Tutorial 3 Feb 12, 2021

Non-closed Form Solution

Unlike the LS regression which has a closed form solution of the optimal weights, logistic regression for classification does not.

We use the fact that convex functions always have a minimum to find the optimal weights

Gradient Descent

1. make initial guess w_1
2. evaluate the negative gradient at point, $-\nabla E(w_1)$
   - take negative since gradient points to direction of increase
3. make new guess $w_2 = w_1 - \lambda \nabla E(w_1)$
4. iterate 3 until solution is found or maximum # of iterations is reached

Picking Initial Guess

1) Random guess
2) Solution to similar that does have a closed form solution

Picking Step Size

1) Random size

2) Used backtracking line search to determine step size at each iteration, check if step increase causes error to go up

Termination Condition

1) Max number of iterations to run
2) Check for convergence, if decrease in error is less than ε.

How to avoid getting trapped in local min?

1) Use momentum to optimize
2) Stochastic gradient descent for cheaper estimation