ML Homework 1

林大权
ID: 85610653

1 (1) ：1) 反欺诈：一些恶意商户为了在某个电商平台或者推荐平台上，刷自己旗下的商品的交易量或访问量，这时平台的安全人员可以用机器学习算法进行数据分析，找到离群点，很快发现欺诈数据。

2）交通拥堵预测：分析以往数据预测未来道路交通情况，及时规避拥堵路段。

2 (2) ：

证明. Suppose a sufficiently small step $\alpha \boldsymbol{d}$, $\alpha > 0$.

We have $f(\boldsymbol{x} + \alpha \boldsymbol{d}) = f(\boldsymbol{x}) + \alpha \nabla f(\boldsymbol{x})^T \boldsymbol{d} + \alpha ||\boldsymbol{d}|| \sigma_{\boldsymbol{x}}(\alpha \boldsymbol{d})$, where $\sigma_{\boldsymbol{x}}(\alpha \boldsymbol{d}) \to 0$ as $\alpha \to 0$

$\Rightarrow$

$$\frac{f(\boldsymbol{x} + \alpha \boldsymbol{d}) - f(\boldsymbol{x})}{\alpha} = \nabla f(\boldsymbol{x})^T \boldsymbol{d} + ||\boldsymbol{d}|| \sigma_{\boldsymbol{x}}(\alpha \boldsymbol{d})$$

Since $\nabla f(\boldsymbol{x})^T \boldsymbol{d} < 0$ and $\sigma_{\boldsymbol{x}}(\alpha \boldsymbol{d}) \to 0$ as $\alpha \to 0$

$\therefore f(\boldsymbol{x} + \alpha \boldsymbol{d}) < f(\boldsymbol{x})$, $\forall \alpha > 0$ sufficiently small.

□

2 (1) ：

证明. $\because$ the Hessian of f : $\boldsymbol{H}$is PSD, and $\forall \boldsymbol{x} \neq \boldsymbol{0}$, we have $\boldsymbol{x^T H x} \geq 0$

$\therefore$ From the key, we can get $f(\boldsymbol{x}) \geq f(\boldsymbol{x'}) + \nabla f(\boldsymbol{x'})^T (\boldsymbol{x} - \boldsymbol{x'})$

$\therefore \forall x, y \in \boldsymbol{dom}(f)$, Suppose $z = \lambda x + (1 - \lambda) y$

we have

$$f(\boldsymbol{x}) \geq f(\boldsymbol{z}) + \nabla f(\boldsymbol{z})^T (\boldsymbol{x} - \boldsymbol{z}) \tag{1}$$

$$f(\boldsymbol{y}) \geq f(\boldsymbol{z}) + \nabla f(\boldsymbol{z})^T (\boldsymbol{y} - \boldsymbol{z}) \tag{2}$$

then $(1) * \lambda + (2) * (1 - \lambda)$, we have:

$$\lambda f(\boldsymbol{x}) + (1 - \lambda) f(\boldsymbol{y}) \geq f(\boldsymbol{z}) + \nabla f(\boldsymbol{z})^T (\lambda \boldsymbol{x} + (1 - \lambda) \boldsymbol{y} - \boldsymbol{z}) = f(\boldsymbol{z}) = f(\lambda \boldsymbol{x} + (1 - \lambda) \boldsymbol{y})$$

$\therefore f$ is convex.

□

2 (2) ：

证明. $\because f$ is convex, be definition:

$$\lambda f(\boldsymbol{y}) + (1-\lambda) f(\boldsymbol{x}) \geq f(\lambda \boldsymbol{y} + (1-\lambda)\boldsymbol{x}), \forall \lambda \in [0,1], x, y \in \boldsymbol{dom}(f)$$

$\Rightarrow$

$$f(\boldsymbol{x}) + \lambda(f(\boldsymbol{y}) - f(\boldsymbol{x})) \geq f(\boldsymbol{x} + \lambda(\boldsymbol{y} - \boldsymbol{x}))$$

$\Rightarrow$

$$f(\boldsymbol{y}) - f(\boldsymbol{x}) \geq \frac{f(\boldsymbol{x} + \lambda(\boldsymbol{y} - \boldsymbol{x}) - f(\boldsymbol{x})}{\lambda(\boldsymbol{y} - \boldsymbol{x})}(\boldsymbol{y} - \boldsymbol{x})$$

$\Rightarrow$

$$f(\boldsymbol{y}) - f(\boldsymbol{x}) \geq \nabla f(\boldsymbol{x})^T (\boldsymbol{y} - \boldsymbol{x})$$

If $\nabla f(\boldsymbol{x}^*) \neq 0$, then $\boldsymbol{d} = -\nabla f(\boldsymbol{x}^*)$ is a descent direction, whereby $\boldsymbol{x}^*$ cannot be a global minimizer.

$\therefore$ the global minimizer satisfies the condition $\nabla f(\boldsymbol{x}^*) = 0$

$\square$

3 4 : Please read file LinearRegression.ipynb