Method of Least Squares

Data : (x_i, y_i) $1 \leq i \leq n$

"$Y = w_1 X + w_2$ is the relation"

Find w_1 and w_2 so that :

$\sum_{i=1}^{n} (y_i - w_1 x_i - w_2)^2$ is Minimized

Solution

- differentiate
- normal Equations
- Can we Learn ?

Gradient Descent

Parameters :-

- Starting Point, w_0
- Step Size, ε
- Iterations, T

Algorithm

- $w(t+1) = w(t) - \varepsilon \nabla L(w(t))$
- stop at $t = T-1$
- take best $\tilde{w}$

How Low Can You Sink?

In Search of Global Minima

Vivek S Borkar

Vivek S Borkar obtained his Ph D from the University of California, Berkeley, in 1980 and is now with the Department of Computer Science and Automation, Indian Institute of Science, Bangalore. He likes to

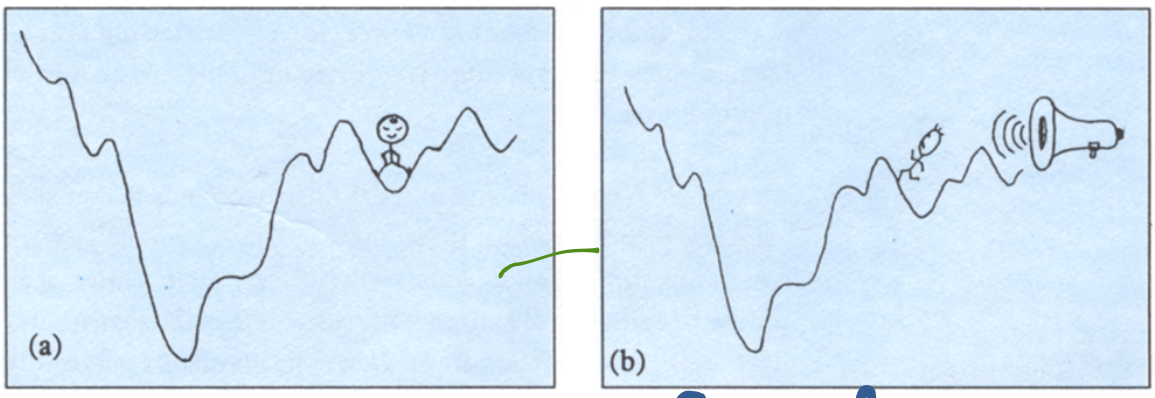
Many scientific and technological problems involve finding the lowest attainable value of a function. This article motivates, describes, and discusses one powerful computational strategy. It is called "simulated annealing" in analogy with the physical process of slowly cooling a system to take it to its ground state.

Getting to the Bottom of It

Vivek S. Borkar

Resonance

February 1997



Stochastic Gradient descent

$$w_{t+1} = w_t - \varepsilon\left(\nabla f(w_t) + \xi_t\right)$$