Gradients and Algorithmic Perspectives

Alexander G. Ororbia II

Biologically-Inspired Intelligent Systems

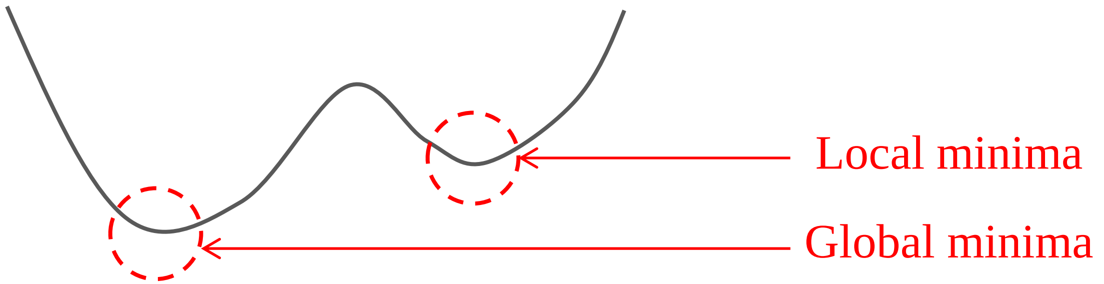
CSCI-633

1/25/2024

Optimization through Derivatives

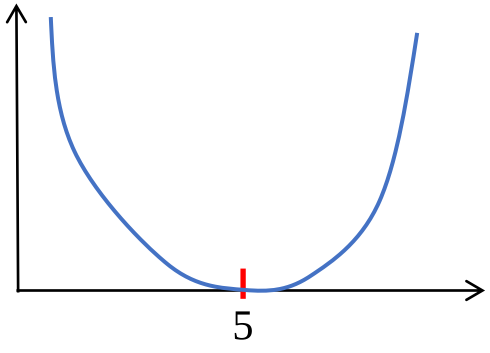
- How would we solve the following problems?

(1)



(2) $\min_{w} f(w) = (w-5)^2$ ⇒

(a) Plot

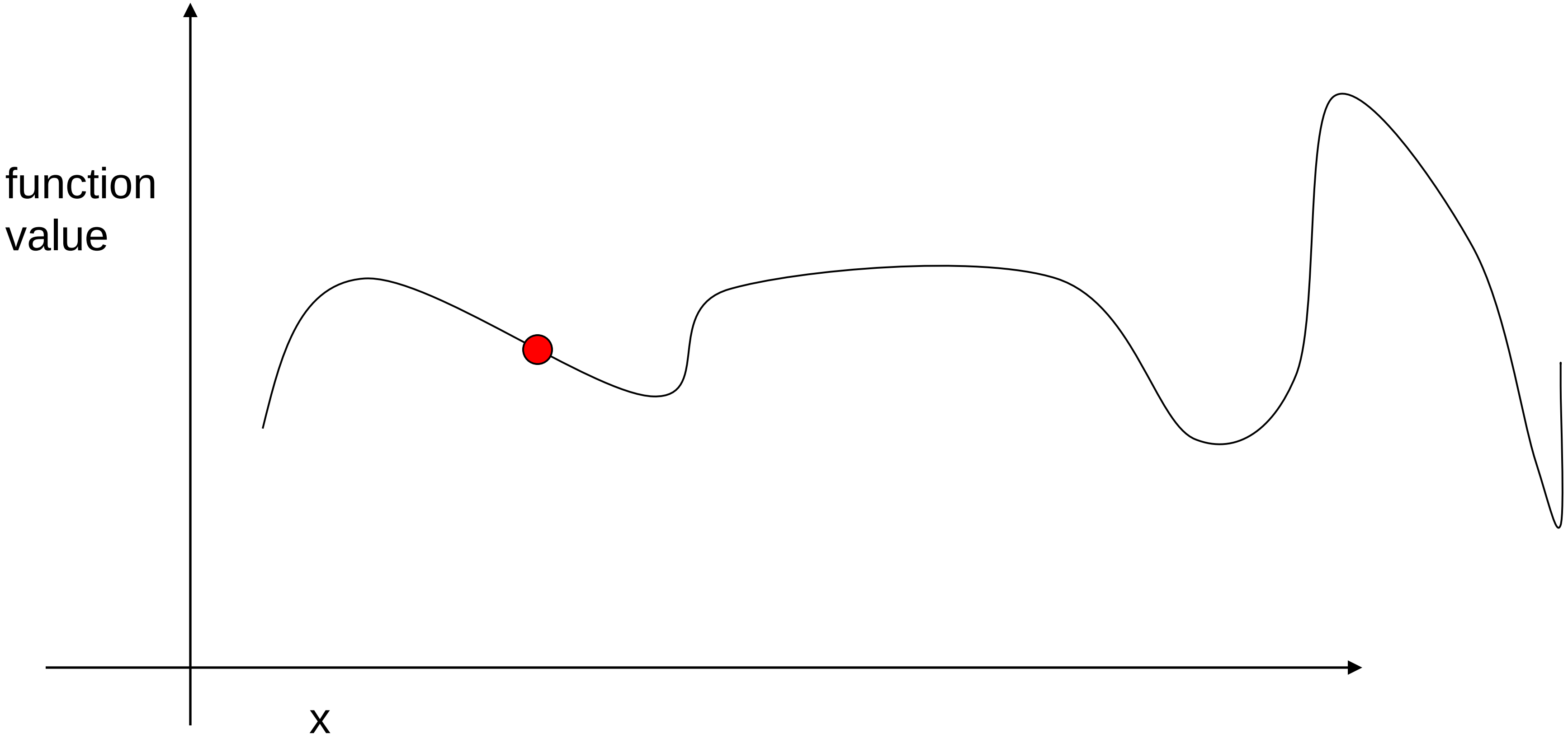


(b) Take ***derivatives***, **check = 0**

Following “Gradient Flows”

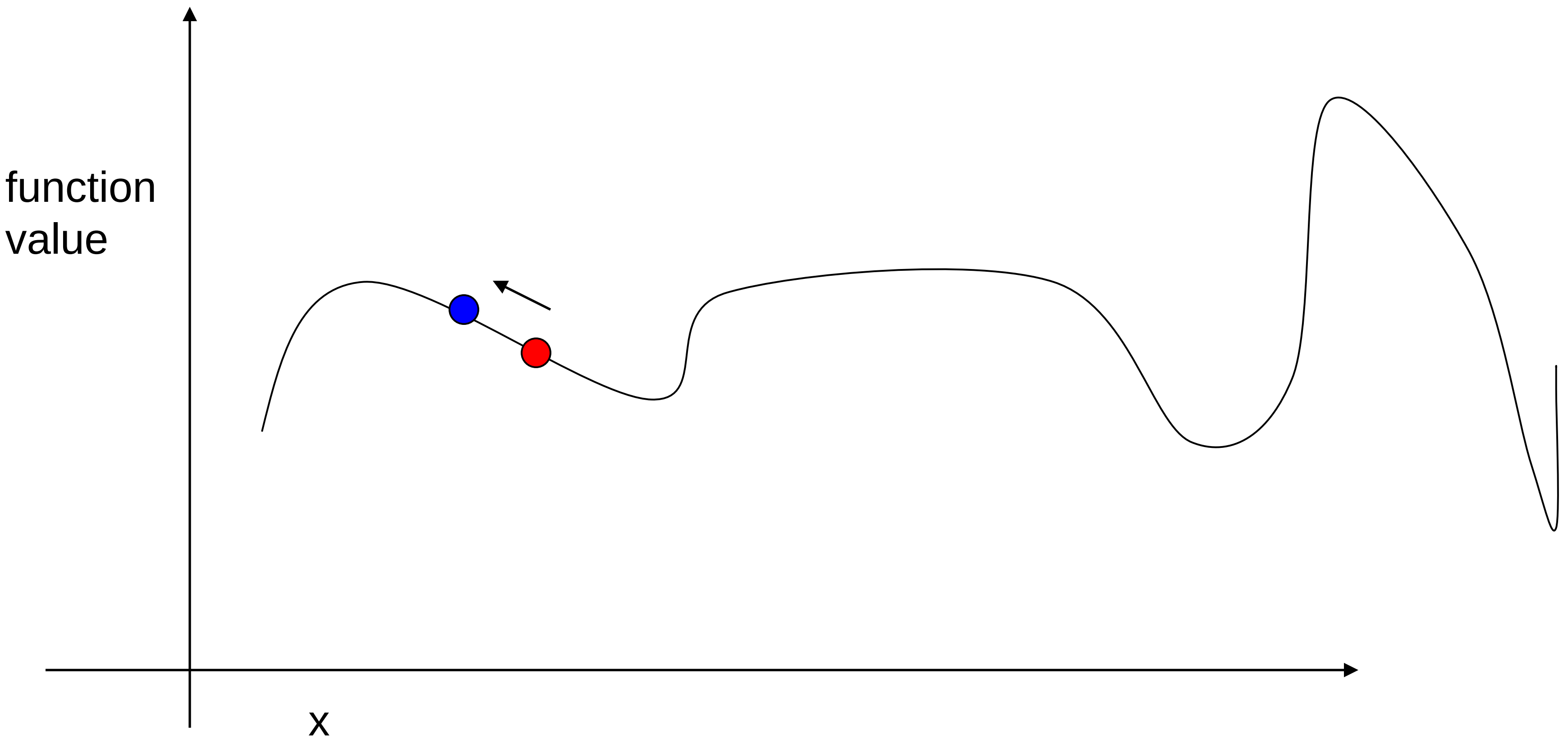
Gradient **Ascent**

- Random Starting Point



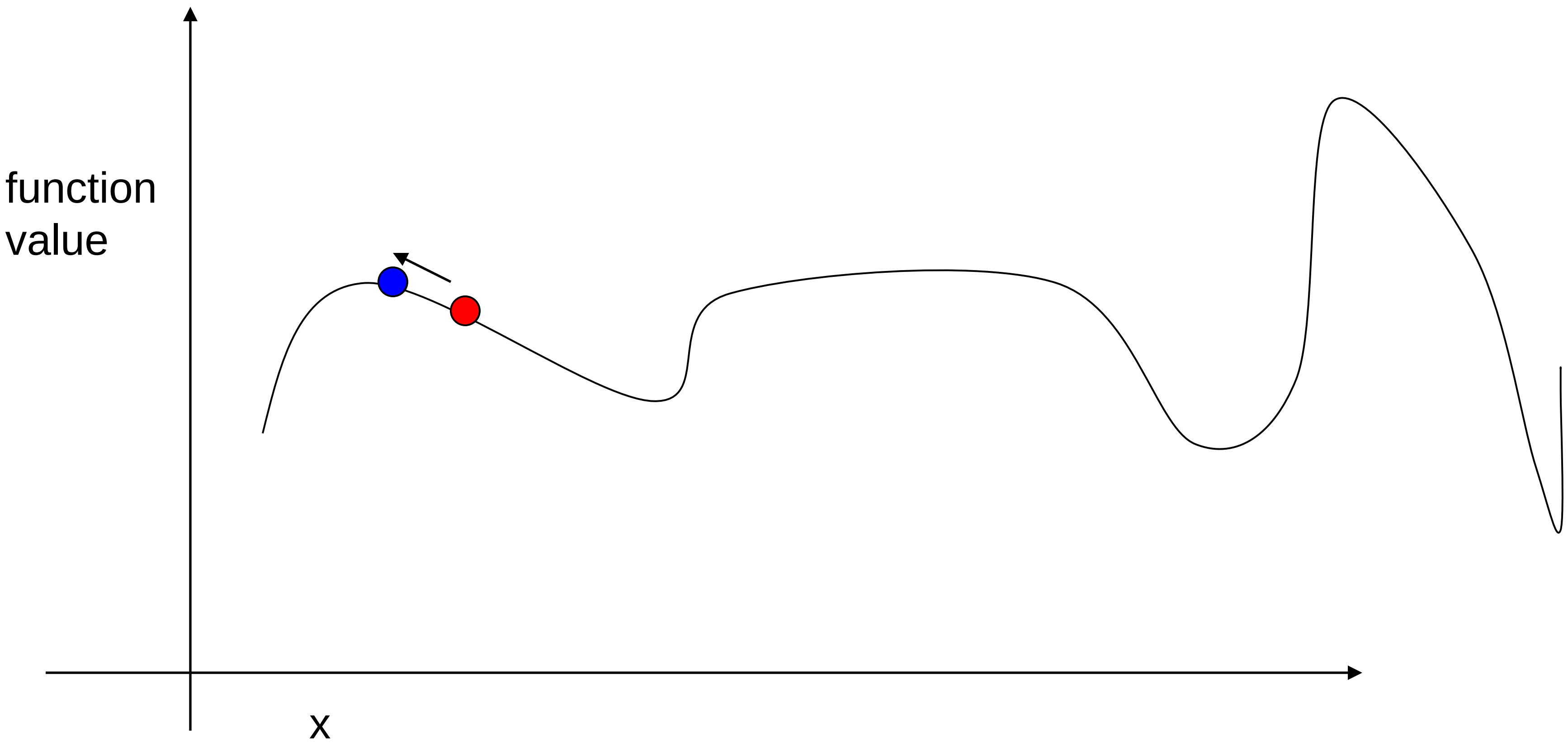
Gradient Ascent

- Take step in direction of largest increase (obvious in 1D, must be computed in higher dimensions)



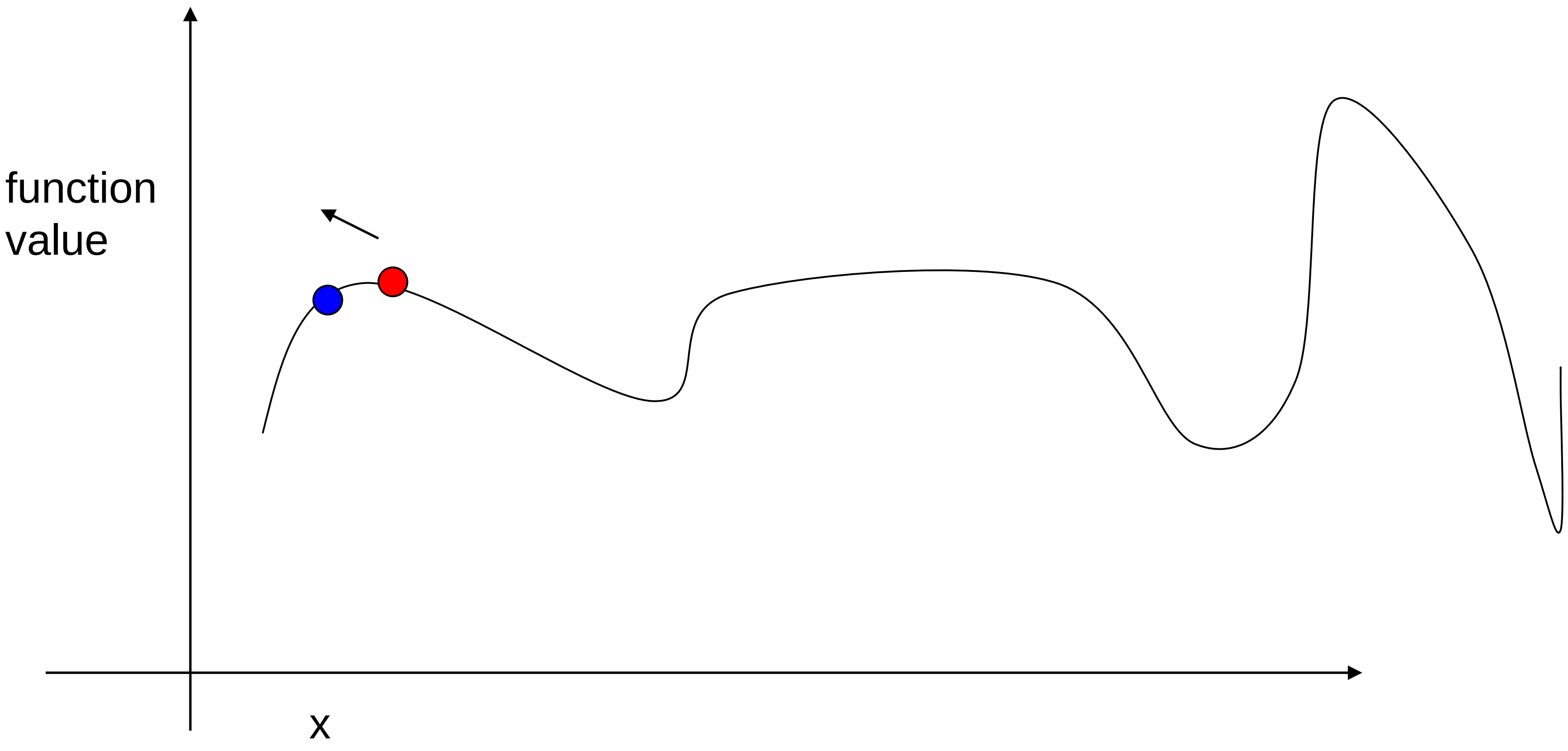
Gradient Ascent

- Repeat



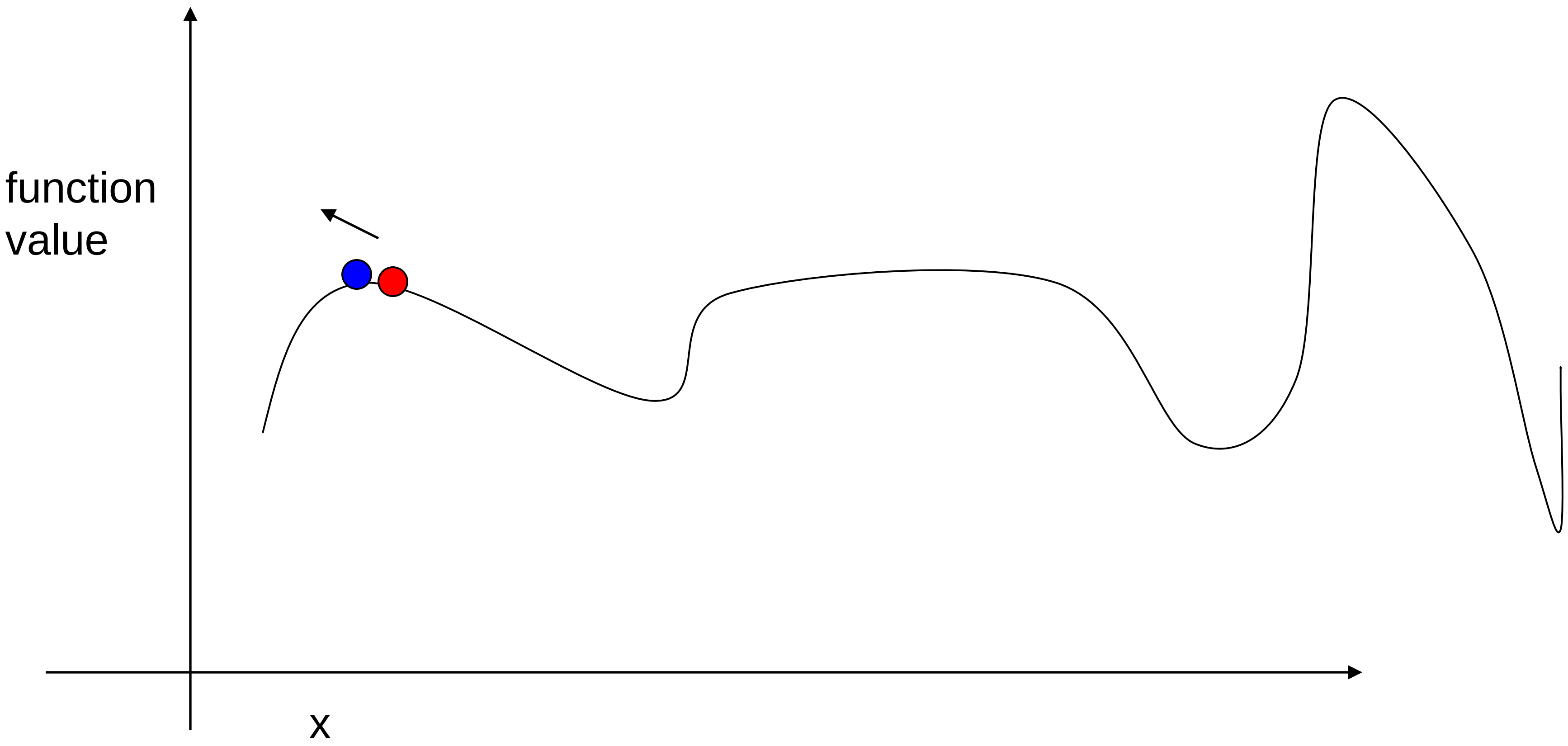
Gradient Ascent

- Next step is actually lower, so stop



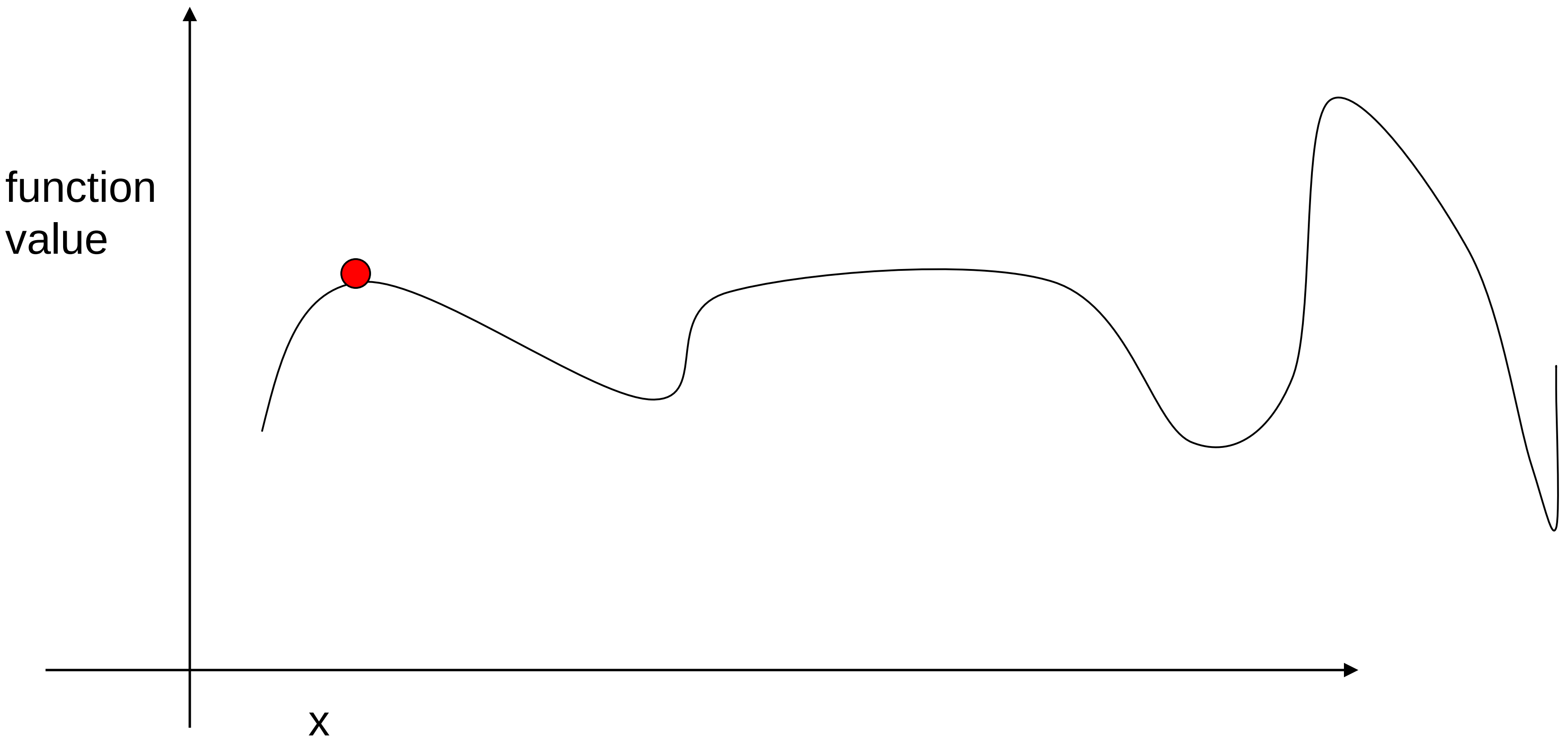
Gradient Ascent

- Could reduce step size to “hone in”



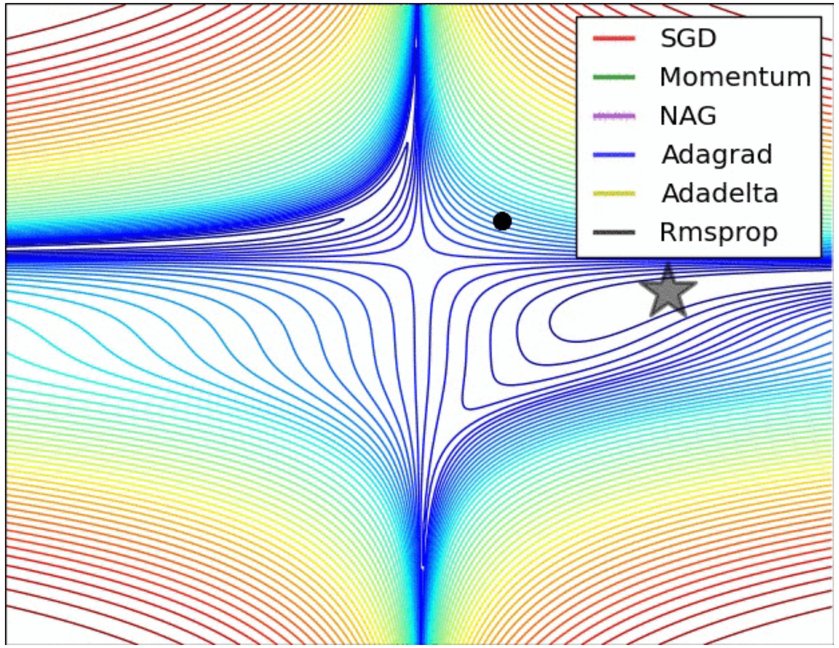
Gradient Ascent

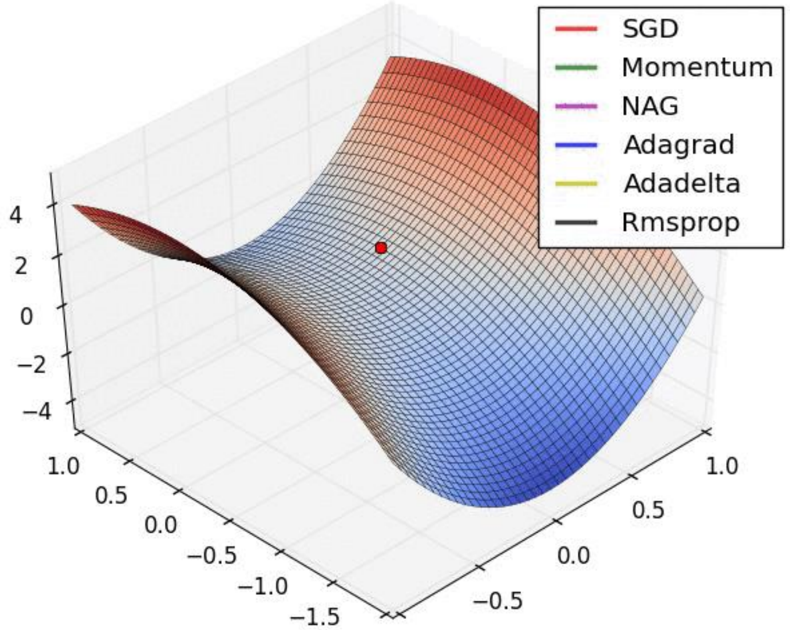
- Converge to (local) maximum



Newton’s Method

Race of the Optimizers!





http://cs231n.github.io/neural-networks-3/#hyper

Problem Specification/Formulation

- **General mathematical optimization (minimization) problem:**

minimize $f_i(\boldsymbol{x}), i = 1, 2, \ldots, M$

subject to $h_j(\boldsymbol{x}) = 0, \quad j = 1, 2, \ldots, J$

$g_k(\boldsymbol{x}) \leq 0, \quad k = 1, 2, \ldots, K$

- f_i : $\boldsymbol{R^d} \rightarrow \boldsymbol{R}$: objective/cost fnctn (maps search/design space -> solution/response space)
- $\boldsymbol{x}=(x_1, \ldots, x_d)^T$: design variables - unknowns of the problem, could be mix of discrete & continuous (contus) values (***x*** is “design vector”)
- h_j : $\boldsymbol{R^d} \rightarrow \boldsymbol{R}$: inequality constraints
- g_k : $\boldsymbol{R^d} \rightarrow \boldsymbol{R}$: equality constraints

- This problem is a constrained optimization problem
  - Linear constraints + linear objectives → linear programming problem

Algorithmic View: An Iterative Process

- ***[White board notes]***

Questions?