Analysis
Homework 6: Variational Inference by Rohan Chandra.
USING 1 LATE DAY.

new_phi:

In new_phi, a new value for the variational parameter phi is calculated. This is accomplished by looking at the appropriate column of the Beta matrix for the given word index. The exponential digamma terms are calculated for the sum of over the gamma vectors and for the vector of gammas. The exponential term is calculated as a vector and is element-wise multiplied by the extracted Beta column. Then the normalization term is calculated as the sum of the unscaled phi result divided by the count term to result in the proper scaling result. The phi vector is returned. All computations are done using numpy calls to improve computation speed.

m_step:

In the m_step function, the new Beta matrix is calculated. In the e_step function call, topic_counts is computed based on the final phi computed in the inner loop of the document iteration loop. This set of counts is passed to the m_step function as an input. Then we know that the Beta columns can be approximated by taking the counts of the phi matrix represented in topic_counts and for each topic, dividing by the total counts. The comments in the code gives us that each topic is a row, so we simply take the sum over the rows then scale each row by one over that relevant sum. The row sums are taken using a numpy call on the topic_counts matrix. The scaling is done in a loop. A check is done to prevent divides by zero. The resulting matrix is returned.

Results:

- The resulting topics.txt list was generated using the default code settings and running the lda script against the train.dat and voc.dat datasets in ../data/ap.

- A successful code would generate clear topics with words that clearly belong in that topic and make sense. For example, "Money" topic would have words like bank, dollar, cent, million and so on. "Law" topic would contain words like attorney, judge, prison, trial etc.

- **My generated list of topic is successful. Topic 2 clearly matches with Law. Topic 6 clearly matches with Money. And topic 3 somewhat deals with buildings like police station, hospital, schools, etc.**

Extra Credit

- α is the dirichlet parameter.

- By differentiating the ELBO objective function w.r.t α,

$$\mathbb{F} = \frac{\partial \mathcal{L}}{\partial \alpha} = \psi(\sum_i \alpha_i) - \psi(\alpha) + \psi(\gamma) - \psi(\sum_i \gamma_i)$$

- To compute optimal α, we set $\mathbb{F} = 0$. We approximate this by Newton's method.

- At each iteration, we compute:

$$\alpha_{new} = \alpha_{old} - \frac{\mathbb{F}}{\partial \mathbb{F}/\partial \alpha}$$