test

Patrick Gardocki

2023-10-24

1. Optimization

a)

$\nabla l(\theta) = l'(\theta)$

$= \sum_{i=1}^{m} \frac{d}{d\theta}(-log(1 + exp(-\theta x^i))) + \frac{d}{d\theta}((y^i - 1)\theta x^i)$

$\nabla l(\theta) = \sum_{i=1}^{m} \left(\frac{x^i exp(-\theta x^i)}{1+exp(-\theta x^i)} + (y^i - 1)x^i \right)$

b)

Initialize: θ, γ, ϵ

While: $|\theta^{t+1} - \theta^t| > \epsilon$

Do: $\theta^{t+1} = \theta^t + \gamma \nabla(\theta)$

c)

Initialize: $\theta, \gamma, \epsilon, K$

While: $|\theta^{t+1} - \theta^t| > \epsilon$

Do: $\theta^{t+1} = \theta^t + \gamma \sum_{i \subset S_k} \left(\frac{x^i exp(-\theta x^i)}{1+exp(-\theta x^i)} + (y^i - 1)x^i \right)$

d)

Given: $\nabla l(\theta) = \sum_{i=1}^{m} \left(\frac{x^i exp(-\theta x^i)}{1+exp(-\theta x^i)} + (y^i - 1)x^i \right)$

$l''(\theta) = \sum_{i=1}^{m} \frac{d}{d\theta} \left(\frac{-x^i}{1+exp(-\theta x^i)} \right) = \sum_{i=1}^{m} \left(\frac{-x^{i\,2}}{(1+exp(-\theta x^i)^2} \right)$

$l(\theta)$ is concave because the hessian is less than 0. There is a global minimum and gradient descent will achieve a unique solution when the gradient is at or near 0.

2. Naive bayes for spam filtering

a)

b)