

Name: Adarsh Srivastava

Email: [e0954759@u.nus.edu](mailto:e0954759@u.nus.edu)

Student ID: A0254358X

### Function definitions:

1. **e\_step**: In this step, we recursively calculate alpha and beta using the forward backward algorithm while using a scaling factor. The emission probabilities are assumed to come from K Gaussian distributions, whose params (mu and sigma) are updated with each step. Once we get the scaled alpha and beta, we calculate the gamma and xi for the sequence and add them to the list to be returned eventually.
2. **calculate\_alpha**: We use 0-indexing, and use the definitions on slides to calculate the values of alpha for each timestamp, with special treatment to the first value. After every calculation, we normalize the value and store the normalizing factor in c, to be used later. Finally, the function is called recursively until the last timestamp is reached.
3. **calculate\_beta**: Similar to *calculate\_alpha*, this function uses the formula on slides to calculate beta in the backward pass, initializing last Betas to 1 and using the previously stored normalizing constant C to normalize the values. The function is called recursively until beta for the first timestep is calculated.
4. **calculate\_xi**: This is again a simple element wise calculation of xi for each timestamp. Here too we use the pre calculated normalizing constant to normalize the value.
5. **m\_step**: Here too we have used the formulae on slides to calculate the value - with one difference. Since we have a list of sequences instead of a single sequence in each iteration, we also sum each numerator and denominator over the list of sequences. This is equivalent to calculating values for the concatenated sequence.
6. **fit\_hmm**: Finally, we use the given initialization function to calculate initial values of params. Using these, we start the iterations of alternating e-step, m-step, and updating the params. We do this until the largest change in any single value of the params is lower than an  $\epsilon = 1e - 3$ , upon which the iteration terminates.