Report 3

Lu, Yiwen A0225573A

```
pro_x_z = calculate_pro_x_z(n_states, input, phi)
```

This function is aiming to calculate $P(X|Z,\theta)$ in batch. scipy.stats.norm.pdf() is used to calculate Gaussian function.

```
alpha_list, C_n, pro_x_z = forward_pass(x_list, pi, A, phi)
```

This function is used to calculate alpha list in e-step. Also, I record C_n and $P(X|Z)$ for m-step. This could help to save time.

```
beta_list = backward_pass(x_list, pi, A, phi, C_n, pro_x_z)
```

This function is used to calculate beta list in m-step. In this step beta_list[-1] is initiate as 1.

```
gamma_list, xi_list = e_step(x_list, pi, A, phi)
```

This function is used to update $\gamma(z_n)$ and $\xi(z_{n-1}, z_n)$.

```
pi, A, phi = m_step(x_list, gamma_list, xi_list)
```

This function is used to update π, A and $\theta = \mu, \sigma$

```
pi, A, phi = fit_hmm(x_list, n_states)
```

This function is combined with iterated e-step and m-step. Until $|\theta_{old} - \theta| < 1e-4$, the loop is over.

```
Running on seq_short
\---------------------
Loaded 200 sequences, with average length = 8.0
Groundtruth pi:
 [0.9 0.1 0. ]
Groundtruth A:
 [[0.3 0.7 0. ]
 [0.  0.6 0.4]
 [0.  0.  1. ]]
Groundtruth phi:
 {'mu': array([-1.5,  0.5, -0.2]), 'sigma': array([0.5, 0.2, 0.3])}
Your pi:
 [0.09 0.91 0.  ]
Your A:
 [[0.6  0.   0.4 ]
 [0.72 0.28 0.  ]
 [0.   0.   1.  ]]
Your phi:
 {'mu': array([ 0.52, -1.5 , -0.21]), 'sigma': array([0.21, 0.51, 0.29])}


Running on seq_long
\---------------------
Loaded 5 sequences, with average length = 1000.0
Groundtruth pi:
 [0.6 0.  0.  0.4]
Groundtruth A:
 [[0.3 0.  0.  0.7]
 [0.7 0.3 0.  0. ]
 [0.  0.6 0.4 0. ]
 [0.  0.  0.4 0.6]]
Groundtruth phi:
 {'mu': array([0. , 1.1, 2. , 1. ]), 'sigma': array([0.5, 0.3, 0.4, 0.2])}
Your pi:
 [0.   0.64 0.   0.36]
Your A:
 [[0.3  0.7  0.   0.  ]
 [0.   0.3  0.   0.7 ]
 [0.58 0.   0.42 0.  ]
 [0.   0.   0.43 0.57]]
Your phi:
 {'mu': array([1.07, 0.  , 2.  , 1.  ]), 'sigma': array([0.31, 0.52, 0.41, 0.19])}
```