Homework 14

Requirements:

1. Digital format (can be typeset or photos, ought to write clearly if written by hand), upload to `https://course.pku.edu.cn/`.

2. Submit by next class

3. A problem is not counted if nobody can work it out

4. Each homework 10 points; 1 point deducted for each week's delay

Problems:

1. Use different methods to solve the logistic regression problem:

$$\min_{\bar{\mathbf{w}}} \frac{1}{N} \sum_{i=1}^{N} \log \left(1 + \exp\left(-y_i(\bar{\mathbf{w}}^T \bar{\mathbf{x}}_i)\right)\right),$$

where $\bar{\mathbf{x}}_i \in \mathbb{R}^{100}$ and $y_i \in \{1, -1\}$ are randomly generated and $N = 10,000$.

1) SGD;

2) Momentum;

3) Nesertov's Accelerated Gradient (NAG);

4) AdaGrad;

5) AdaDelta;

6) Adam;

7) Adan.

Compare their convergence speed, where the horizontal axis is the computing time. Hand in your code and report.

2. Use different methods to solve the logistic regression problem:

$$\min_{\bar{\mathbf{w}}} \frac{1}{N} \sum_{i=1}^{N} \log \left(1 + \exp\left(-y_i(\bar{\mathbf{w}}^T \bar{\mathbf{x}}_i)\right)\right),$$

where $\bar{\mathbf{x}}_i \in \mathbb{R}^{10,000}$ and $y_i \in \{1, -1\}$ are randomly generated and $N = 1,000$.

1) ADAM;

2) RCD(1) (where you have to estimate the elementwise β_i).

Compare their convergence speed, where the horizontal axis is the computing time. Hand in your code and report.

3. Use stochastic ADMM to solve the constrained logistic regression problem:

$$\min_{\bar{\mathbf{w}}} \frac{1}{N} \sum_{i=1}^{N} \log \left(1 + \exp\left(-y_i(\bar{\mathbf{w}}^T \bar{\mathbf{x}}_i)\right)\right), s.t. \mathbf{1}^T \bar{\mathbf{w}} = 1,$$

where $\bar{\mathbf{x}}_i \in \mathbb{R}^{100}$ and $y_i \in \{1, -1\}$ are randomly generated and $N = 10,000$. Hand in your code and report.