

Princeton University MAT 202 Spring 2008

Willie W. Wong

Feb 25 - 29

FEB 25, 2008

- Recall vaguely from last Friday, the final example that we did. We had 8 vectors $\mathbf{v}_1 \dots \mathbf{v}_8$ in \mathbb{R}^7 . We want to find a basis for their span. Remember that the span of such 8 vectors in \mathbb{R}^7 can be thought of as the image of the linear transformation given by the 7×8 matrix

$$A = [\mathbf{v}_1 \quad \dots \quad \mathbf{v}_8]$$

we have that the solution set of $A\mathbf{x} = 0$ gives us some information about the linear dependence of the vectors. In particular, we have that the number of free-parameters to $A\mathbf{x} = 0$ is the number of extra vectors we have from the list; this is because each free parameter corresponds to 1 linear relation, and hence 1 redundant vector.

- Let's make precise this way of "counting stuff".
- To start, first recall the following: given vectors $\mathbf{v}_1, \dots, \mathbf{v}_M$ in \mathbb{R}^N , the below statements are equivalent
 - They are linearly independent
 - None of the vectors are redundant (written as a linear combination of the preceding ones)
 - None of the vectors can be written as a linear combination of the others
 - The only linear relation among them is the trivial one
 - $\ker([\mathbf{v}_1 \quad \dots \quad \mathbf{v}_M]) = \{0\}$
 - $\text{rank}([\mathbf{v}_1 \quad \dots \quad \mathbf{v}_M]) = M$
- Fact: all bases (notice, the plural of "basis" is "bases") of a subspace V have the same number of vectors.
 - First we establish the following: let $\mathbf{v}_1, \dots, \mathbf{v}_M$ in V be linearly independent. And let $\mathbf{w}_1, \dots, \mathbf{w}_Q$ be such that they span V , then $Q \geq M$.

- To do so, we use the definition of matrix multiplication. Since the w s span V , we can write every one of v_i as a linear combination of the w s:

$$v_i = c_{1,i}w_1 + c_{2,i}w_2 + \cdots + c_{Q,i}w_Q$$

Now, write B for the matrix formed by the column vectors v_i and A for the matrix formed by the column vectors w_j . Then the j th term of the vector v_i is the entry b_{ji} of the matrix B . So the equation here becomes

$$b_{ji} = c_{1,i}a_{j1} + c_{2,i}a_{j2} + \cdots + c_{Q,i}a_{jQ}$$

which is the same as

$$B = AC$$

the multiplication by matrices. Suppose that V is a subspace of \mathbb{R}^N , then B is an $N \times M$ matrix. A is an $N \times Q$ matrix. C is an $Q \times M$ matrix.

- Now, let's look at the solution set $Cx = 0$. Suppose $x \in \ker(C)$, then $ACx = A0 = 0$. So $Bx = 0$. This means that every x in $\ker(C)$ is in $\ker(B)$. (Draw Venn-type diagram.)
- But we know that the column vectors of B are linearly independent, so $\ker(B) = \{0\}$. Hence the kernel of C is necessarily the set $\{0\}$. Which means that $Q \geq \text{rank}(C) = M$, so $Q \geq M$.
- The proof is technical and not too illuminating. What you should keep in mind is the geometry. Morally speaking, since w_1, \dots, w_Q spans V , we first remove all the redundant ones and re-label them $u_1 \dots u_U$ where $U \leq Q$. Give that the u s are linearly independent, we would like to find a linear transformation such that $T(u_1) = v_1, T(u_2) = v_2$ etc... obviously, this scheme cannot work if $U \leq Q < M$, since we will run out of u s.
- Now we can state the fact that all bases of V have the same number of vectors. (Why? let $v_1 \dots v_M$ and $w_1 \dots w_Q$ both be bases of V . By the previous argument, we have that $M \leq Q$ and $Q \leq M$. So $Q = M$.)
- So this number, the number of vectors in a basis, is a property of the subspace V that does not depend on which basis you take. This is the intrinsic property of a subspace called its *dimension*. We write it as $\dim(V)$.
- Now, suppose we have a subspace V of \mathbb{R}^N . Suppose $\dim(V) = P$ (which, presumably, we know because we have an explicit basis of V with P vectors). Then we have the following facts concerning a list of Q vectors v_1, \dots, v_Q all sitting in V
 - If the list is linearly independent, then $Q \leq P$.
 - If the list spans V , then $Q \geq P$.
 - If $Q = P$ exactly, and the list is linearly independent, then the list is a basis.

– If $Q = P$ exactly, and the list spans V , then the list is a basis.

- Example, consider the example from last Friday again

$$A = \begin{bmatrix} 1 & 5 & 3 & 5 & 3 & 4 & 2 & 1 \\ 0 & 4 & 5 & 7 & 2 & 3 & 1 & 4 \\ 1 & 0 & 4 & 6 & 3 & 0 & 1 & 1 \\ 0 & 1 & 1 & 4 & 3 & 0 & 1 & 4 \\ 2 & 3 & 3 & 4 & 3 & 2 & 3 & 1 \\ 2 & 3 & 3 & 2 & 1 & 1 & 4 & 5 \\ 1 & 5 & 4 & 5 & 2 & 3 & 3 & 4 \end{bmatrix}$$

We've found a basis for its image. Let's recall how we did it: we try to eliminate redundant column vectors. To do so, we want to establish linear relations between the column vectors. This means that we try to solve the equation

$$A\mathbf{x} = 0$$

Whose solution we calculated to be the 2-parameter family.

$$x_1 = -s - t$$

$$x_2 = -t$$

$$x_3 = s$$

$$x_4 = -s$$

$$x_5 = s$$

$$x_6 = t$$

$$x_7 = t$$

$$x_8 = 0$$

The vectors corresponding to the free parameters (the vectors \mathbf{v}_5 and \mathbf{v}_7) are then the redundant vectors, as we can see by treating setting all but one of the free parameters to 0 each time, and the remaining one to 1.

- How about finding a basis for the kernel? The linear relations that we can derive from treating each of the free parameters separately (i.e., setting first $s = 0, t = 1$ and then $s = 1, t = 0$) are

$$-\mathbf{v}_1 - \mathbf{v}_2 + \mathbf{v}_6 + \mathbf{v}_7 = 0$$

and

$$-\mathbf{v}_1 + \mathbf{v}_3 - \mathbf{v}_4 + \mathbf{v}_5 = 0$$

This means that we know at least two solutions to the equation: let

$$\mathbf{y} = \begin{bmatrix} -1 \\ -1 \\ 0 \\ 0 \\ 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{z} = \begin{bmatrix} -1 \\ 0 \\ 1 \\ -1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

we have that

$$A\mathbf{y} = A\mathbf{z} = 0$$

The claim is now that the vectors obtained by setting the free-parameters to 0 and 1 form a basis for the kernel. First we need to check that they are linearly independent. This is quite obvious since there are only two of them and neither is a multiple of the other. Next we need to check that they span the solution set. Well, the kernel, i.e. the solution for $A\mathbf{x} = 0$ has the 2-parameter family of solutions given above. You can see that the family can be re-written as

$$\mathbf{x} = t\mathbf{y} + s\mathbf{z}$$

- So, in general: let $T : \mathbb{R}^M \rightarrow \mathbb{R}^N$ be a linear transformation with matrix A . Find bases for its kernel and image.
 - Write A in column vectors $[\mathbf{v}_1 \ \dots \ \mathbf{v}_M]$.
 - Solve the equation $A\mathbf{x} = 0$ by taking the rref of the augmented matrix $[A \ : \ 0]$.
 - The columns corresponding to the free-parameters are redundant. Once you remove those, the remaining column vectors form a basis for the image.
 - Writing the solution as a parametrized family (by reading off from the rref of the augmented matrix), we will have

$$\mathbf{x} = s_1\mathbf{w}_1 + \dots + s_L\mathbf{w}_L$$

where L is the number of free parameters. The vectors $\mathbf{w}_1 \dots \mathbf{w}_L$ are the basis of the kernel.

- Notice that a basis of the image of a transformation consists of the column vectors from the columns with the leading ones in rref of the matrix. This means:

$$\dim(\text{im}(A)) = \text{rank}(A)$$

further more, the number of vectors in the basis of the kernel (i.e., the dimension of the kernel) is the same as the number of free parameters. And now remember from chapter 1 that, if A is $N \times M$

$$\dim(\text{ker}(A)) = \{ \text{number of free parameters} \} = M - \text{rank}(A)$$

so we have

- The Rank-Nullity Theorem: we call the value $\dim(\ker(A))$ the *nullity* of A . Then the rank-nullity theorem states that

$$\{\text{rank of } A\} + \{\text{nullity of } A\} = M$$

and

$$\dim(\text{im}(A)) + \dim(\ker(A)) = M$$

- Geometrically, this means that every linear transformation contains, in some sense, a projection operation. Remember in 2 dimensions, the projections onto a line collapses one of the 2 dimensions and leaves only 1 dimension. The rank-nullity theorem says that that of the M dimensions in the domain, some number of dimensions (the nullity) is collapsed in the linear transformation (the nullity being 0 implies that no dimension is collapsed), while some number survive (the rank), and every dimension is accounted for in this decomposition.
- A small digression: now that we have considered bases of subspaces, how about the subspace \mathbb{R}^N of \mathbb{R}^N ? We know 1 basis, the standard one given $e_1 \dots e_N$. We also know that a basis of \mathbb{R}^N must have N vectors. So the matrix corresponding to the column vectors is $N \times N$. Since the basis is independent, the matrix has rank N . So: *the vectors v_1, \dots, v_N of \mathbb{R}^N forms a basis if and only if the matrix*

$$[v_1 \quad \dots \quad v_N]$$

is invertible.

FEB 27, 2008

Today is the first quiz of the semestre. Covers chapters 1 and 2.

FEB 29, 2008

- Today's topic: coördinates. We begin with a figure. In the picture, the purple arrow is some vector v . Up until now, we would have been unhesitant about writing the expression

$$v = \begin{bmatrix} 6.5 \\ 4.5 \end{bmatrix}$$

Indeed, in the coördinate system defined by the black axes (which corresponds to the standard or Cartesian coördinates we have been using up until now), v can be read off as the ordered pair $(6.5, 4.5)$.

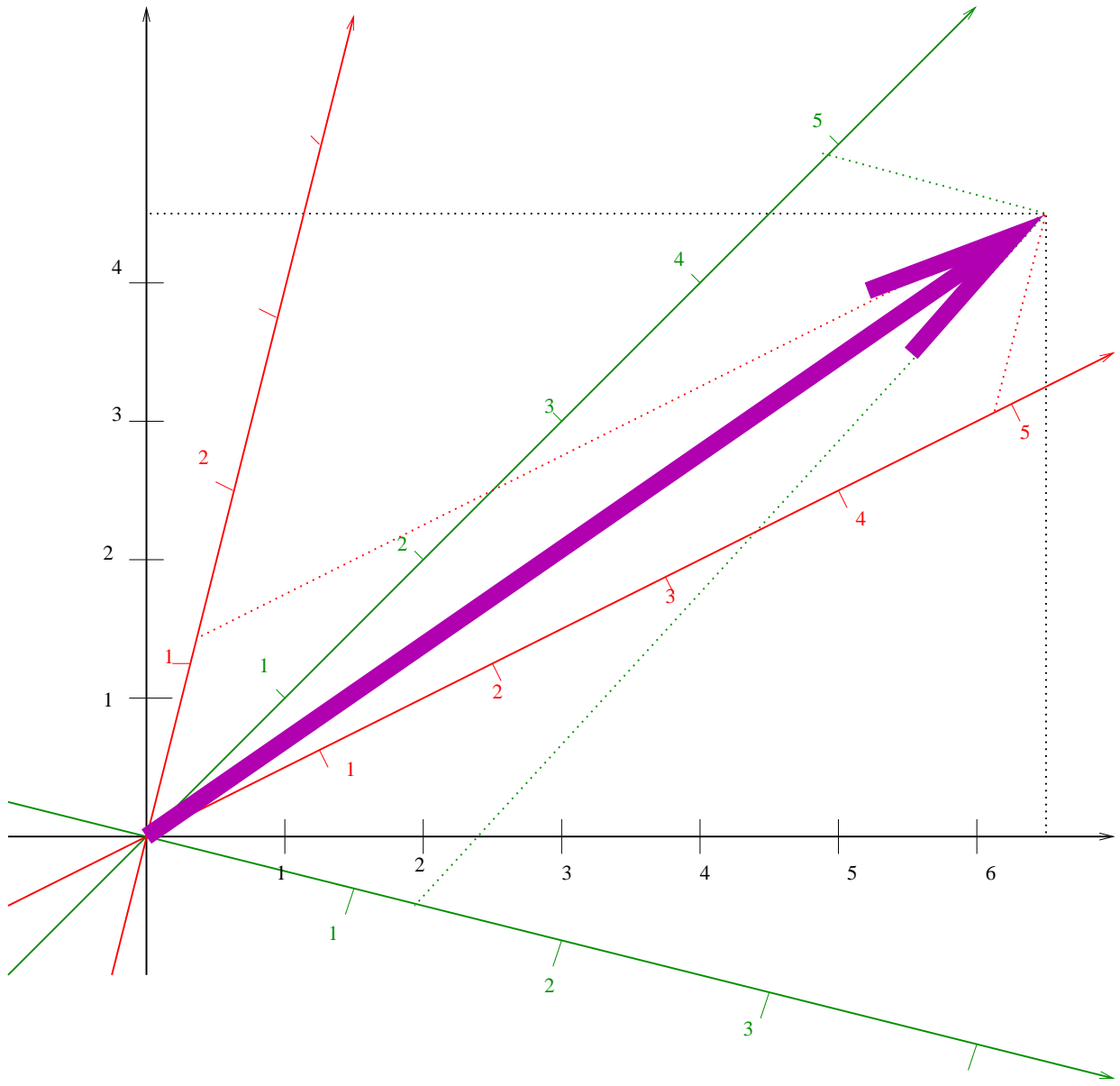


Figure 1: This figure illustrates how one vector (the purple arrow) is viewed in three separate coordinate systems (the black, red, and the green sets of arrows). The dotted lines are the “drop lines” from which we can read off the coordinates of the vector in the various different coordinate systems; notice that the drop lines are parallel to the coordinate axes. The black system corresponds to the standard or Cartesian coordinates we are most used to.

- While the Cartesian coordinate is a convenient tool of calculation (and Descartes was certainly one cool cat), especially in analytic geometry, we want to, for now, free the vector from the encumbrance of a fixed set of $x - y$ axes.
- Philosophically we are trying to give recognition to vectors as *Ding an sich*. For starters, even if we stick with the Cartesian system, what the vector looks like depends strongly on your point of view: suppose an arrow is drawn on the ground pointing northwesterly. For a passerby who thinks of the left-right direction as his x axis and the front-back direction as his y -axis, if he approaches from the east, he will see the arrow as pointing in the $(1, 1)$ direction, whereas if he approaches from the south, he will see the arrow pointing in the $(-1, 1)$ direction. Furthermore, while the Cartesian system has the nice property that the axes are perpendicular to each other, this is certainly not a requirement for a system of coordinates: a system of coordinates merely needs to assign to every point in space a unique identifier.¹
- So, even though most of our calculations will be done in the framework of Cartesian coordinates, I want to motivate some things slightly more generally.
- Going back to the picture. What does it mean for the purple vector to be $\begin{bmatrix} 6.5 \\ 4.5 \end{bmatrix}$? It means that the purple vector extends 6.5 units of length in the x -direction and 4.5 units of length in the y -direction. In vector notation, this is

$$\mathbf{v} = 6.5\mathbf{e}_1 + 4.5\mathbf{e}_2$$

- Remember from last Friday that, given a basis $\mathcal{B} = \{\mathbf{w}_1, \dots, \mathbf{w}_P\}$ of a subspace W , any vector $\mathbf{w} \in W$ has a unique expression as

$$\mathbf{w} = c_1\mathbf{w}_1 + \dots + c_P\mathbf{w}_P$$

we call the numbers c_1, c_2, \dots, c_P the \mathcal{B} -coordinates of the vector \mathbf{w} . Let's think about how we assign the standard coordinates to a figure in the plane: we basically take two meter sticks (or yard sticks, if you want to go English), hold them at 90 degree angles to each other, choose a point as the origin, and draw two axes. Then, for a given point, we drop a line parallel to the y -axis (see the figure) and find its intersect with the x -axis, and call it the x -coordinate of the point. We drop a line parallel to the x -axis and find its intersect with the y -axis to call it the y -coordinate of the point. Now imagine that we have two sticks marked at regular intervals, but they need not be the same: we can have, say, a meter stick marked with centimeters, and a yard stick marked with inches. Now we hold them up, not perpendicular to each other (but also not parallel), at some crooked angle. We can still draw two axes. We can still call the axes x and y . We can still drop the parallel lines and find the intercepts and give the points coordinates. But the numbers will be different from before. This is the meaning of the \mathcal{B} -coordinates. You "measure" where a point is based on several crookedly-held, non-standard sticks.

¹For those of you who have read Kant, the vector is the *noumenon*, while the expression of the vector in a coordinate system, e.g. as an N -tuple of numbers, is the *phenomenon*.

- Look at the two green arrows in the picture: they are not parallel, so they span the whole \mathbb{R}^2 . So do the two red arrows. We can use either set to define coordinates: in the green system, the purple vector is roughly $\begin{bmatrix} 1.25 \\ 4.9 \end{bmatrix}$ and in the red system it is $\begin{bmatrix} 4.9 \\ 1.15 \end{bmatrix}$. Let \mathbf{g}_1 be the vector that has coordinate $(1, 0)$ on the green system, and \mathbf{g}_2 be the one with $(0, 1)$, and similarly for \mathbf{r}_1 and \mathbf{r}_2 , then in vector notation we have

$$\mathbf{v} = 1.25\mathbf{g}_1 + 4.9\mathbf{g}_2 = 4.9\mathbf{r}_1 + 1.15\mathbf{r}_2$$

- A bit of a notation that Bretscher uses (but we'll try not to use, because it is clumsy and prone to confusion). He writes $[\mathbf{w}]_{\mathcal{B}}$ for the *column vector whose entries are the \mathcal{B} -coordinates of \mathbf{w}* .
- In this notation, going back to the picture, we have

$$[\text{purple vector}]_{\text{standard (black) basis}} = \begin{bmatrix} 6.5 \\ 4.5 \end{bmatrix}$$

and

$$[\text{purple vector}]_{\text{green basis}} = \begin{bmatrix} 1.25 \\ 4.9 \end{bmatrix}, \quad [\text{purple vector}]_{\text{red basis}} = \begin{bmatrix} 4.9 \\ 1.15 \end{bmatrix}$$

HOMWORK FOR THIS WEEK

Homework will be due on Wednesday of next week, instead of Monday, since we didn't finish covering section 3.4.

3.2: 6, 8, 18, 28, 43, 45, 53, 54

3.3: 25, 27, 32, 35, 38, 52, 61

3.4: 17, 30, 45, 53, 56, 71