

Stochastic Descent Algorithms

The goal of this repository is to minimize a non-convex function

$$\min_x F(x),$$

In order to attain a global minimum we add a noise to three different descent algorithms

Gradient Descent:

$$X^{n+1} = X^n - \frac{1}{\gamma} \nabla_X F(X^n) \Delta t_n + \frac{\sigma_n}{\gamma} \Delta W^n$$

Moment 1 (Langevin):

$$\begin{cases} V^{n+1} = (1 - \mu \Delta t_n) V^n - \frac{1}{\gamma} \nabla_X F(X^n) \Delta t_n + \frac{\sigma_n}{\gamma} \Delta W^n \\ X^{n+1} = X^n + V^{n+1} \Delta t_n \end{cases}$$

Moment 2:

$$\begin{cases} z^{n+1} = -\lambda_1 z^n \Delta t + \lambda_2 V^n dt + \frac{\sigma_n}{\gamma} \Delta W^n \\ V^{n+1} = (1 - \mu \Delta t_n) V^n - \frac{1}{\gamma} \nabla_X F(X^n) \Delta t_n - z^{n+1} \Delta t_n. \\ X^{n+1} = X^n + V^{n+1} \Delta t_n \end{cases}$$

where $\sigma_n \rightarrow 0$, $\Delta t_n \rightarrow 0$, $\Delta W^n \sim N(0, 1)$