MAT 653: Statistical Simulation

Instructor: Dr. Wei Li Scribe: Yuan Liu

Oct 28th, 2021

Simulated Annealing

Consider maximize $h(\theta)$ (global maximization):

Assume $h : \Theta \mapsto \mathbb{R}$, $\int_H h(\theta) < \infty$, $H(\theta) \propto exp(\frac{h(\theta)}{T}), T > 0$, where T is the temperature.

Define the so-called Boltzman-Gibbs transformation of h to be

$$\pi_t(\theta) \propto exp\Big(\frac{h(\theta)}{T_t}\Big).$$

As $T_t \downarrow 0$, the kernel of density will become more and more concentrated in a narrower neighborhood of the maximizer of h.

Algorithm

Suppose $g(\cdot)$ symmetric density function $g(-x) = g(x)$, e.g. N(0,1).

1. simulate $\zeta^{(t)}$ from g
2. generate $\theta^{(t+1)}$ as

$$\theta^{(t+1)} = \begin{cases} \theta^{(t)} + \zeta^{(t)} & \textit{with prob } \rho = exp(\frac{\triangle h}{T_t}) \wedge 1, \triangle h = h(\theta^{(t)} + \zeta^{(t)}) - h(\theta^{(t)}) \\ \theta^{(t)} & \textit{with prob } 1 - \rho \end{cases}$$

3. update $T_{t+1} \leftarrow T_t$. For example, one can let $T_t = \frac{1}{log(1+t)}$ or $(\frac{1}{1.2})^t$.

Remarks:

1. If the perturbation increases h (i.e., $\triangle h \geq 0$), the new proposal is automatically accepted. If $\triangle h < 0$, then still accept it with certain probability $\rho (< 1)$.
2. When T_t is relatively large, many "bad" proposals (that is those $\triangle h < 0$) are still accepted. Therefore a larger part of the space can be explored. But at larger stages, as $T_t \downarrow 0$, that will limit the "bad" proposal that we allow.
3. When T_t is sufficiently small, one essentially "quenches" the process by accepting only "good" proposals (i.e., $\triangle h \geq 0$).