Report: Experiments on α updates

Arijus Pleska

This report assesses some experiments performed on reproducing α parameters used in generating synthetic data. Note that the report is structured in the following order: 1) defining the experiment settings; 2) assessing the experiment results.

The Experiment Settings

The intention of the carried experiments is to identify the optimal settings for the Metropolis–Hastings algorithm application. To start with, I have generated a synthetic corpus; the parameters used in the corpus generation will allow to assess the performance achieved in the experiments. The corpus generation parameters are set as follows:

- The number of topics: K = 2;
- The number of documents (time-slices): T = 20;
- The size of vocabulary: V = 10;
- The number of words per document t: $N_t \sim \text{Pois}(\lambda), \quad \lambda = 1000.$

Further, to consider the initial settings of α_k development over documents, α_0 is a sine curve and α_1 is a cosine curve; the corresponding softmax expressions of the curves, i.e $\mu = \text{softmax}(\alpha)$, are illustrated in Figure 1 below.

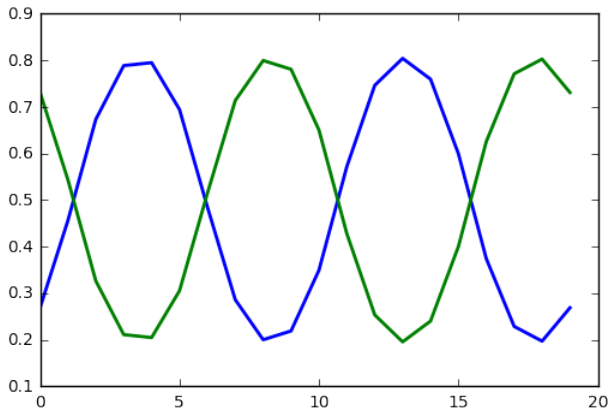


Figure 1: The values of softmax(α) used in the generative process.

Speaking of β, it was initially predefined and kept constant throughout the dynamic generative process; β is illustrated in Figure 2 below.

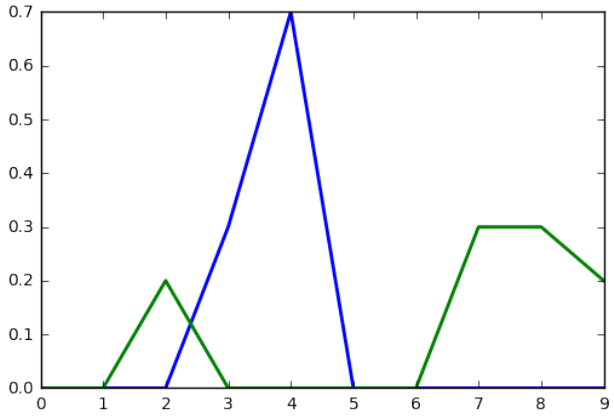


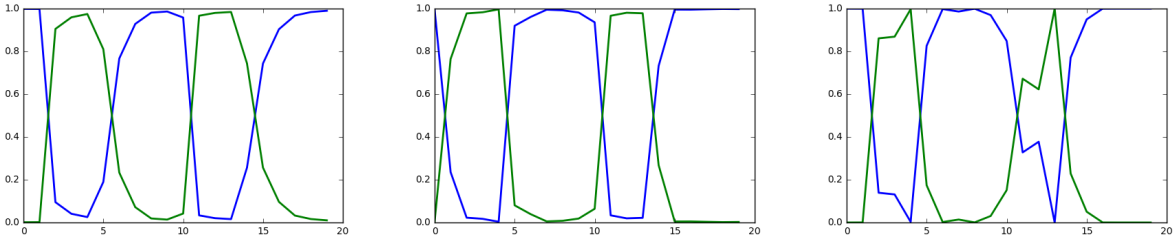
Figure 2: The values of β used in the generative process.

Note that the latter β values were applied to the autoregressive topic model for the α update experiments.

The experiment results

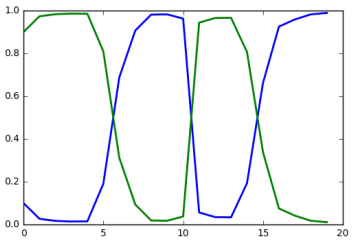
The first experiment is focused on discovering the choice of the variances. To be more specific, the alpha update is based on three different variances were used: the 'initial' variance $\sigma_0^2 I$ to induce α_t at $t = 0$, the 'basic' variance $\sigma^2 I$ to induce α_t at $t > 0$, and the 'proposed' variance $\delta^2 I$ to induce α'_t at $t = 0$; also, note that α'_t at $t > 0$ were induced using the 'basic' variance.

For both experiments, the number of autoregressive iterations was set to 500 and σ^2 was set to 0.01. The resulting plots of μ with different values of σ_0^2 and δ^2 are illustrated in Figures 3, 4, 5 below.

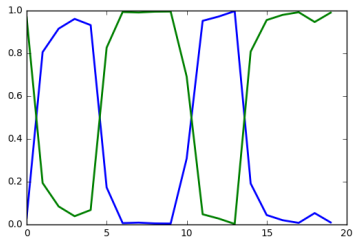


(a) $\sigma_0^2 = 1, \quad \delta^2 = 0.5$ (b) $\sigma_0^2 = 1, \quad \delta^2 = 2$ (c) $\sigma_0^2 = 1, \quad \delta^2 = 5$

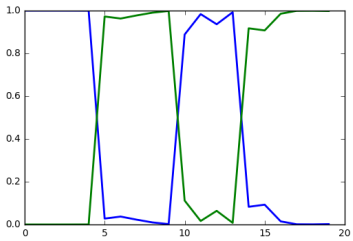
Figure 3: softmax(α) with the lowest value of the initial variance.



(a) $\sigma_0^2 = 5, \quad \delta^2 = 0.5$

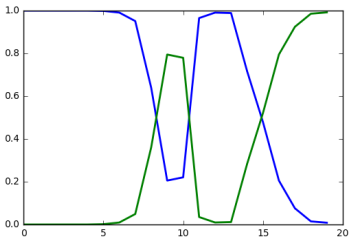


(b) $\sigma_0^2 = 5, \quad \delta^2 = 2$

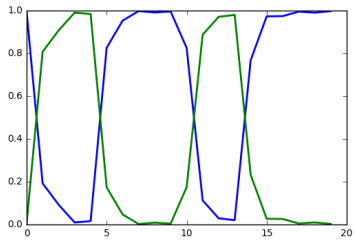


(c) $\sigma_0^2 = 5, \quad \delta^2 = 5$

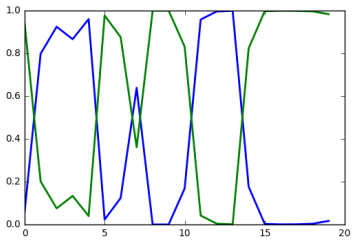
Figure 4: softmax(α) with the average value of the initial variance.



(a) $\sigma_0^2 = 10, \quad \delta^2 = 0.5$



(b) $\sigma_0^2 = 10, \quad \delta^2 = 2$



(c) $\sigma_0^2 = 10, \quad \delta^2 = 5$

Figure 5: softmax(α) with the highest value of the initial variance.

The second experiment was carried to determine the impact of the α update in recovering the original topic fluctuations in the synthetic corpus. For this reason, the autoregressive part of the dynamic topic was disabled. The topic assignments to the documents of the last iteration are visualised in Figure 6 below.

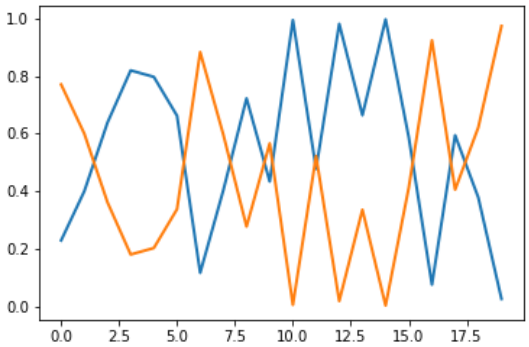


Figure 6: The topic assignments to the documents.