# Linear algebra notes part 2

Yoel Sanchez Araujo

October 2020

## 1 Systems of equations

You'll remember that at some point in your grade school years you solved a set of equations that looked something like:

$$\begin{aligned} x_1 + 2x_2 &= 5 \\ x_1 + 6x_2 &= 13 \end{aligned} \tag{1}$$

we can write this system of equations in a form that looks more like what we've seen in linear algebra so far. That system in 1 is equal to:

$$\begin{bmatrix} 1 & 2 \\ 1 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 5 \\ 13 \end{bmatrix} \tag{2}$$

notice that we collected the unknown variables into their own vector, all of the coefficients into a matrix and the right hand side of the equal signs into its own vector. We'll write this same system more compactly as:

$$X\mathbf{w} = \mathbf{y} \tag{3}$$

where $X$ is the matrix, $\mathbf{w}$ is the vector of unknowns, and $\mathbf{y} = [5, 13]$.

### 1.1 Inverses

One way to solve the above is to perform row reduction and use the final result as what is set equal to the unknowns. I won't go into detail on row-reduction but if you want to learn it or want a refresher on it, see *this lecture*. In essence, the point of row reduction is to try and multiply $X$ by it's inverse:

$$X^{-1} \tag{4}$$

doing so will leave the identity matrix on the left hand side and produce the equation:

$$I_2\mathbf{w} = X^{-1}\mathbf{y} \tag{5}$$

notice that $I_2\mathbf{w} = \mathbf{w}$, and that $X^{-1}\mathbf{y}$ will result in some vector $\mathbf{b}$ so we are left with:

$$\begin{aligned} &\mathbf{w} = \mathbf{b} \\ &\mathbf{w} = [1, 2] = \mathbf{b} = X^{-1}\mathbf{y} \end{aligned} \tag{6}$$

As you've seen, an inverse will be defined as that same matrix raised to the negative one power. **However, this is only notation! You can't actually raise every element in a matrix to the negative one power and arrive at the inverse of that matrix**. In regular algebra the inverse of 5 is 1/5 and multiplying them gives you back the number one. Similarly in linear algebra, the inverse of a matrix $A$ is $A^{-1}$, and multiplying $A$ and $A^{-1}$ will result in the matrix with the same meaning as the number one, the identity: $I_n$ (the subscript $n$ denotes the appropriate dimension for the identity matrix)

# 2 Spaces and transformations

Think of the matrix $I_2$, it defines the standard space that we are use to looking at when we look at 2-D plots with an $x$-axis and $y$-axis. The two columns of $I_2$, [1, 0] and [0, 1] together are a set of vectors that create what we call a basis for 2-D space. Formally, a **basis** is a set of vectors that is linearly independent from one another (e.g. no pair of vectors is on the same line), and that **span** the space. For a set of vectors to span a space means that using only those vectors, we can write any other vector in that space as a linear combination of those two vectors. Example:

The vector [2, 6] is defined by the linear combination of:

$$2 \times \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 6 \times \begin{bmatrix} 0 \\ 1 \end{bmatrix} \tag{7}$$

which can also be written as:

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 6 \end{bmatrix} \tag{8}$$

now remember that any vector times the identity results in the same vector, this hopefully lets you see that you could get to any point in 2-D space by linear combinations of the identity matrix, and it is exactly because you can do this that the identity forms a basis for the space. There are however, an infinite amount of bases for 2-D space, for example:

$$\begin{bmatrix} 1 & 5 \\ 3 & 2 \end{bmatrix} \tag{9}$$

also defines a basis for 2-D space.

## 2.1 Columns and rows as spaces

A vector with $N$ elements lives in $N$-dimensional space. For example a vector with 2 elements is a point on a line stretching forward and back forever. A vector with 3 elements is a point in 3-dimensional space and so on. What if we have a matrix? If we have a 2-by-2 matrix, then the number of rows tell us the space each vector lives in, and the number of columns tell us the space that linear combinations of the columns span. For instance, two columns tells us we can get anywhere on a plane by linear combinations of the columns, assuming they are independent.

## 2.2 Rank

The above brings us to the definition of **rank**, which is defined as the dimension that a matrix spans. Sometimes people distinguish between row rank and column rank for matrices, however, they are the same number so we'll just say rank. This last fact is important, if a matrix has columns that look independent, but you notice that any pair of rows is dependent, then consequently this reduces the rank of the columns as well.

Example: A matrix with 3 rows and 2 columns can have at most rank 2:

$$X = \begin{bmatrix} 1 & 2 \\ 4 & 3 \\ 5 & 1 \end{bmatrix} \tag{10}$$

has rank 2 because the third row can be written as a linear combination of the first 2 rows. The process for figuring this out can be a little cumbersome, but if you want to work it out yourself you can use reduced row echelon form to have a clear picture. To be clear, I'm skipping steps here but if we define a new row 2 of $X$, $R_2^*$, as the linear combination:

$$R_2^* = \frac{\Big((4 \times R_1) - R_2\Big)}{5}$$

(note that $R_1$ is the first row of $X$, $R_2$ is the second row), and then a new row 1 as:

$$R_1^* = -2 \times R_2^* + R_1$$

then the third row is equal to $(5 \times R_1^*) + R_2^*$. Because all of the operations have only been linear combinations of the first 2 rows, we see that the third row can be written as a linear combination of the first 2, and so it is redundant, thus reducing the rank. If you want to see this play out in Julia you can copy the following into your REPL:

```
X = [1 2; 4 3; 5 1]
R2 = (4 * X[1, :] - X[2, :]) / 5;
R1 = -2 * R2 + X[1, :];
R3 = 5 * R1 + R2;
```

Example: A matrix with 2 rows and 3 columns has at most rank 2:

$$X = \begin{bmatrix} 1 & 2 & 5 \\ 4 & 3 & 1 \end{bmatrix} \tag{11}$$

by following the same logic as seen above we will arrive at a point where one of the columns will be a linear combination of the other 2, thus resulting in the rank being 2 instead of 3. What does this mean geometrically? If we plot the first two columns as vectors, then we can create a third vector that is a linear combination of them and is equal to the last column in the matrix. Demonstrating that the third column is already in our span by using only the first two columns, (i.e. we can already plot that last vector: [5, 1]) withouth needing it explicity in our set of columns.

## 2.3 Basis

Although we have already defined basis as a set of linearly independent vectors that span a particular space. In addition to having this idea in mind, it's important to have a conceptual intuition for what a basis does. Viewed a particular way, the basis of $I_2$ for 2-D space defines the lengths of the units. That is, to get from 1 to 2 on the x-axis you take a step of 1, precisely because the first column vector is $[1, 0]$. To get from 1 to 3, you take 3 steps of 1: e.g. the vector $[3, 0]$ is just $3 \times [1, 0]$. To reiterate, if you have a new matrix with a set of linearly independent columns that span a space, you can view that matrix as defining a new coordinate space by which to move in, or to measure things in.

## 2.4 Matrix - vector multiplication as mapping into a new coordinate space

With the ideas outlined above in mind, we can see that if we have a matrix $X$ of dimensions $N$ by $P$, and a vector $\mathbf{v}$ with $P$ elements in it, then the operation $\mathbf{w} = X\mathbf{v}$ can be though of $X$ taking in $\mathbf{v}$, and taking you to where in the space of $X$ that vector $\mathbf{v}$ is, which we now call $\mathbf{w}$.