eCornell

CHEAT SHEET

Gradient Descent

Algorithm Name	Gradient Descent
Description	Gradient descent is a minimization method that uses only the gradient information. Essentially, you update the parameters by stepping in $-\nabla f$, which is the steepest decreasing direction for function f.
Applicability	Minimization problems.
Assumptions	The objective function has to be differentiable, namely, the gradient exists.
Underlying Mathematical Principles	• Gradient • Partial derivatives
Additional Details	• Gradient descent gives optimal solution if the loss function is convex. • If loss function is not convex, gradient descent might produce a local minimum. • The learning rate is a hyperparameter.
Example	You can use gradient descent to fi nd the optimal weight vector for the logistic loss function, which has no analytical solution.