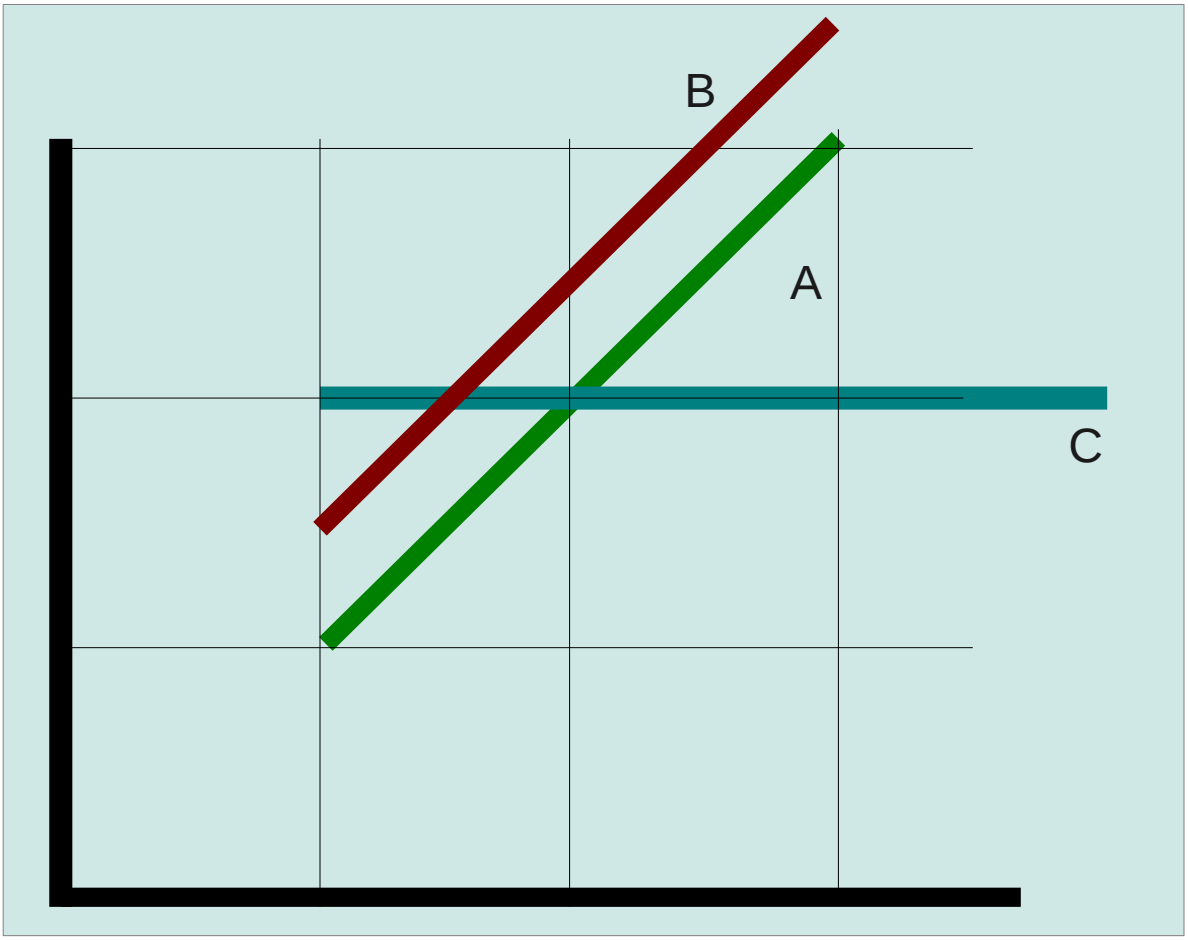
A quick rundown of the process.

Lets see which line (B or C) is more similar to line A:

line A (our “Query”) (1,1 - 2,2 - 3,3) a 45 degree line of three points

line B (1,1.5 - 2,2.5 - 3,3.5) (a 45 degree line slightly higher)
line C (1,2 – 2,2 – 3,2) a flat line on y=2



Each line is turned into a vector, in this case, having three dimensions. Datapoint 1 on the line becomes the first dimension (X), datapoint 2 becomes Y and datapoint 3 becomes Z ... so we have a point in a three dimensional space that represents the line. Long lines could easily have 200+ dimensions.

1. “Relative Magnitude”

Lets see how well our dimensions line up:
In the case of a document, this number acts as a way to increase the likelyhood of documents that have highly correlated words frequencies to appear together.
For each matched dimension to the ref, sum the product of that dimension.

RelMagnitudeAB = sqrt((1*1)+(2*2)+(3*3)) = sqrt(14) = 3.741
RelMagnitudeBA = sqrt((1.5*1.5)+(2.5*2.5)+(3.5*3.5)) = sqrt(2.25+6.25+12.25) = 4.555

RelMagnitudeAC = sqrt((1*1)+(2*2)+(3*3)) = sqrt(14) = 3.741
RelMagnitudeCA = sqrt((2*2)+(2*2)+(2*2)) = sqrt(12) = 3.464

In the algo, I compare the vectors both ways (compare A to B and then B to A) , since we can have wildly differing results. I then multiply them together... so

ActualRelMagnitudeAB = 17.040255
ActualRelMagnitudeAC = 12.958

2. Relativity

Purely based on overlapping dimensionality... each line has the same 3 dimensions. so... 3
RelativityAB = 3
RelativityAC = 3

3. Dot Product

DotProductAB = (1*1.5)+(2*2.5)+(3*3.5) = 17
DotProductAC = (1*2)+(2*2)+(3*2) = 12

4. Result

Our base result... a multiplier on the output that helps eliminate bad matches... since we have equal numbers and designations of dimensions, this will be 1.

DotProductAB / RelMagnitudeAB = 1.3119
DotProductAC / RelMagnitudeAC = 0.9260

5. Score

One final step in text, which won't matter here is I usually multiply this score by the count of dimensions in our “query” vector divided by the number of matching dimensions in the tested vector ... but in this case it would be 1 .. since all vectors have matching dimensions

RESULT here:

A score of 1 means a perfect match
So Line B is actually LESS SIMILAR to A than C in this small section,
mainly due to point 2 being equal.