EIGENVALUES, EIGENVECTORS

1. Vector spaces

No study of Linear Algebra would be complete without an understanding of vector spaces. We already saw that matrices can be seen as a set of column vectors and the system of linear equations can be broken down as a set of linear combinations of these columns from Week 1. Now, let's take it further and ask questions on the space of solutions that can exist. Given a matrix A and a constraint vector b, we know how to find the solution, if it exists. Now, can we ask the question that for a matrix A, what is the space of all possible solutions? What kinds of limitations does A impose on the set of possible solutions? Does it impose any constraints on the possible values of x, even when we don't know any specific b vector?

Let's begin by defining some common vector spaces: $\mathbf{R^1}, \mathbf{R^2}, \mathbf{R^3}, \ldots, \mathbf{R^n}$. By R^n, we mean an n-dimensional column-vector taking real-numbered values. For instance $\mathbf{R^3}$ contains the space of all 3-dimensional real-valued vectors, or the space we live in, if the 3 values represent the x, y, and z co-ordinates!

There are two fundamental operations on vectors in $\mathbf{R^n}$ that we should pay attention to. One is multiplying with a scalar. When any vector in $\mathbf{R^n}$ is multiplied by a scalar, the result stays in $\mathbf{R^n}$. Similarly, adding two vectors in $\mathbf{R^n}$ results in a vector that is in $\mathbf{R^n}$. Further, for every $\mathbf{R^n}$, there is a unique $\mathbf{0^n}$, or zero-vector which guarantees $\mathbf{x} + \mathbf{0} = \mathbf{x}$ in $\mathbf{R^n}$. In general, we can lay out 8 properties that vector spaces must satisfy:

(1) $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$
(2) $(\mathbf{x} + \mathbf{y}) + \mathbf{z} = \mathbf{x} + (\mathbf{y} + \mathbf{z})$
(3) $\mathbf{0} + \mathbf{y} = \mathbf{y}$. There exists a $\mathbf{0}$ vector in every space.
(4) For every $\mathbf{x}$, there is a unique vector $-\mathbf{x}$ such that $\mathbf{x} + (-\mathbf{x}) = \mathbf{0}$.
(5) 1 times $\mathbf{x}$ is $\mathbf{x}$.
(6) $(c1c2)\mathbf{x} = c1(c2\mathbf{x})$
(7) $c(\mathbf{x} + \mathbf{y}) = c\mathbf{x} + c\mathbf{y}$
(8) $(c1 + c2)\mathbf{x} = c1\mathbf{x} + c2\mathbf{x}$

We'll define a vector space as a set of all vectors together with these above properties for vector addition and multiplication by real numbers. Similar to $\mathbf{R^n}$, we can see that the space of matrices of any given dimensions also satisfy the above 8 properties and therefore are also *vector spaces*. Similarly a plane in $\mathbf{R^3}$ going through the origin is a vector space in its own right. You can easily demonstrate that the above 8 properties hold for any vectors lying in that plane. You may be tempted to consider the plane to be $\mathbf{R^2}$ but it is not. Every vector living in that plane has 3 distinct co-ordinates and therefore it is a plane in $\mathbf{R^3}$, and it is a subspace of $\mathbf{R^3}$ that is distinct from $\mathbf{R^2}$.

2. Column space of A

We saw that in $Ax = b$, we take linear combinations of the columns of A. In general, if A is not invertible, there may be some bs for which there are solutions. The space of all possible bs is the column-space of A. We can arrive at the column-space of A by taking all possible linear combinations of A. Now we can see that for a solution to exist for $Ax = b$, b must be in the column-space of A. The system of equations is solvable if and only if b is in the column-space of A!

We can now define the nullspace of A as the vector space where $Ax = 0$. For non-invertible matrices A, there are non-zero vectors x that satisfy $Ax = 0$ and these define the nullspace of A. For invertible matrices, the only solution for $Ax = 0$ is $x = 0$.

When we deal with matrices A which are not invertible, we will typically encounter two cases: In the first case, the matrix is square but one or more columns can be expressed as a linear combination of other columns. In the second case, the matrix is rectangular and has more columns than rows. If you really think these two cases, you'll see that they are essentially only one case – there are more variables than there are constraints. We can still perform the pivot and elimination procedure but we'll have free variables in addition to the pivot variables. The free variables can take any values. The simplest choices are 1s and 0s as all other solutions can be expressed as linear combinations of these.

3. Rank of a matrix and linear independence

The rank of a matrix is the number of pivots it has. Further, for rectangular matrices the rank has to be no greater than the smaller of the row or column dimension. For invertible (square) matrices, the rank is the same as the dimension of the matrix. For a matrix with a rank r, there are r independent columns and all other columns can be expressed as linear combinations of these r columns.

Let's define a **basis** as a set of independent vectors that span a space. That is, the set of r independent columns that can be used to construct other columns in a rectangular is the **basis** of that matrix since all solutions for $Ax = b$ can be obtained from linear combinations of those r columns.

A sequence of vectors are considered linearly independent if and only if there is no non-zero linear combination of these vectors that will produce the 0 vector. That is, given **x1** and **x2** as two column vectors, the only way we get $c1\mathbf{x1} + c2\mathbf{x2} = 0$ is when $c1 = c2 = 0$, then **x1** and **x2** are considered linearly independent.

4. Determinants

The determinant of a square matrix is a scalar. When the determinant is zero, we know that the matrix is not invertible.

Here are some properties of determinants

(1) The determinant is a real number.
(2) The determinant can be a negative number.
(3) The determinant only exists for square matrices.

(4) The inverse of a matrix will exist only if the determinant is not zero.

We can define the determinant for a 2x2 matrix as follows:

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \tag{1}$$

det(A) $= ad - bc$. For a 3x3 matrix and above, the determinant is defined in terms of the determinants of their sub-matrices. For

$$A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \tag{2}$$

we define det(A) as a x det($\begin{bmatrix} e & f \\ h & i \end{bmatrix}$) - b x det($\begin{bmatrix} d & f \\ g & i \end{bmatrix}$) + c x det($\begin{bmatrix} d & e \\ g & h \end{bmatrix}$). Notice how we alternate + and - between odd and even columns at the top. Also note, when we consider any element in the first row (e.g. element a), we remove that row and column and compute the determinant of the smaller matrix with that row and column removed. This pattern is repeated recursively for determinants of matrices larger than 3x3. The procedure is to take the element of the first row, one by one and compute the determinant of the sub-matrix with that row and column removed and then multiply by that element and take the appropriate sign (+ for odd and - for even numbered columns).

If $det(A) = 0$, then A is singular. If $det(A) \neq 0$ then A is invertible.

5. Eigenvalues and Eigenvectors

When you compute Ax, you are multiplying the matrix A onto vector x and this typically causes x to change its direction. However, there are some special vectors x which lie in the same direction as Ax. In other words, $Ax = \lambda x$ where λ is a scalar. If we are lucky enough to find such vectors, we call them *eigenvectors* and the corresponding λ as *eigenvalues*.

To find eigenvalues and eigenvectors, we take the determinant of a special matrix and set it to zero: $det(A - \lambda I) = 0$. This gives rise to a polynomial equation in terms of λ called the *characteristic polynomial*. The roots of this polynomial equation are the eigenvalues. Once we have the eigenvalues, we can solve for $(A - \lambda I)x = 0$ and find the eigenvectors x.

Eigenvalues and Eigenvectors play an important role in many calculations. A very useful technique called the Principal Components Analysis is frequently used in Machine Learning contexts and it relies on Eigenvalues and Eigenvectors. We'll see how PCA is used to solve regression and classification problems later. The PageRank algorithm at the heart of the Google search engine uses Eigenvalues to help rank webpages. We'll visit PageRank when we cover probabilities and statistics later in this course.

Let's compute the eigenvalues and eigenvectors for a 2x2 matrix. Let

$$A = \begin{bmatrix} 1 & 4 \\ 3 & 2 \end{bmatrix} \tag{3}$$

To get the eigenvalues, we need to solve

$$det(\begin{bmatrix} 1 & 4 \\ 3 & 2 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}) = 0 \tag{4}$$

That is,

$$det(\begin{bmatrix} 1-\lambda & 4 \\ 3 & 2-\lambda \end{bmatrix}) = 0 \tag{5}$$

which reduces to

$$(1-\lambda) \times (2-\lambda) - 4 \times 3 = 0 \tag{6}$$

Which gives us the characteristic polynomial as

$$\lambda^2 - 3\lambda - 10 = 0 \tag{7}$$

which factorizes as

$$(\lambda - 5) \times (\lambda + 2) = 0 \tag{8}$$

Which gives us two eigenvalues as $\lambda_1 = 5$ and $\lambda_2 = -2$. When you use λ_1 in the definition of eigenvalue, you get

$$Av - \lambda v = (A - \lambda I)v = 0 \tag{9}$$

which reduces to

$$(\begin{bmatrix} 1 & 4 \\ 3 & 2 \end{bmatrix} - \begin{bmatrix} 5 & 0 \\ 0 & 5 \end{bmatrix}) \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = 0 \tag{10}$$

or

$$(\begin{bmatrix} -4 & 4 \\ 3 & -3 \end{bmatrix}) \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = 0 \tag{11}$$

This is just a system of equations problem. If you take the first row, you get

$$-4v_1 + 4v_2 = 0 \tag{12}$$

or

$$v_1 = v_2 \tag{13}$$

Which is the equation for a line in the 2-D plane. This line has a slope of 1 and passes through the origin. So, $v = (1, 1)$ is an eigenvector. Customarily, we write unit vectors are eigenvectors. So, we can instead write $v = (0.7071, 0.7071)$ as the first eigenvector. The eigenspace corresponding to the first eigenvalue is the line that passes through the origin at a 45 degree angle.

We can similarly solve for the second eigenvalue and we get an equation for a line again. This time the equation of the line is $-3v_1 + 4v_2 = 0$. So, all the eigenvectors corresponding to the second eigenvalue lie along this line.