TAM 598 Lecture 23:

Global Bayesian Optimization

Announcements:

- HW 6 covers lectures 21–23; due on Fri May 2 problem:

find $\underline{x}^* = \underset{\underline{x}}{\operatorname{argmax}} f(\underline{x})$

given

- we can evaluate $f(\underline{x})$ at any $\underline{x}$
- evaluating $f(\underline{x})$ takes time / money
- we cannot evaluate $\nabla f(\underline{x})$
- dimensionality of $\underline{x}$ is not very high (< 8)

Bayesian Approach: sequential information acquisition for optimization. A decision making strategy to decide where to evaluate the function next use a one-step-look ahead information acquisition policy:

(1) start with data set of n_0 input-output observations

$$D_{n_0} = \left(\underline{x}_{1:n_0}, \underline{y}_{1:n_0}\right)$$

(2) For $n = n_0, n_0+1, \ldots$

a) use current dataset to build a regression model for $f(\underline{x})$, eg Gaussian process regression

$$f(\cdot) \mid D_{n_0} \sim p\left(f(\cdot) \mid D_{n_0}\right)$$

b) pick the most important point to evaluate next by maximizing an acquisition function $a_{n_0}(\underline{x})$ which depends on our current state of knowledge

b) pick the most important point to evaluate next by maximizing an acquisition function $a_{n_0}(\underline{x})$ which depends on our current state of knowledge

$$\underline{x}_{n+1} = \underset{\underline{x}}{\operatorname{argmax}} \; a_n(\underline{x})$$

$$\text{and} \quad a_n(\underline{x}) \geq 0$$

c) If this max value $a_n(\underline{x}_{n+1}) \leq G$, some threshold, then we stop. We are done.

d) Otherwise, we evaluate $y_{n+1} = f(\underline{x}_{n+1})$

e) Add new data point to our data set

$$D_{n+1} = \left((\underline{x}_{1:n}, \underline{x}_{n+1}), (\underline{y}_{1:n}, y_{n+1}) \right)$$

f) use Bayes' Rule to update our state of knowledge:

$$f(\cdot) \mid D_{n+1} \sim p(f(\cdot) \mid D_{n+1})$$

$$\propto p(y_{n+1} \mid \underline{x}_{n+1}, f(\cdot))\, p(f(\cdot) \mid D_n)$$

g) continue loop (until you are satisfied or run out of evaluation budget)

3) Report your current state of knowledge about the maximum of the function, eg the observed maximum

index of obs. max $\quad i^* = \underset{1 \le i \le n}{\operatorname{argmax}}\ y_i \implies \max\ y_{i^*}$ at x_{i^*}

Common acquisition functions:

- maximum mean (bad)
- max upper interval
- probability of improvement
- expected improvement