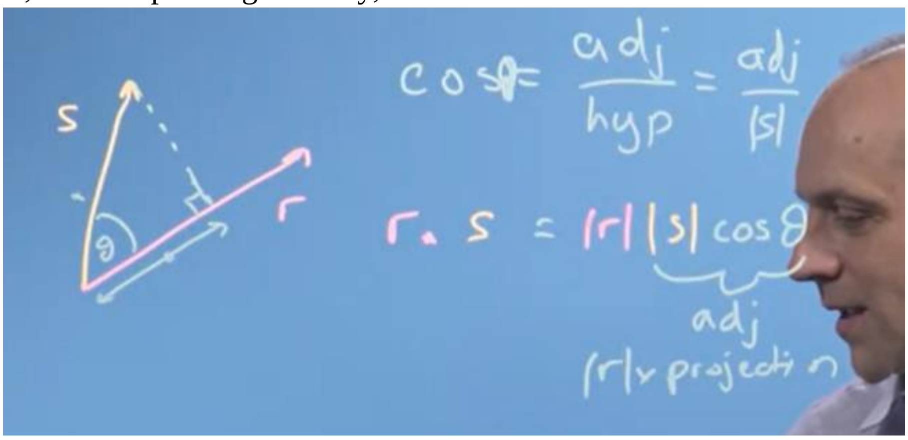
Mathematics-for-Machine-Learning

1. Getting handle on vectors:
   a. Definition:
      i. Vectors are usually viewed by computers as an ordered list of numbers which they can perform "operations" on - some operations are very natural and, as we will see, very useful!
      ii. A vector in space-time can be described using 3 dimensions of space and 1 dimension of time according to some co-ordinate system.
      iii. Vectors can be thought of in a variety of different ways - some geometrically, some algebraically, some numerically. In this way, there are a lot of techniques one can use to deal with vectors.
   b. Vector is just something that are based on two rules. Firstly, addition, and secondly, multiplication by a scalar number.
      i. Addition:
         1. $r = \begin{bmatrix} 3 \\ 2 \end{bmatrix}, s = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$
         2. $r + s = s + r(associativity) = \begin{bmatrix} 3-1 \\ 2+2 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$
      ii. Multiplication:
         1. $r = \begin{bmatrix} 3 \\ 2 \end{bmatrix},$
         2. $2r = \begin{bmatrix} 3 \cdot 2 \\ 2 \cdot 2 \end{bmatrix} = \begin{bmatrix} 6 \\ 4 \end{bmatrix}$
2. Finding the size of vectors, its size and projection
   a. $r = \begin{bmatrix} 3 \\ 2 \end{bmatrix},\ s = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$
   b. Size of vectors = $|r| = \sqrt{3^2 + 2^2}$ = $\sqrt{9+4}$
   c. Dot product of vectors:
      i. $r \cdot s = s \cdot r = r_{11} \cdot s_{11} + r_{21} \cdot s_{21} = 3 \cdot (-1) + 2 \cdot 2 = 1$
   d. Cosine rule = $r \cdot s = |r| \cdot |s| \cdot \cos\theta$
   e. Projection = dot product gives us, is it gives us the projection here of S on to R times the size of R. And one thing to notice here is that if S was perpendicular to R, if S was pointing this way, it would have no shadow.



      i. Vector projection = $\frac{r \cdot s}{r \cdot r} \cdot r$
      ii. Scalar projection = $\frac{r \cdot s}{r}$
3. Changing basis:

a. Any vector space has multiple bases, so the question naturally arises: what are the relationships between bases of a vector space? In the first place, there must be the same number of elements in any basis of a vector space. Then, given two bases of a vector space, there is a way to translate vectors in terms of one basis into terms of the other; this is known as change of basis.
Change of basis is a technique applied to finite-dimensional vector spaces in order to rewrite vectors in terms of a different set of basis elements. It is useful for many types of matrix computations in linear algebra and can be viewed as a type of linear transformation.

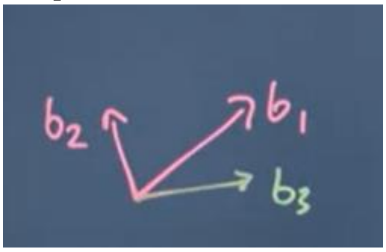
b. $v = \begin{bmatrix} 10 \\ -5 \end{bmatrix}, b_1 = \begin{bmatrix} 3 \\ 4 \end{bmatrix}, b_2 = \begin{bmatrix} 4 \\ -3 \end{bmatrix}$

c. 1) $\frac{v \cdot b_1}{|b_1|^2} = \frac{2}{5}$

d. 2) $\frac{v \cdot b_2}{|b_2|^2} = \frac{11}{5}$

e. $v_b = \begin{bmatrix} \frac{2}{5} \\ \frac{11}{5} \end{bmatrix}$

4. Basis, vector space and linear independence
   a. Basis is a set pf n vectors that:
      i. Are not linear combination of each other(Linear independent)
      ii. Span the space
      iii. The space is n-dimensional
   b. Linear independence:

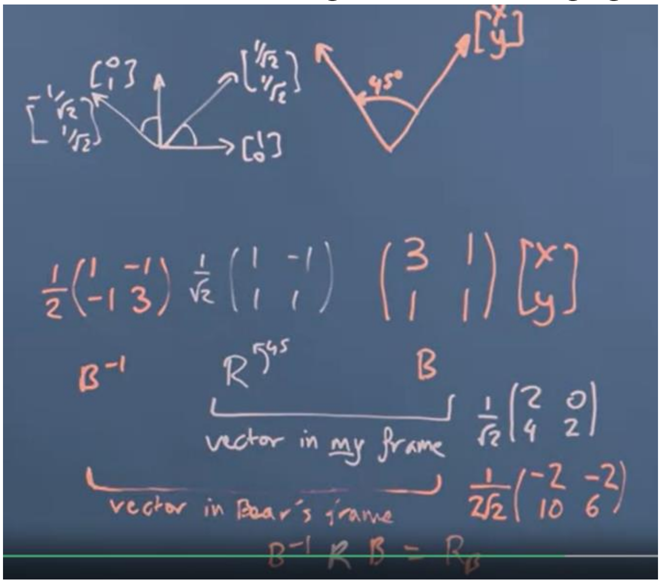


      i. $b_3 \neq a_1 b_1 + a_2 b_2$
      ii. b_3 *does not lie in the plane spanned by* b_1 *and* b_2
      iii. if detM≠0, the given vectors are linearly independent

5. Gaussian elimination
   a. Gaussian elimination, also known as row reduction, is an algorithm in linear algebra for solving a system of linear equations. It is usually understood as a sequence of operations performed on the corresponding matrix of coefficients.

$$\left[\begin{array}{ccc|c} 1 & 3 & 1 & 9 \\ 1 & 1 & -1 & 1 \\ 3 & 11 & 5 & 35 \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & 3 & 1 & 9 \\ 0 & -2 & -2 & -8 \\ 0 & 2 & 2 & 8 \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & 3 & 1 & 9 \\ 0 & -2 & -2 & -8 \\ 0 & 0 & 0 & 0 \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & 0 & -2 & -3 \\ 0 & 1 & 1 & 4 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

6. Transformation in a changed basis and changing basis:



7. Orthogonal matrices
   a. $A^T_{ij} = A_{ji}$, $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$, $A^T = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}$
   b. Orthogonal matrixes –
      i. $Q^T \cdot Q = I$
      ii. $Q^T = Q^{-1}$
   c. A set of unit length basis vectors that are all perpendicular to each other are called an orthonormal basis set, and the matrix composed of them is called an orthogonal matrix. $a_i \cdot a_j = 0 \; if \; i \neq j, a_i \cdot a_j = 1 \; if \; i = j$
8. The Gram–Schmidt process

We define the projection operator by

$$\mathrm{proj}_{\mathbf{u}}(\mathbf{v}) = \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\langle \mathbf{u}, \mathbf{u} \rangle}\mathbf{u},$$

where $\langle \mathbf{u}, \mathbf{v} \rangle$ denotes the inner product of the vectors **u** and **v**. This operator projects the vector **v** orthogonally onto the line spanned by vector **u**. If **u** = **0**, we define $\mathrm{proj}_{\mathbf{0}}(\mathbf{v}) := \mathbf{0}$. i.e., the projection map $\mathrm{proj}_{\mathbf{0}}$ is the zero map, sending every vector to the zero vector.

The Gram–Schmidt process then works as follows:

$$\begin{aligned}
\mathbf{u}_1 &= \mathbf{v}_1, & \mathbf{e}_1 &= \frac{\mathbf{u}_1}{\|\mathbf{u}_1\|} \\
\mathbf{u}_2 &= \mathbf{v}_2 - \mathrm{proj}_{\mathbf{u}_1}(\mathbf{v}_2), & \mathbf{e}_2 &= \frac{\mathbf{u}_2}{\|\mathbf{u}_2\|} \\
\mathbf{u}_3 &= \mathbf{v}_3 - \mathrm{proj}_{\mathbf{u}_1}(\mathbf{v}_3) - \mathrm{proj}_{\mathbf{u}_2}(\mathbf{v}_3), & \mathbf{e}_3 &= \frac{\mathbf{u}_3}{\|\mathbf{u}_3\|} \\
\mathbf{u}_4 &= \mathbf{v}_4 - \mathrm{proj}_{\mathbf{u}_1}(\mathbf{v}_4) - \mathrm{proj}_{\mathbf{u}_2}(\mathbf{v}_4) - \mathrm{proj}_{\mathbf{u}_3}(\mathbf{v}_4), & \mathbf{e}_4 &= \frac{\mathbf{u}_4}{\|\mathbf{u}_4\|} \\
&\;\;\vdots & &\;\;\vdots \\
\mathbf{u}_k &= \mathbf{v}_k - \sum_{j=1}^{k-1} \mathrm{proj}_{\mathbf{u}_j}(\mathbf{v}_k), & \mathbf{e}_k &= \frac{\mathbf{u}_k}{\|\mathbf{u}_k\|}.
\end{aligned}$$

   a. Gram–Schmidt process is a method for orthonormalizing a set of vectors in an inner product space, most commonly the Euclidean space
9. Transformation matrix for reflecting vectors in an arbitrarily angled mirror, construct an orthonormal basis that spans a set of input vectors $T = ET_E E^{-1}$, $r^{-1} = T \cdot r$

10. Eigenvalues and Eigenvectors:

a. In linear algebra, an **eigenvector** or **characteristic vector** of a linear transformation is a nonzero vector that changes by a scalar factor when that linear transformation is applied to it. The corresponding eigenvalue, often denoted by λ is the factor by which the eigenvector is scaled.

b. How to find eigenvectors and eigenvalues

   i. $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$

   ii. $Ax = \lambda x; (A - \lambda I)x = 0$

   iii. $det\left(\begin{bmatrix} -\lambda & -1 \\ 1 & -\lambda \end{bmatrix}\right) = \lambda^2 + 1; \lambda = 1 \ and \ \lambda = -1$

   iv. $@\lambda = 1$

      1. $\begin{bmatrix} -1 & -1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

      2. $\begin{cases} -x_1 - x_2 = 0 \\ x_1 - x_2 = 0 \end{cases}$

      3. $\begin{cases} x_2 = 0 \\ x_1 = 0 \end{cases}$

   v. $@\lambda = -1$

      1. $\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

      2. $\begin{cases} x_1 - x_2 = 0 \\ x_1 + x_2 = 0 \end{cases}$

      3. $\begin{cases} x_1 = x_2 \\ x_2 = 0 \end{cases}$

c. To find T^n we use diagonalization for ex:

   i. $T = \begin{bmatrix} 1 & 0 \\ 2 & -1 \end{bmatrix}$

   ii. The find eigenvectors:

      1. $(1 - \lambda)(-1 - \lambda) = 0$

      2. $\lambda = 1 \ and \ \lambda = -1$

      3. $@\lambda = 1$

         a. $\{x_1 = x_2$

         b. $v_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

      4. $@\lambda = -1$

         a. $\{x_1 = 0$

         b. $v_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$

   iii. $T^n = C \cdot D \cdot C^{-1}$

   iv. $C = [v_1, v_2] = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$

   v. $C^{-1} = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}$

   vi. $D = \begin{bmatrix} \lambda_1 & \\ & \lambda_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$

   vii. $T^n = C \cdot D^n \cdot C$