Lecture notes by John Noonan (02/05/2015)

1 Continuation of Vector Spaces

1.1 Show that scalar multiplication distributes over vector addition

For some vector space $\mathbf{V} \in (\mathbf{R}^2, \mathbf{R})$

$$(\alpha + \beta)\mathbf{V} = \alpha\mathbf{V} + \beta\mathbf{V} \tag{1}$$

$$\alpha(\mathbf{V} + \mathbf{W}) = \alpha\mathbf{V} + \alpha\mathbf{W} \tag{2}$$

To prove this, we should reduce to just addition and scalar multiplication so that we don't implicitly assume anything that we are trying to prove.

(1) Proof:

For the purpose of the proof, let $\mathbf{V} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

$(\alpha + \beta)\mathbf{V} = (\alpha + \beta) \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

$= \begin{bmatrix} (\alpha + \beta)x_1 \\ (\alpha + \beta)x_2 \end{bmatrix}$

$= \begin{bmatrix} \alpha x_1 + \beta x_1 \\ \alpha x_2 + \beta x_2 \end{bmatrix}$ (Distributive property of real numbers)

$= \begin{bmatrix} \alpha x_1 \\ \alpha x_2 \end{bmatrix} + \begin{bmatrix} \beta x_1 \\ \beta x_2 \end{bmatrix}$

$= \alpha \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \beta \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ (Take out α, β)

(2) Can do as an exercise

1.2 Basis of a Vector Space – Illustrative Example

To describe the basis of a vector space, we will use an example of choosing an eating place at Berkeley.

Consider all vectors of length 5: $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix}$ in the vector space
($\mathbf{V}$, {0, 1}) such that $x_i \in$ {0, 1}, GF(2). This requires all elements of the vector space to be 0 or 1.
We can also represent this as:
$\mathbf{V} : \{[x_1, x_2, ..., x_5]^T | x_i \in \{0, 1\}\}$
(GF(2) denotes the binary field, i.e the only numbers are $\{0, 1\}$ and addition and multiplication are defined as follows:)
Addition:

$$0 + 0 = 0$$
$$0 + 1 = 0$$
$$1 + 0 = 1$$
$$1 + 1 = 0$$

Multiplication:

$$0 * x = 0$$
$$1 * x = x$$

Here, we are creating different notions of addition and multiplication. Addition in this field is the storing of bits where the 2 vectors are "wrapped" back to each other. With the multiplication field, we implicitly look at the remainder of dividing by 2.
We can use this field to study the different food preferences of the students and instructors in the class by representing the data in a utility matrix of 0s and 1s.

Food Places:
- Brewed Awakening
- Stuffed Inn
- Frozen Yogurt
- Nefeli
- Celia's

Preferences:

Professors:	G	E	C	B	V	
	1	0	1	0	1	(Brewed Awakening)
	1	1	0	0	0	(Stuffed Inn)
	0	0	1	0	1	(Frozen Yogurt)
	1	1	0	1	1	(Nefeli)
	0	1	1	1	1	(Celia's)

1: like, 0: doesn't like G: Gireeja, E: Elad, C: Claire, B: Babak, V: Vivek

At this point, one might wonder if a set of only 5 instances is large enough to express anyone's preferences in the whole class. The answer is *yes*! In other words, the 5 instructors serve as a **basis** for the vector space.

For example, suppose Student1 and Student2 have the following preferences:

Student1:
0
0
1
0
0
Student2:
1
0
0
1
0

So, again, the question is: Can Student1 and Student2's preferences be represented by the *basis* created by the 5 instructors' preferences such that

Student1(or Student2) = $\alpha_G * G + \alpha_E * E + \alpha_C * C + \alpha_B * B + \alpha_V * V$?

We see that we *can* obtain Student1's preferences-vector by *linearly* combining the instructors' preferences.

This means that we can re-create Student's preferences by doing the following:
Student1 = 1 * G + 1 * E + 1 * C + 0 * B + 0 * V

$$= 1 * \begin{bmatrix}1\\1\\0\\1\\0\end{bmatrix} + 1 * \begin{bmatrix}0\\1\\0\\1\\1\end{bmatrix} + 1 * \begin{bmatrix}1\\0\\1\\0\\1\end{bmatrix} + 0 * \begin{bmatrix}0\\0\\0\\1\\1\end{bmatrix} + 0 * \begin{bmatrix}1\\0\\1\\1\\1\end{bmatrix}$$

$$= \begin{bmatrix} 1+0+1 \\ 1+1+0 \\ 0+0+1 \\ 1+1+0 \\ 0+1+1 \end{bmatrix}$$

$$= \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$ where we utilize our special addition notion to simplify the matrix.

Similarly, we can compute Student2's preferences-vector as a linear combination of the instructors' vectors.

Student2 = 1 * G + 1 * E + 0 * C + 1 * B + 0 * V

$$= 1 * \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} + 1 * \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \\ 1 \end{bmatrix} + 0 * \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 1 \end{bmatrix} + 1 * \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} + 0 * \begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} 1+0+0 \\ 1+1+0 \\ 0+0+0 \\ 1+1+1 \\ 0+1+1 \end{bmatrix}$$

$$= \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

So, again, the set of instructors form a **basis** for this vector space. If we only had 4 instructors, would the instructors' set form a basis? *No*, the dimension of the space is too large.
Also, if 2 instructors had identical food choices, this would not form a basis. Each vector of the basis needs to be linearly independent.

This way of working with preferences is very applicable. For example, companies like Amazon or Netflix use this to understand how users are similar to each other.

This brings us to the formal definition of *Linear Independence*: (We will formally define the notion of a basis later.)

1.3 Linear Independence/Dependence

Given a vector space $(\mathbf{V}, \mathbf{F})$, the set of vectors
$\{\mathbf{V}_1, \mathbf{V}_2, ..., \mathbf{V}_p\}$ where $\mathbf{V}_i \in \mathbf{V}$ are linearly independent if for scalars
$(\alpha_1, \alpha_2, ..., \alpha_p), \alpha_1 \mathbf{V}_1 + \alpha_2 \mathbf{V}_2 + ... + \alpha_p \mathbf{V}_p = 0 \Rightarrow$ (implies) $\alpha_i = 0 \ \forall$ (for every i) $1 \leq i \leq p$

In other words, if the only way to get 0 is to multiply each of the vectors by 0, then that set of vectors is linearly independent.

Example (Application): Suppose we have a robot with wheels, and it only has 1 motor which allows it to drive left or right in some playing field.

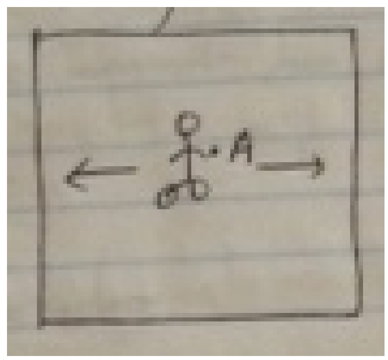

Figure 1: Robot with 1 motor in a playing field

However, with this robot it is not possible to reach every part of the playing field. If we were to put a second motor into the robot, what type of motor would we put in?
If we put in another motor that only moves left to right, we return to the same issue we had before, but if we add a motor so that the robot can move up and down as well, the robot would now be able to access every part of the field.

In other words, when we added a similar motor (making the motors, or basis, linearly dependent), our problem was not solved; in contrast, when we added a motor that allowed the robot to move in a different set of directions, we solved our problem since our motors were "linearly independent."

We will now define the notion of ***linear dependence***:

Given a vector space $(\mathbf{V}, \mathbf{F})$, the set of vectors
$\{\mathbf{V}_1, \mathbf{V}_2, ..., \mathbf{V}_p\}$ where $\mathbf{V}_i \in \mathbf{V}$ are linearly dependent if for scalars
$(\alpha_1, \alpha_2, ..., \alpha_p), \alpha_1\mathbf{V}_1 + \alpha_2\mathbf{V}_2 + ... + \alpha_p\mathbf{V}_p = 0$ where not all α_i's are zero.

Example 1 Suppose we have the set of vectors $\{\mathbf{V}_1, \mathbf{V}_2, 0\}$ such that $\mathbf{V}_i \in \mathbf{R}^3$, are these elements of $\mathbf{R}^3$ linearly independent?

The answer is **no**, because there is a **0** vector. For example, if $1 * \mathbf{V}_1 = -1 * \mathbf{V}_2$, then for any scalar α, $\alpha * \mathbf{0} = 0$. This means that for $\alpha_i \neq 0$, it is possible to have
$\alpha_1\mathbf{V}_1 + \alpha_2\mathbf{V}_2 + ... + \alpha_p\mathbf{V}_p = 0$.

Example 2 Are the following 2 vectors linearly dependent or independent?

$$\begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

The answer is that they are **linearly independent**. This is because the only way that
$$\alpha * \begin{bmatrix} 0 \\ 1 \end{bmatrix} + \beta * \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \mathbf{0}$$
is if $\alpha = 0$ or $\beta = 0$.

We will now formally define a **basis**:

1.4 Definition of a Basis

Given a vector space $(\mathbf{V}, \mathbf{F})$, a set of vectors $\{\mathbf{b}_1, \mathbf{b}_2, ..., \mathbf{b}_n\}$ is a basis if

1. $\{\mathbf{b}_1, \mathbf{b}_2, ..., \mathbf{b}_n\}$ are linearly independent.

2. $\forall$ vector $\mathbf{b} \in \mathbf{V}$, there exists $\alpha_1, \alpha_2, ..., \alpha_n$ such that $\mathbf{V} = \alpha_1 \mathbf{b}_1 + \alpha_2 \mathbf{b}_2 + ... + \alpha_n \mathbf{b}_n$

This means that if we can express 1 vector as some linear combination of another vector, the independence is lost, and these 2 vectors do not form a basis.

We can now describe the meaning of a *dimension*:

Dimension The dimension of a vector space is simply the number of basis vectors.

Example

The vector space that is the set of vectors used to form the "5 x 5 identity matrix" is of dimension 5 x 5 since it requires 5 vectors to form the basis. Specifically,

$$\begin{bmatrix}1\\0\\0\\0\\0\end{bmatrix}, \begin{bmatrix}0\\1\\0\\0\\0\end{bmatrix}, \begin{bmatrix}0\\0\\1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\0\\0\\1\\0\end{bmatrix}, \begin{bmatrix}0\\0\\0\\0\\1\end{bmatrix} \Rightarrow \begin{bmatrix}1&0&0&0&0\\0&1&0&0&0\\0&0&1&0&0\\0&0&0&1&0\\0&0&0&0&1\end{bmatrix}$$

Note: All bases will have exactly the same size.

1.5 "Geometry" of a Basis

We can think of the *rotation* of vectors as essentially being a change of bases.

First of all, in order to rotate a vector, we multiply by a *rotation matrix*.

If we can find the x and y coordinates of a point in the standard basis, then we can also find the x and y coordinates of the point in our new basis. In this example, our original bases were

$$\begin{bmatrix}1\\0\end{bmatrix}, \begin{bmatrix}0\\1\end{bmatrix}$$

Now, after we multiply these vectors by the rotation matrix below, we have a new set of bases:

$$\begin{bmatrix}\cos(\theta) & -\sin(\theta)\\ \sin(\theta) & \cos(\theta)\end{bmatrix} \quad \cdot \quad \begin{bmatrix}1\\0\end{bmatrix} \quad = \quad \begin{bmatrix}\cos(\theta)\\ \sin(\theta)\end{bmatrix}$$

Rotation matrix   Original Vector $\Rightarrow$ Rotated Vector

Similarly,

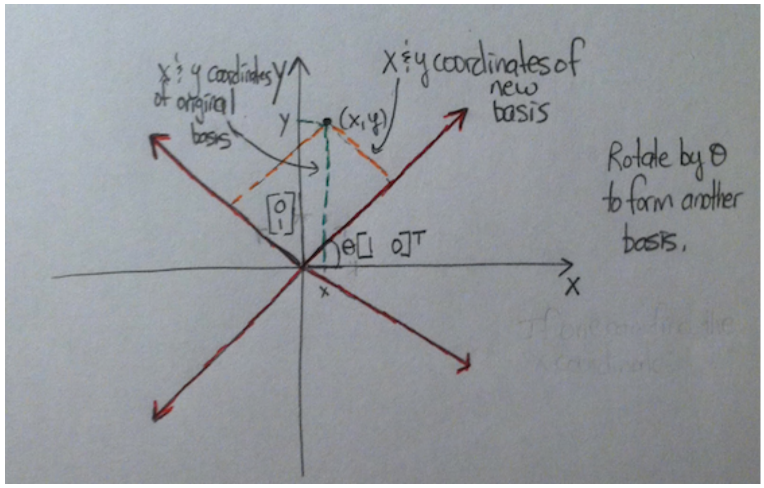

Figure 2: Geometric Depiction of Basis Transformation

$$\begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \quad \cdot \quad \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -\sin(\theta) \\ \cos(\theta) \end{bmatrix}$$

Thus, our new basis that was formed by being rotated is:

$$\begin{bmatrix} \cos(\theta) \\ \sin(\theta) \end{bmatrix}, \begin{bmatrix} -\sin(\theta) \\ \cos(\theta) \end{bmatrix}$$

More generally, we can think of matrix multiplication as a transformation of space, the same way the rotation matrix is a transformation of space. Multiplication by the rotation matrix maps a vector in $\mathbb{R}^2$ to another vector in $\mathbb{R}^2$. This property is actually true for any 2×2 matrix, not just the rotation matrix. Multiplication by any 2×2 matrix maps a 2×1 vector to another 2×1 vector. Thus, matrix multiplication is just another way of capturing space transformations.

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x' \\ y' \end{bmatrix} \text{ (Transforming the space)}$$

Demo of Basis Transformation To see a demo of how shapes get transformed when we change the bases through rotation, go to the following link:
`http://demonstrations.wolfram.com/MatrixTransformation/`

1.6 Matrix Multiplication i.e. Transformation of Spaces

$$\underset{A}{\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}} \underset{B}{\begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}} = \underset{AB}{\begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{12}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{bmatrix}}$$

We can think of matrix multiplication as cascading – we take the dot product of each row vector with the column vector of the other matrix starting from the first column of matrix B and going to the right.

Another description:

$$\begin{bmatrix} ----R_1---- \\ ----R_2---- \\ . \\ . \\ . \\ ----R_N---- \end{bmatrix} \begin{bmatrix} | & | & | \\ | & | & | \\ C_1 & C_2 \ldots & C_N \\ | & | & | \\ | & | & | \end{bmatrix} = \begin{bmatrix} R_1C_1 & R_1C_2 & \ldots & R_1C_N \\ R_2C_1 & R_2C_2 & \ldots & R_2C_N \\ & . & & \\ & . & & \\ & . & & \\ R_NC_1 & R_NC_2 & \ldots & R_NC_N \end{bmatrix}$$

where each matrix is an $N \times N$ matrix.

In fact, we can think of solving systems of equations as matrix multiplication:

$$a_{11}x_1 + a_{12}x_2 = b_1$$
$$a_{21}x_1 + a_{22}x_2 = b_2$$
$$\Downarrow$$
$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

Here, we solve for x_1 and x_2 where b_1 and b_2 capture all the information for x_1 and x_2.

Example $$\begin{bmatrix} 2 & 4 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} (2)(1)+(4)(3) & (2)(2)+(4)(4) \\ (3)(1)+(1)(3) & (3)(2)+(1)(4) \end{bmatrix}$$

$$= \begin{bmatrix} 14 & 20 \\ 6 & 10 \end{bmatrix}$$

1.7 The Inverse of a Matrix

Definition Given a square matrix $\mathbf{A}$ (with dimension n x n), if there exists $\mathbf{X_{nxn}}$ such that $\mathbf{X} \cdot \mathbf{A} = \mathbf{I_{nxn}}$, then $\mathbf{X}$ is the *inverse* of $\mathbf{A}$, and we write $\mathbf{X} = \mathbf{A^{-1}}$.

Special Properties of Square Matrices

* $(\mathbf{A^{-1}})^{-1} = \mathbf{A}$
* $\mathbf{A^{-1}A} = \mathbf{AA^{-1}} = \mathbf{I_{nxn}}$

We can use inverses of matrices to solve systems of linear equations. For example to solve, $\mathbf{A} \cdot \begin{bmatrix} \mathbf{v_1} \\ \mathbf{v_2} \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$ we can use the inverse of $\mathbf{A}$.

$$\Rightarrow \mathbf{A^{-1}A} \begin{bmatrix} \mathbf{v_1} \\ \mathbf{v_2} \end{bmatrix} = \mathbf{A^{-1}} \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \Rightarrow \mathbf{I} \begin{bmatrix} \mathbf{v_1} \\ \mathbf{v_2} \end{bmatrix} = \mathbf{A^{-1}} \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \Rightarrow \begin{bmatrix} \mathbf{v_1} \\ \mathbf{v_2} \end{bmatrix} = \mathbf{A^{-1}} \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

1.8 Examples of Inverses of Matrices

(1) $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ (Exactly the same)

(2) $\begin{bmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{bmatrix}^{-1} = \begin{bmatrix} 1/a_{11} & 0 & 0 \\ 0 & 1/a_{22} & 0 \\ 0 & 0 & 1/a_{33} \end{bmatrix}$,

if $a_{11}, a_{22}, a_{33} \neq 0$. If any of the diagonal elements are zero, the matrix is not invertible, or "singular". To check if we computed the inverse correctly, we simply multiply our original matrix by the inverse matrix we just solved:

$$\begin{bmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{bmatrix}^{-1} \cdot \begin{bmatrix} \frac{1}{a_{11}} & 0 & 0 \\ 0 & \frac{1}{a_{22}} & 0 \\ 0 & 0 & \frac{1}{a_{33}} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Note: However, not all matrices are invertible.

1.9 Computing the Inverse of a Matrix

Suppose $\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$

Claim: $\mathbf{A}^{-1} = \frac{1}{?} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}$

Find ?, i.e. what is the constant?

We can multiply the two matrices to get:

$$\begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}^{-1} \cdot \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} a_{11}a_{22} - a_{12}a_{21} & a_{22}a_{12} - a_{12}a_{22} \\ -a_{11}a_{21} + a_{11}a_{21} & -a_{21}a_{12} + a_{11}a_{22} \end{bmatrix} = \begin{bmatrix} a_{11}a_{22} - a_{12}a_{21} & 0 \\ 0 & -a_{21}a_{12} + a_{11}a_{22} \end{bmatrix}$$

In order for this matrix to be the identity matrix, we would need to multiply by $\frac{1}{a_{11}a_{22}-a_{12}a_{21}}$. Thus, ? = $a_{11}a_{22} - a_{12}a_{21}$. Thus, the formula to compute the inverse of a 2×2 matrix is:

$$\mathbf{A}^{-1} = \frac{1}{a_{11}a_{22}-a_{12}a_{21}} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}$$

Applications of Inverses of Matrices Matrix inversion is important to translate data represented in one form in another form, and recover different representations of data from each other. For example, the ideas of inversion are useful for encoding and decoding wireless transmissions, say from a cellphone to a cell tower.

For example, data that was sent could be $\begin{bmatrix} x_1 \\ x_2 \\ . \\ . \\ . \\ x_n \end{bmatrix}$ while the received data could be $\begin{bmatrix} ? \\ ? \\ x_3 \\ ? \\ ? \end{bmatrix}$

Linear independence and bases are useful to understand how data can be represented in different formats and how to recover these formats from each other.