Forecast and Simulation **Summer Term 2023**
http://ml3.leuphana.de/lectures/summer23/FS/
Machine Learning Group, Leuphana University of Lüneburg

Ulf Brefeld (brefeld@leuphana.de)
Yannick Rudolph (yannick.rudolph@leuphana.de)

Exercise 9

Due: Thursday, June 29, 2023

Task 1: Baum-Welch algorithm (continues)

This is a copy of Task 2 of Exercise 8 and relates to Task 1 of Exercise 1.

i) Implement the Baum-Welch algorithm for Task 1 (taking the data from *exercise-02-data.npz*).

ii) As in Task 1 of Exercise 1, apply the Viterbi algorithm to find the most probable sequence of states $z = (z_1, z_2, \ldots, z_5)$ given the following sequence of activities reported by Bob $x = (\mathsf{Walk}, \mathsf{Shop}, \mathsf{Clean}, \mathsf{Shop}, \mathsf{Walk})$, and using the HMM learned with the Baum-Welch algorithm. Try to provide an interpretation for the two possible states.

iii) Compare your learned HMM to the model provided in Task 1 of Exercise 1.

Answer

i) See *solution-09-Baum-Welch.ipynb*.

ii) and iii) Discussed in class.

Task 2: Revisiting the forward-backward algorithm

This is a copy of Task 3 of Exercise 8 and relates to Task 2 of Exercise 2.

i) You might have re-used the forward-backward algorithm that we have implemented in Task 2 of Exercise 2. In our implementation of the algoithm, we have used the *log-sum-exp trick* to prevent numerical problems. Both [Rabiner, 1989] (Section V.A.) and [Bishop, 2006] (Section 13.2.4) rather suggest the use of *scaling factors*. What might be the advantage of scaling compared to the *log-sum-exp trick*?

ii) *Bonus Task: Re-implement the forward-backward algorithm with scaling factors.*

Answer

i) Discussed in class.

ii) *No solution provided.*

Task 3: HM-SVM

This task relates to the Hidden Markov SVM as introduced in [Altun et al., 2003].

i) Please be able to describe the working set optimization for HM-SVMs (Algorithm 2 in [Altun et al., 2003]) in your own words. Make sure that you understand the notation.

ii) Derive the dual of the soft margin HM-SVM with L_2 penalties in Equation 18 of [Altun et al., 2003].

iii) Show why the penalty term (Equation 20 of [Altun et al., 2003]) can be absorbed into the kernel (Equation 21).

Answer

See classnotes.

References

[Altun et al., 2003] Altun, Y., Tsochantaridis, I., and Hofmann, T. (2003). Hidden markov support vector machines. In *Proceedings of the 20th international conference on machine learning (ICML-03)*, pages 3–10.

[Bishop, 2006] Bishop, C. M. (2006). *Pattern recognition and machine learning*. Springer.

[Rabiner, 1989] Rabiner, L. R. (1989). A tutorial on hidden markov models and selected applications in speech recognition. *Proceedings of the IEEE*, 77(2):257–286.