Find the following, using the formulae in vectors1.pdf, slides 20-24.

1.1. \( p(w = \text{politician}, c = \text{liberal}) \)

1.2. \( p(w = \text{politician}) \)

1.3. \( p(c = \text{liberal}) \)

1.4. Compute the PPMI score for word politician and context liberal.

1.5. Suppose \( P(w = i \mid c = j) \) is equal to \( P(w = i). \) What can we say about \( \text{PPMI}(w=i,c=j)? \) If you don’t remember the discussion of this case in class, use the chain rule to turn this into a fact about \( P(i,j). \)

1.6. Is it possible for the PMI value of target word \( i \) and context word \( j \) to be negative?

1.7. When is PPMI undefined?
2 Cosine similarity

Use slides 33-37 to help with the following.

2.1. Using the counts (rather than the PPMI values), compute the cosine similarity of target words conservative and politician. Note: It should be a number between 0 and 1.

2.2. Same two words: Now compute the cosine similarity using vectors with PPMI values. Note: For this problem, think of the log of a probability of 0 as a negative number with a very high absolute value. So for the purposes of PPMI any 0 probabilities are going to yield a PPMI of 0. Show at least this much of your work: What are the two PPMI vectors? What are the vectors after they are divided by their length? (We say the have been normalized).