Simulated Annealing (Recozimento Simulado)

UFMG

10/11/2025

Sumário

Recozimento Simulado - Introdução

Também chamado de método meta-heurístico, este tem como intuito resolver problemas de otimização de grande complexidade. Este algoritmo foi introduzido por volta de 1980 por três pesquisadores, sendo eles: Kirkpatrick, Gelatt e Vecchi. O método traz soluções sub-ótimas, ou seja, eficientes e sem grande esforço computacional.
Tem como objetivo encontrar resultados satisfatórios, ainda que não exatos, sendo que muitas vezes estes resultados são inviáveis e/ou difíceis de serem alcançados.
Se enquadra na classe *Markov chain Monte Carlo* (MCMC) e, por ter caráter estocástico alinhado à sua capacidade de escapar de mínimos locais, este é um ótimo candidato para resolução de problemas complexos de otimização.

Processo de Recozimento

O conceito foi introduzido da ideia de recozimento, onde um sólido é levado para um estado de baixa energia após aumentar sua temperatura.
O processo consiste em dua estapas, sendo elas:

- "Derretimento" da sua estrutura ao levá-lo a uma temperatura muito elevada.
- Resfriamento esquematizado buscando atingir um estado sólido de energia mínima.

Uma vez em estado líquido as partículas do material exposto são distribuídas de forma aleatória.
O estado mínimo de energia só é alcançado com uma temperatura inicial que seja suficientemente alta e um tempo de resfriamento que seja logo o suficiente.

Aplicabilidade

O algoritmo **Simulated Annealing** tem aplicabilidade em várias áreas e aqui serão expostas algumas delas:

- Engenharia de software
- *Machine Learning*
- Teoria de filas
- Processos de manufatura
- logística e transporte com otimização

Teoria

Fundamentos - Inferência Estatística

Busca por estimação a posteriori dos parâmetros θ condicionado pelos dados observados em $\mathbf{y}$. Tem como fórmula, o seguinte:

$$f_{\theta|\mathbf{Y}}(\theta|\mathbf{y}) = \frac{f_{\theta,\mathbf{Y}}(\theta,\mathbf{y})}{f_{\mathbf{Y}}(\mathbf{y})} = \frac{f_{\mathbf{y}|\theta}(\mathbf{y})f_{\theta}(\theta)}{\int_0^1 f_{\theta,\mathbf{Y}}(\theta,\mathbf{y})d\theta}$$

Onde cada item representa:

- $f_{\mathbf{y}|\theta}(\mathbf{y}) \rightarrow$ função de verossimilhança;
- $f_{\theta}(\theta) \rightarrow$ distribuição a priori;
- $f_{\theta,\mathbf{Y}}(\theta,\mathbf{y}) \rightarrow$ distribuição conjunta de θ e $\mathbf{Y}$;
- $f_{\mathbf{Y}}(\mathbf{y}) \rightarrow$ distribuição preditiva (independente de θ);
- $f_{\theta|\mathbf{Y}}(\theta|\mathbf{y}) \rightarrow$ distribuição a posteriori.

Tem como ideia central resolver o problema simulando o modelo em vez de se calcular a verossimilhança.
Sua ideia básica consiste em:

$$f_{\theta|\mathbf{Y}}(\theta|\mathbf{y}) \propto \int f_{\mathbf{X}|\theta}(\mathbf{x}) f_{\theta}(\theta) \delta(\mathbf{x} - y) dx$$

É evidente que $\delta(\mathbf{x} - y)$ é um cálculo inviável, portanto se é aplicada uma função de tolerância $exp\{-\rho(x, y)/\varepsilon\}$:

$$\pi_{\varepsilon}(\theta, \mathbf{x}) \propto f_{\mathbf{X}|\theta}(\mathbf{x}) f_{\theta}(\theta) exp\{-\rho(x, y)/\varepsilon\}$$

Se a simulação $\mathbf{x}$, portanto estiver próxima de y, será aceita o θ que neste caso é controlado por ε