Linear algebra

> Reference: https://www.coursera.org/learn/linear-algebra-machine-learning

Vector operations

Basics

$$\vec{r} + \vec{s} = \vec{s} + \vec{r}$$

$$2\vec{r} = \vec{r} + \vec{r}$$

$$\|\vec{r}\|^2 = \sum_i r_i$$

Dot or inner product

$$\vec{r}.\vec{s} = \sum_i r_i s_i$$

- *commutative:* $\vec{r}.\vec{s} = \vec{s}.\vec{r}$
- *distributive:* $\vec{r}.(\vec{s} + \vec{t}) = \vec{r}.\vec{s} + \vec{r}.\vec{t}$
- *associative:* $\vec{r}.(a\vec{s}) = a(\vec{r}.\vec{s})$

$$\vec{r}.\vec{r} = \|\vec{r}\|^2$$

$$\vec{r}.\vec{s} = \|\vec{r}\|\|\vec{s}\|cos\theta$$

Scalar and vector projection

- scalar projection ($\vec{s}$ on $\vec{r}$): $\frac{\vec{r}.\vec{s}}{\|\vec{r}\|}$
- vector projection: $\frac{\vec{r}.\vec{s}}{\vec{r}.\vec{r}}\vec{r}$

Orthonormal vectors

$$\vec{e_i}.\vec{e_j} = 0 \text{ (i.e. orthogonal)}$$

$$\vec{e_i}.\vec{e_i} = 1 \text{ (i.e. unit size)}$$

Basis

A **basis** is a set of *n* vectors that:

- are not linear combinations of each other
- span the space

The space is then n-dimensional.

Matrices

Basics

$$A\vec{r} = \vec{r'}$$

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} e \\ f \end{bmatrix} = \begin{bmatrix} ae + bf \\ ce + df \end{bmatrix}$$

- *associative:* $A(n\vec{r}) = n(A\vec{r}) = n\vec{r'}$
- *distributive:* $A(\vec{r} + \vec{s}) = A\vec{r} + A\vec{s}$
- *not commutative*: $AB \neq BA$

$$\text{Identity:} \quad I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\text{Clockwise rotation by θ:} \quad R = \begin{bmatrix} cos\ \theta & sin\ \theta \\ -sin\ \theta & cos\ \theta \end{bmatrix}$$

$$\text{Determinant of 2x2 matrix:} \quad det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc$$

$$\text{Inverse of 2x2 matrix:} \quad \begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

Einstein's Summation Convention

For multiplying matrices a and b:

$$ab_{ik} = \sum_j a_{ij} b_{jk}$$

Change of basis

Change from an original basis to a new, primed basis.

The columns of the transformation matrix B are the *new basis vectors in the original coordinate system*. So

$$B\vec{r'} = \vec{r}$$

where r' is the vector in the B-basis, and r is the vector in the original basis. Or

$$\vec{r'} = B^{-1}\vec{r}$$

If a matrix A is **orthonormal** (all the columns are of unit size and orthogonal to each other) then the inverse of an orthonormal matrix is the transposed matrix:

$$A^{-1} = A^T$$

and

$$A^T A = I \text{ (identity matrix)}$$

Orthonormal basis vector set

- Vectors are **linearly independent** if the determinant of the matrix (having these vectors in columns) != 0
- Vectors are **orthogonal** is their dot product is 0.
- **Orthonormal matrix** = all columns (vectors) are *orthogonal* and of *unit size*

Gram-Schmidt process for constructing an orthonormal basis

How to build an *orthonormal basis vector set* from a basis list of vectors? Apply the **Gram-Schmidt process**: https://www.coursera.org/learn/linear-algebra-machine-learning/lecture/28C1t/the-gram-schmidt-process

Start with *n* linearly independent basis vectors $\vec{v} = \{\vec{v_1}, \vec{v_2}, ..., \vec{v_n}\}$. Then

$$\vec{e_1} = \frac{\vec{v_1}}{||\vec{v_1}||}$$

$$\vec{u_2} = \vec{v_2} - (\vec{v_2}.\vec{e_1})\vec{e_1} \text{ and } \vec{e_2} = \frac{\vec{u_2}}{||\vec{u_2}||}$$

... and so on for $\vec{u_3}$ being the remnant part of $\vec{v_3}$ not composed of the preceding e-vectors, etc...

Transformation in a plane or other object

If we want to apply a transform on a vector (e.g. a reflexion in a plane) but the vector basis is not orthonormal (which means complex interpretation and calculation), we can do the following:

- First transform into the basis referred to the reflection plane, or whichever, by applying a Gram-Schmidt transform: E^{-1}
- Then do the reflection, or other transformation, in the plane of the new basis: T_E.
- Then transform back into the original basis: E (orthonormal vectors $\{\vec{e_i}\}$ in original basis).
  So our transformed vector is:

$$\vec{r'} = ET_E E^{-1}\vec{r}$$

Eigenvalues and eigenvectors

An **eigenvector** or **characteristic vector** of a linear transformation A is a nonzero vector $\vec{v}$ that changes at most by a scalar factor λ when that linear transformation is applied to it (A is a square matrix):

$$A\vec{v} = \lambda\vec{v}$$

The corresponding **eigenvalue** λ is the factor by which the eigenvector is scaled.

There might be several eigenvalues for a matrix A. Eigenvalues will satisfy the following condition

$$(A{-}\lambda I)\vec{v} = 0$$

where I is an *n* by *n* dimensional identity matrix.

This equation has a nonzero solution $\vec{v}$ if and only if the determinant of the matrix $(A{-}\lambda I)$ is zero. Therefore, the eigenvalues of A are values of λ that satisfy the equation

$$|A - \lambda I| = 0$$

which can be factored into the product of *n* linear terms,

$$|A - \lambda I| = (\lambda_1 - \lambda)(\lambda_2 - \lambda)\cdots(\lambda_n - \lambda)$$

where each λ_i may be real but in general is a complex number. The numbers $\lambda_1, \lambda_2, ...\lambda_n$, which may not all have distinct values, are roots of the polynomial and are the **eigenvalues of A.**

Geometrically, an *eigenvector*, corresponding to a real nonzero eigenvalue, points in a direction in which it is stretched by the transformation and the *eigenvalue* is the factor by which it is stretched. If the eigenvalue is negative, the direction is reversed. Loosely speaking, in a multidimensional vector space, the eigenvector is not rotated.

> Very cool tool to visualize matrix transformations and their eigenvectors: https://www.coursera.org/learn/linear-algebra-machine-learning/ungradedWidget/AVEfF/visualising-matrices-and-eigen

Diagonalization and the eigendecomposition

> Reference: https://en.wikipedia.org/wiki/Eigendecomposition_of_a_matrix

Suppose the eigenvectors of A form a basis, or equivalently A has *n* **linearly independent eigenvectors** $\vec{v_1}, \vec{v_2}, ..., \vec{v_n}$ with associated eigenvalues $\lambda_1, \lambda_2, ..., \lambda_n$. The eigenvalues need not be distinct. Define a square matrix Q whose columns are the *n* linearly independent eigenvectors of A,

$$Q = \begin{bmatrix} \vec{v_1} & \vec{v_2} & \cdots & \vec{v_n} \end{bmatrix}$$

Since each column of Q is an eigenvector of A, right multiplying A by Q scales each column of Q by its associated eigenvalue,

$$AQ = \begin{bmatrix} \lambda_1 \vec{v_1} & \lambda_2 \vec{v_2} & \cdots & \lambda_n \vec{v_n} \end{bmatrix}$$

With this in mind, define a diagonal matrix Λ where each diagonal element Λ_{ii} is the eigenvalue λ_i associated with the *i*th column of Q. Then

$$AQ = Q\Lambda$$

Because the columns of Q are linearly independent, Q is *invertible*. Right multiplying both sides of the equation by Q^{-1},

$$A = Q\Lambda Q^{-1}$$

or by instead left multiplying both sides by Q^{-1},

$$Q^{-1}AQ = \Lambda$$

A can therefore be decomposed into:

- a matrix composed of its eigenvectors,
- a diagonal matrix with its eigenvalues along the diagonal, and
- the inverse of the matrix of eigenvectors.

This is called the **eigendecomposition** and it is a similarity transformation. Such a matrix A is said to be *similar* to the diagonal matrix Λ or **diagonalizable** (see https://en.wikipedia.org/wiki/Diagonal_matrix). The matrix Q is the change of basis matrix of the similarity transformation. Essentially, the matrices A and Λ represent the same linear transformation expressed in two different bases. The eigenvectors are used as the basis when representing the linear transformation as Λ.

Note that the following equation allows to compute more quickly the matrix exponential of A:

$$A^m = Q\Lambda^m Q^{-1}$$

where

$$\Lambda^m = \begin{bmatrix} \lambda_1^m & 0 & \cdots & 0 \\ 0 & \lambda_2^m & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n^m \end{bmatrix}$$

Moreover, the eigenvalues of A^2 and A^{-1} are λ^2 and λ^{-1}, with the same eigenvectors.