Convex Optimization Homework 5, Assignment 4

Dinghuai Zhang, School of Mathematical Sciences, 1600013525

1 Problem Settings

$$\min_x \frac{1}{2}\|Ax-b\|_2^2 + \mu\|x\|_1 \tag{1}$$

It's a standard LASSO problem without constraints.

1.1 Data Building

```
n = 1024;
m = 512;
A = randn(m,n);
u = sprandn(n,1,0.1);
b = A*u;
mu = 1e-3;
x0 = rand(n,1);
```

2 Solve the problem using Adagrad

If we denote gradient of x as g, in each iteration we compute:

$$r \leftarrow r + g \odot g \tag{2}$$

$$x \leftarrow x - \left(\frac{\epsilon}{\delta + \sqrt{r}} \odot g\right) \tag{3}$$

3 Solve the problem using Adam

In each iteration we compute:

$$g \in \partial f\left(x\right) \tag{4}$$
$$s = \rho_1 s + (1 - \rho_1^{\text{iter}})g \tag{5}$$
$$r = \rho_2 r + (1 - \rho_2^{\text{iter}})g \odot g \tag{6}$$
$$\hat{s} = \frac{s}{1 - \rho_1} \tag{7}$$
$$\hat{r} = \frac{r}{1 - \rho_2} \tag{8}$$
$$x = x - \frac{\epsilon \hat{s}}{\delta + \sqrt{\hat{r}}} \tag{9}$$

4 Solve the problem using RMSProp

In each iteration we compute:

$$g \in \partial f(x) \tag{10}$$
$$r = \rho r + (1 - \rho)g \odot g \tag{11}$$
$$v = -\frac{\epsilon}{\sqrt{r}} \odot g \tag{12}$$
$$x = x + v \tag{13}$$

5 Solve the problem using Momentum

In each iteration we compute:

$$g \in \partial f(x) \tag{14}$$
$$v = \alpha v - \epsilon g \tag{15}$$
$$x = x + v \tag{16}$$

6 Numerical results and interpretations

We can see the cpu time, the optimal value and error with cvx mosek in the tabluation:

Method	cpu time/s	error with cvx-mosek	optimal value
cvx-call-mosek	0.86	0.00e+0	7.2844399533e-02
adagrad	0.53	3.70e-6	7.2844392983e-02
adam	0.31	3.19e-6	7.2844342800e-02
rmsprop	0.39	5.07e-6	7.2844385600e-02
momentum	0.44	3.74e-6	7.2844389612e-02

We can see that Adam is the strongest algorithm among these. It compose the tricks from other optimization methods.