However, since  $\text{im}(A)$  is in  $\mathbb{R}^3$ ,  $\dim(\text{im}(A)) \leq 3$ . Recall again that the dimension of the image is the dimension of the col space, which has 3 entries per column, so it is in  $\mathbb{R}^3$ . So, then  $\dim(\text{null}(A)) \geq 2$

### 1.3.3 Extending into a Basis

To extend a set of vectors into a basis for some  $\mathbb{R}^n$ , we add  $e_1, e_2, \dots, e_n$  to the matrix, augmenting it. Then, we row reduce it and take the columns of augmented  $A$  corresponding to the columns in  $A_{REF}$  with leading ones.

### 1.3.4 More Properties

- For invertible  $U$  and  $V$ ,  $\text{rank}(AV) = \text{rank}(UA)$ .
- The basis of  $\text{im}(A)$  is the set of columns corresponding to the leading ones in a row echelon equivalent of  $A$ .
- A basis of  $\text{null}(A)$  is the  $(n-r)$  basic solutions to  $A\vec{x} = \vec{0}$ . Solve the homogenous system and take the basic solutions for the parameters!

## 1.4 Equivalent Statements for a $M \times N$ matrix $A$

Theorem 1

The following are equivalent:

1.  $\text{rank}(A) = n$
2. The rows of  $A$  span  $\mathbb{R}^n$
3. The columns of  $A$  in  $\mathbb{R}^m$  are linearly independent
4.  $A^T A$  is invertible (Symmetric matrix)
5. There exists some  $C$  such that  $CA = I_n$
6.  $A\vec{x} = \vec{0}$  has only the trivial solution.

Theorem 2

1.  $\text{rank}(A) = m$
2. The cols of  $A$  span  $\mathbb{R}^m$
3. The rows of  $A$  in  $\mathbb{R}^n$  are linearly independent
4.  $A^T A$  is invertible (Symmetric matrix)
5. There exists some  $D$  such that  $AD = I_m$
6.  $A\vec{x} = \vec{b}$  has a unique solution for every  $\vec{b} \in \mathbb{R}^n$

## 1.5 Similarity

$A \sim B$  ( $A$  is similar to  $B$ ) if there exists some invertible  $P$ , such that  $B = P^{-1}AP$ . Some properties:

- $A \sim A$
- $A \sim B \Rightarrow B \sim A$
- $A \sim B, B \sim C \Rightarrow C \sim A$

The Trace ( $\text{tr}(A)$ ) of a matrix  $A$  gives the sum of the main diagonal entries  $\sum a_{i,i}$ . This may or may not be very useful.  $\text{tr}(AB) = \text{tr}(BA)$

Theorem: if  $A \sim B$  for square A, B, they share:

- Trace
- Determinants
- Characteristic Polynomials
- Eigenvalues

but NOT EIGENVECTORS!

Theorem: A is diagonalizable if there exists a basis of  $R^n$  composed of eigenvectors. Also, if  $\{\vec{x}_1 \dots \vec{x}_w\}$  for some DISTINCT eigenvalues  $\lambda_1 \dots \lambda_k$ , the set  $\{\vec{x}_1 \dots \vec{x}_k\}$  is linearly independent. If  $k = n$ , it is diagonalizable.

Thm: A is diagonalizable if and only if the number of basic eigenvalues  $E_\lambda A$  is equal to the multiplicity of  $\lambda$  in  $C_A x$ .

## 1.6 Least Squares Approximations

For  $A\vec{x} = \vec{b}$ , we can solve it if it's consistent. If it's inconsistent, we can use LSA to get the closest solutions for  $\vec{x}$ .

Geometrically, inconsistency appears as a vector outside of the plane formed by  $\text{im}(A)$ . It starts on the plane somewhere and comes off it. The point made by dropping the perpendicular from  $b$  (vector off the plane) onto the plane of  $\text{im}(A)$ , which we'll call  $z$ , is the closest we can get to a solution.

This gives us the normal equations

$$A^T \vec{b} = (A^T A) \vec{z}$$

. We know that a symmetric matrix is invertible, and we are solving for  $\vec{z}$ . We can use  $(A^T A)^{-1} A^T \vec{b} = \vec{z}$  to solve directly, minimizing  $|\vec{b} - A\vec{z}|$