

Probabilistic Models of Human and Machine Learning

CSCI 7222

Assignment 8

Assigned 11/16/2015
Due 12/3/2015

Two paths for Assignment 8 (and 9)

In this assignment, I'll have you experiment with Bayesian optimization. We have one more assignment for the course, which will involve sequential or time series models. In lieu of assignments 8 and 9, you are welcome to do a mini course project. I have spoken to several of you who have research ideas that are a bit more interesting and involved than a single homework assignment, and I have suggested that you work on your ideas instead of assignments 8 and 9. For the rest of you, if you see avenues for using probabilistic models in your own research, the mini-project would be an opportunity to start such an investigation. If you want to discuss a potential mini-project, let's chat.

If you choose to do the mini project, I will ask you to report on the project during our last meeting during finals week.

Bayesian optimization

If you choose to do the standard assignments 8 and 9, I would like you to get experience running Bayesian optimization with Gaussian process surrogate functions. I would like you to get first hand experience with initializing hyperparameters, optimizing the hyperparameters, iteratively running experiments, and observing the convergence on a global optimum.

You are welcome to write as much of the software on your own as you'd like, but it's perfectly acceptable for you to use existing software packages. The packages I know about are as follows:

- MOE
- Spearmint
- BayesOpt
- PyBO

Task 1

Write a function to generate an experiment result given an input vector. I'd suggest using a 2-dimensional input space in order that you can easily visualize the function as well as the progress of Bayesian optimization. For example, your function might be based on a mixture of Gaussians added to a linear function of the inputs. Make the function complex enough that there are some local optima and that the global optimum isn't in one corner of the space. I'd like your function to add noise to the returned value, e.g., Gaussian mean zero noise. Make a plot of the expected value of the function.

Task 2

Run Bayesian optimization with experiment results returned by your function. For the Bayesian optimization, do not build in any knowledge you have about the

solution, e.g., allow the optimization procedure to estimate the function mean and observation noise variance. Describe the assumptions you made in your model and report on the outcome of experiments, e.g., report how close the estimated optimum is from the true optimum as more experiments are performed. To get a reliable estimate of convergence, you'll need to run the optimization process multiple times.

Test at least two different GP covariance functions, and run the optimization process enough times you can determine which covariance function converges most rapidly.

Task 3 [OPTIONAL]

If you're enjoying Bayesian optimization, you might also try to implement multiarm bandit optimization. (The code itself is pretty trivial.) Or you might try to do GP surrogate-based optimization with a non-Gaussian observation model, e.g., suppose the observations are binary (yes/no) and the likelihood is the probability of a 'yes' response using a logistic function of the latent GP value.