Chapter 1

Simulated Annealing

1.1 Introduction

In traditional processing of metals, a standard method to improve the quality of metal is to heat it up to high temperatures, then slowly cool it down. This method is called "Annealing". Sometimes this is done in repetitive cycles [1]. Thus, we consider the cooling process by these four core steps:

Step 1: Start at a very high temperature.

Step 2: Temperature decreases slowly by steps.

Step 3: Transition from a high-energy unordered regime to a low energy (partially) ordered regime.

Step 4: Optimization process stops when the system is frozen in a (quasi) optimum state at a low temperature.

The idea here is to introduce a variable T, then a generalized Metropolis scheme is used to simulate the system at the temperature T, while T is gradually cooled down from some high starting temperature [1].

Therefore, we look at the method as a Metropolis algorithm:

Metropolis algorithm: In statistics and in statistical physics, the Metropolis algorithm is a Markov chain Monte Carlo (MCMC) method for obtaining a sequence of random samples from a probability distribution for which direct sampling is difficult [2].

Markov chain: A Markov process is a random and memoryless process in which in each transaction, the next state of the process depends only on the current state and not on the sequence of event that already proceed. Therefore, for each given present state, the future and past states are completely independent [2].

Formally,

$$P(X_{n+1} = x | X_1 = x_1, X_2 = x_2, X_3 = x_3, ..., X_n = x_n) = P(X_{n+1} = x | X_n = x_n)$$

Thus, the possible values of X_i form a countable set S called the "state space" of Markov chain.

Generalized Metropolis algorithm:

We consider a system with a state described by an N-dimensional vector x , for which the function to be minimized is $f(x)$. The function should return a scalar quantity. The generalized temperature T is just a scalar quantity which has the same dimensions as f [1]. To generate Metropolis algorithm we follow these steps [1]:

1. (a) Select the initial configuration x.
   (b) Set the number of Monte Carlo steps $n_{\text{MCS}} = 0$.
   (c) Set the initial temperature to some high value T_0.
2. Choose a transition $\triangle x$ at random
3. Calculate the function value before the transition $f_b = f(x)$.
4. Do the trial transition as $x = x + \triangle x$.
5. Calculate the function value after the transition $f_a = f(x)$.
6. Calculate $\triangle f = f_a - f_b$, then:
7. If $\triangle f \leq 0$ accept the state.
8. If $\triangle f > 0$:
   (a) Generate a random number u between 0 and 1.
   (b) Accept the state only if $u < e^{-\triangle f/T}$.
9. If the state is rejected, return to the previous state: $x = x - \triangle x$.
10. Reduce the temperature by some small value: $T = T - \varepsilon_T$.
11. Set $n_{\text{MCS}} = n_{\text{MCS}} + 1$.
12. If $T > 0$ return to step 1.

Canonical equilibrium distribution of classical systems:
If Γ set of states, $\sigma \in \Gamma$ and $H(\sigma)$ the system energy in state σ, then:

$$P_{eq}(\sigma) = \frac{1}{Z} exp(\frac{H(\sigma)}{k_B T}), \text{where } Z = \sum_{\sigma \in \Gamma} exp(\frac{H(\sigma)}{k_B T})$$

K_B: a physical constant

Thus, at equilibrium:

$$\frac{P(\sigma \to \tau)}{P(\tau \to \sigma)} = exp(-\frac{\triangle H}{k_B T})$$

In this stage, we have two different choices:

1. **Metropolis criterion:**

$$P(\sigma \to \tau) = \begin{cases} exp(-\frac{\triangle H}{k_B T}) & \text{if } \triangle H > 0 \\ 1 & \text{if not} \end{cases}$$

a: If the new state is better or as good as the actual state, it is always accepted.
b: If is worse, there is a positive probability to be accepted.

2. **Heat bath condition:**

$$P(\sigma \to \tau) = \frac{1}{1 + exp(\frac{\triangle H}{k_B T})}$$

Any new state is accepted with a positive probability.

1.2 SA Method

Simulated Annealing (SA) is a probabilistic algorithm method that is proposed by Kirkpatrick, Gelette and Vecchi (1983) and Gerny (1985) to find the global minimum of a cost function.

The idea is to generate a path through the space of the possible solutions. Thus, it should be from one solution to another nearby solution, leading ultimately to the optimum solution. For generating this path, solutions are chosen from the locality of the preceding solution by a probabilistic function of the improvement gained by this move. So, we restrict ourselves to the case of a cost function defined on a finite set. The algorithm consists of several steps that are not strictly required to produce improved solutions. But each of theme has a certain probability of leading to improve the last result.

At the start all steps all equally likely, but as the algorithm progresses, the tolerance for solutions worse than the current one decreases, eventually to the point where only improvements are accepted. In this way the algorithm can attain the optimum solution without becoming trapped in local optima.

The main steps of method can be mentioned as follow:

Step 0: Initialization
Set initial temperature T and initial configuration $\hat{\theta}_0 = \theta_{curr}$; calculate $L(\theta_{curr})$.

Step 1: Candidate value
Relative to θ_{curr}, randomly determine new configuration θ_{new} and calculate $L(\theta_{new})$.

Step 2: Accept transition / Reject transition
Let $\triangle L = L(\theta_{new}) - L(\theta_{curr})$. If $\triangle L < 0$ accept θ_{new}. If not, generate σ uniform on $(0, 1)$ and accept θ_{new} if $\sigma < exp(- \triangle L/T)$; otherwise keep θ_{curr}.

Step 3: Iterate at fixed temperature
Repeat steps 1 and 2 until T is changed (an equilibrium is reached).

Step 4: Decrease temperature
Lower T according to the annealing schedule and return to Step 1. Continue till effective convergence.

Bibliography

[1] Kai Nordlund, *Basics of Monte Carlo simulations.* Gould-Tobochnik p. 607-610, 2006.

[2] William M. Bolstad, *Study Guide for Understanding Computational Bayesian Statistics* 2012.

[3] Dimitris Bertsimas, John Tsitsiklis, *Theory of multidimensional scaling.* Statistical Science, Vol. 8. No. 1. p10-15, 1993.