Summary of Day 11

William Gunther

June 4, 2014

1 Objectives

- Define dimension, and give geometric interpretation.
- Be able to write a vector in a different coordinate system relative to a different basis.
- Begin talking about linear transformations.

2 Summary

- In linear algebra, we often want to capture invariants. That is, things that stay the same even when you alter them. We know lots of them: for instance, the row space is invariant under row operations (so if you do row operations, the row space does not change).

  Along with invariants are particular parameters or characteristics. For instance, we know (sort of-we haven't prove it) that the number of nonzero rows in a matrix's row echelon form does not dependent on the row echelon form you chose. Therefore, this is a parameter that we called rank.

  Now we will learn a new parameter called dimension. The definition will not make sense (as with rank) until we prove a certain invariance.

- The **dimension** of a subspace is the size of a basis. We denote the dimension of V by $\dim V$

  Theorem If V is a subspace of $\mathbb{R}^n$, then any two bases of V have the same number of vectors.

  Proof. This is theorem 3.23 in the book. We will prove it in class.

  Remark What should the dimension of the trivial subspace be?

- Intuitively, the dimension is exactly what we want to capture. 1 dimensional subspaces of $\mathbb{R}^n$ are lines since their basis is of size 1, so it's all vectors that are $c\mathbf{v}$. 2 dimensional objects are like planes, as it's vectors like $s\mathbf{v} + t\mathbf{u}$. 3 dimensional objects are...well, whatever they are called, which is a 3 dimensional hyperplanes, but that's not so important. Just like planes look like a copy of $\mathbb{R}^2$ in $\mathbb{R}^3$, so does a 3 dimensional subspace look like a copy of $\mathbb{R}^3$ in higher dimensions, like $\mathbb{R}^4$, it's just harder to picture.

- **Theorem** The row and column space of a matrix have the same dimension

  Proof. Let A be the matrix. Clearly $\dim(\text{row}(A)) = \dim(\text{row}(A'))$ where A' is in rref. $\dim(\text{row}(A'))$ is the number of nonzero rows of the matrix which, by the many connections we've made, is the same as the rank of the matrix (i.e. the number of zero rows is the same as the number of leading entries of A').

  Now, although $\text{col}(A') \neq \text{col}(A)$ they have a fundamental relationship: A is linearly independent if and only if A' is, and moreover, the dependency is demonstrated in the same way (e.g. if the first column is a multiple of the second in A' then it is so in A as well).

  We now make a claim that a moments thought makes clear: in A' all the columns without leading entries can we written as a linear combination of columns with leading entries. Why? Because all the columns with leading entries are elements of the standard basis of $\mathbb{R}^m$ and all the columns not containing leading entries only have nonzero components in places that have a leading entry!

  Therefore, the dimension of A' (which as noted is the same as the dimension of A) is equal to the number of leading entries, which is the rank, which is the number of nonzero rows, which is the dimension of the row space! Wow, cool. □

- We can then give an equivalent meaning for rank (which actually is not the last one we will see) which is more important when thinking about matrices rather than systems: The **rank** of a matrix is the dimension of the row and column spaces.
- What can you say about the connection between $\text{rank}(A)$ and $\text{rank}(A^T)$?
- With this new version of rank, one would expect that we could say more about The Rank Theorem, and rephrase it in this new matrix-centric way. You'd be right.

  With systems there is a tug of war: The more the rank has, the less free variables there will be. What is the tug of war between with regards to the dimension of the column and row spaces?
- The **nullity** of a matrix is the dimension of its null space.

  Theorem If A is $m \times n$ then:
  $$\text{rank}(A) + \text{nullity}(A) = n$$

  Proof. View A as a homogeneous system of equations. Then if the rank of A is r then there are $n - r$ total free variables by the rank theorem. The solution set of this system is exactly the nullspace, and as there are $n - r$ free variables, it has dimension $n - r$. Therefore, the nullity is $n - r$, and as we knew the rank is r, which makes the sum n.
- We can now add some things to the fundamental theorem of invertible matrices:

  Theorem (The fundamental theorem of invertible matrices version 2 (thm 3.27)) If A is $n \times n$ then the following are equivalent:

  1. A is invertible.
  2. $\text{rank}(A) = n$.
  3. $\text{nullity}(A) = 0$
  4. The column space of A is a basis for $\mathbb{R}^n$.
  5. The row space of A is a basis for $\mathbb{R}^n$.

  Proof. This is proved in the book with lots of other results filled in too. You should try to work out the proofs by yourself as practice.
- Now, why is all this basis and dimension stuff important?

  Theorem Let V be a k dimensional subspace of $\mathbb{R}^n$ with bases $B = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$. Then for every $\mathbf{v} \in V$ there exists a **unique** way to write $\mathbf{v}$ as a linear combination of the elements of B.

  Proof. Clearly there is a way to write $\mathbf{v}$ as a linear combination, since B spans the space (that is part of the definition of basis). If there were two ways then:
  $$\mathbf{v} = c_1\mathbf{v_1} + \cdots + c_k\mathbf{v_k}$$
  $$\mathbf{v} = d_1\mathbf{v_1} + \cdots + d_k\mathbf{v_k}$$
  Take their difference:
  $$\mathbf{0} = (c_1 - d_1)\mathbf{v_1} + \cdots + (c_k - d_k)\mathbf{v_k}$$
  As B is linearly independent (as it is a basis) $c_i = d_i$ for all i. □
- Now we can formalize something I mentioned long ago. When you take the span of a set of vectors we now see that you get a subspace. The subspace has some basis, which has some size (no matter what basis you choose it will be the same) called the dimension. The basis then creates a coordinate system for the subspace. Just like $\mathbf{e}_1, \ldots, \mathbf{e}_n$ is a coordinate system for $\mathbb{R}^n$ so is any basis by the last theorem.

  Formally: If $B = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is a basis for an n dimensional subspace of $\mathbb{R}^m$, and $\mathbf{v}$ is a element of this subspace then we already know we can write v uniquely as:
  $$\mathbf{v} = c_1\mathbf{v_1} + \cdots + c_n\mathbf{v_n}$$

So we define the **coordinates of v with respect to the basis** B to be the $mlistc_n$ that are the coefficents. We can write this as a vector:

$$[\mathbf{v}]_B = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}$$

which we call the **coordinate vector of v with respect to** B.

<u>**Example**</u> The vectors

$$B = \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \end{pmatrix} \right\}$$

are a basis for $\mathbb{R}^2$. This means that every vector in $\mathbb{R}^2$ can be written uniquely as a linear combination of these two vectors. You already know how to determine how to figure out how: we can set up an augmented matrix:

$$\left(\begin{array}{cc|c} 1 & -1 & a \\ 1 & 1 & b \end{array}\right)$$

Then you see that to write $[a, b]$ as a linear combination of these vectors, you take:

$$\frac{a+b}{2}\begin{pmatrix} 1 \\ 1 \end{pmatrix} + \frac{b-a}{2}\begin{pmatrix} -1 \\ 1 \end{pmatrix} = \begin{pmatrix} a \\ b \end{pmatrix}$$

Thus, for instance, the vector $[5, 7]$ can be written as:

$$\left[\begin{pmatrix} 5 \\ 7 \end{pmatrix}\right]_B = \begin{pmatrix} 6 \\ 1 \end{pmatrix}$$

- The moral of the story: there is nothing special about the way we write vectors. When we write a vector as $[5, 2]$ we are just writing with with respect to the *standard basis* $\mathbf{e_1}, \mathbf{e_2}$. We could similarly write this vector with respect to any other basis for $\mathbb{R}^2$.

- So, matrices are (as I've said repeatedly) a special type of function. Let's unlock exactly what that means. First we need to review what a function is:

  A **function** f is a mapping from a **domain** A to a **codomain** B. A function has a few guarantees, namely:

  - Every element from A gets mapped to somewhere.
  - Every element gets mapped somewhere in B.
  - There is only one thing in B that each element of A gets mapped to.

  The last condition actually implies the rest, but it's nice to say them all separately.

  The **range** of the function is the stuff in the codomain that actually gets hit.

  <u>**Example**</u> $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = x^2$ has domain $\mathbb{R}$, codomain $\mathbb{R}$, and range $\{ x \in \mathbb{R} \mid x \geq 0 \}$

- $T : \mathbb{R}^n \to \mathbb{R}^m$ is called a **linear transformation** if:

  - $T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$ for all vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$.
  - $T(c\mathbf{u}) = cT(\mathbf{u})$ for all scalars $\mathbf{u}$ and all vectors $\mathbf{u} \in \mathbb{R}^m$.

  As with subspaces, we can abbreviate this to say:

  - $T(0) = 0$
  - $T(\mathbf{u} + c\mathbf{v}) = T(\mathbf{u}) + cT(\mathbf{v})$ for all $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ and all scalars c.

- Matrices are important because...

  <u>**Theorem**</u> Let A be a $m \times n$ matrix. Defined $T_A : \mathbb{R}^n \to \mathbb{R}^m$ by:

  $$T_A(\mathbf{x}) = A\mathbf{x}$$

  this is a linear transformation.