Homework 7

Zhen

November 19, 2021

Prblem 1

$$\min_{x_1,x_2} f(x_1,x_2)=e^{x_1+3x_2-0.1}+e^{x_1-3x_2-0.1}+e^{-x_1-0.1}$$

Solve:

(a) By mean value theorem:

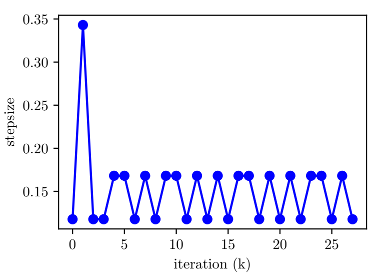
$$\begin{aligned} f(x_1,x_2) &= e^{-0.1}\left[e^{x_1}\left(e^{3x_2}+e^{-3x_2}\right)+e^{-x_1}\right] \\ &\geq e^{-0.1}\left(2e^{x_1}+e^{-x_1}\right) \\ &\geq 2\sqrt{2}e^{-0.1} \end{aligned}$$

The equality condition is: $x_2=0, x_1=-\ln\sqrt{2}$.
Thus $\boldsymbol{x}^*=(-\ln\sqrt{2},0)^{\mathcal{T}},\ f(\boldsymbol{x}^*)=2\sqrt{2}e^{-0.1}$

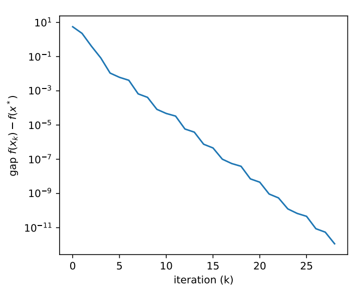
(b) The output is:

```
gradient descent with Armijo
  number of iterations in outer loop: 28
  total number of iterations in inner loop: 151
  solution: [-3.46574284e-01  3.04072749e-07]
  value: 2.5592666966593645
```

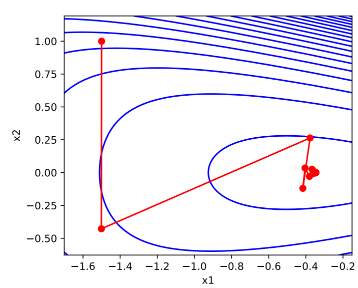
Some plots:



the step sizes t_k



the error $f(\boldsymbol{x}_k)-f(\boldsymbol{x}^*)$



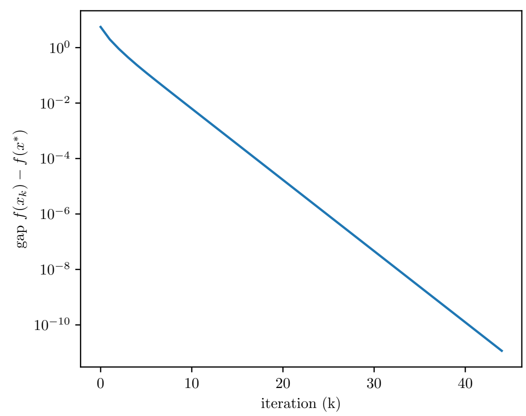
the trajectory of $\boldsymbol{x}_k$

(c) The output is:

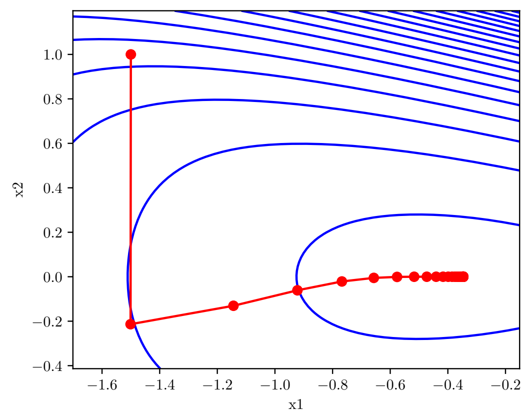
```
gradient descent with constant stepsize  0.1
  number of iterations: 44
  solution: [-3.46576607e-01  3.21465960e-18]
  value: 2.559266696669859

gradient descent with constant stepsize  0.01
  number of iterations: 489
  solution: [-3.46577419e-01  8.65140907e-18]
  value: 2.559266696676969
```

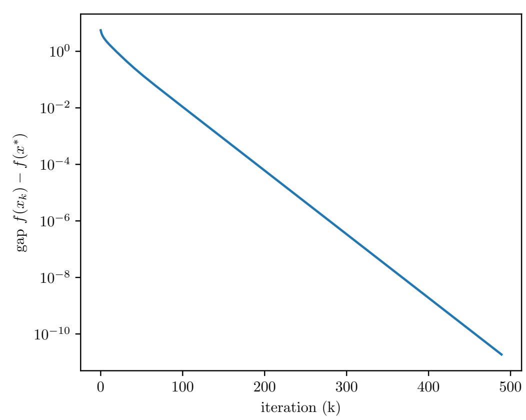
Some plots:



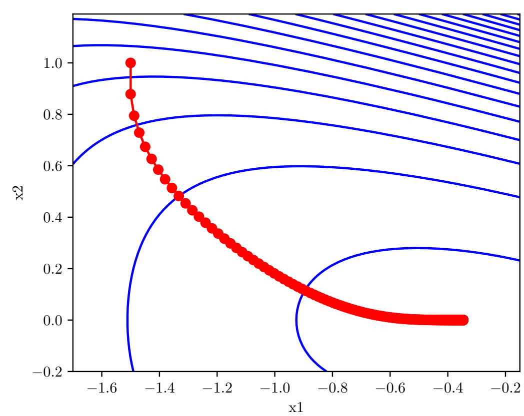
the error $f(\boldsymbol{x}_k) - f(\boldsymbol{x}^*)$



the trajectory of $\boldsymbol{x}_k$



the error $f(\boldsymbol{x}_k) - f(\boldsymbol{x}^*)$

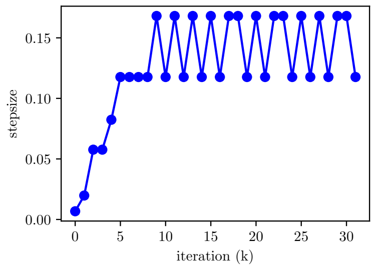


the trajectory of $\boldsymbol{x}_k$

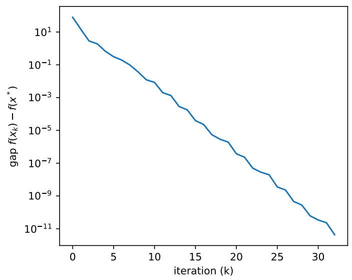
(d) The output is:

```
gradient descent with Armijo
  number of iterations in outer loop: 32
  total number of iterations in inner loop: 197
  solution: [-3.4657238e-01  6.5447655e-07]
  value: 2.5592666966625575
```

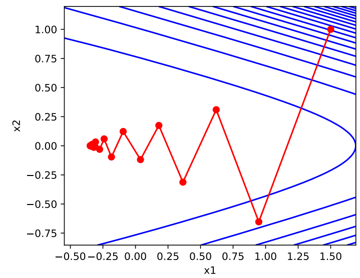
Some plots:



the step sizes t_k



the error $f(\boldsymbol{x}_k) - f(\boldsymbol{x}^*)$



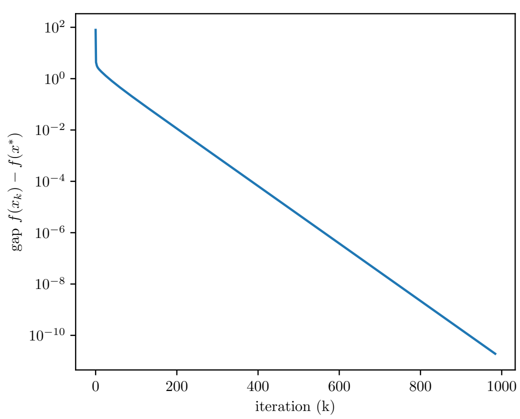
the trajectory of $\boldsymbol{x}_k$

(e) The output is:

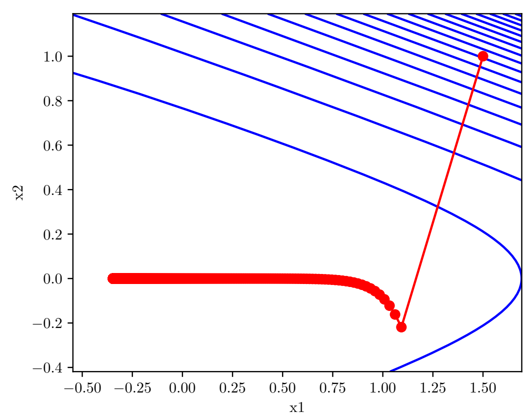
```
gradient descent with constant stepsize  0.1
  number of iterations: 44
  solution: [-3.46576607e-01  3.21465960e-18]
  value: 2.559266696669859

gradient descent with constant stepsize  0.01
  number of iterations: 489
  solution: [-3.46577419e-01  8.65140907e-18]
  value: 2.559266696676969
```

Some plots:



the error $f(\boldsymbol{x}_k) - f(\boldsymbol{x}^*)$



the trajectory of $\boldsymbol{x}_k$

Using the step sizes in part (c) will generate overflow which means that the step sizes are too large and we could not get the optimal solution.

Prblem 2

Noisy gradient:

$$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k - t(\nabla f(\boldsymbol{x}_k) + \boldsymbol{\varepsilon}_k), \text{ where } \|\boldsymbol{\varepsilon}_k\| \le E$$

Proof:

(a) Trivially,

$$\begin{aligned}
\|\boldsymbol{x}_{k+1} - \boldsymbol{x}^*\| &\le \|\tilde{\boldsymbol{x}}_{k+1} - \boldsymbol{x}^*\| + \|\boldsymbol{x}_{k+1} - \tilde{\boldsymbol{x}}_{k+1}\| \\
&= \|\tilde{\boldsymbol{x}}_{k+1} - \boldsymbol{x}^*\| + t\|\boldsymbol{\varepsilon}_k\| \\
&\le \|\tilde{\boldsymbol{x}}_{k+1} - \boldsymbol{x}^*\| + tE
\end{aligned}$$

(b) We just need to prove: $\|\tilde{\boldsymbol{x}}_{k+1} - \boldsymbol{x}^*\| \le \sqrt{1-mt}\|\boldsymbol{x}_k - \boldsymbol{x}^*\|$

$$\begin{aligned}
\|\tilde{\boldsymbol{x}}_{k+1} - \boldsymbol{x}^*\|^2 &= \|\boldsymbol{x}_k - t\nabla f(\boldsymbol{x}_k) - \boldsymbol{x}^*\|^2 \\
&= \|\boldsymbol{x}_k - \boldsymbol{x}^*\|^2 + t^2\|\nabla f(\boldsymbol{x}_k)\|^2 + 2t\nabla f(\boldsymbol{x}_k)^{\mathcal{T}}(\boldsymbol{x}^* - \boldsymbol{x}_k) \\
&\le \|\boldsymbol{x}_k - \boldsymbol{x}^*\|^2 + 2t[f(\boldsymbol{x}_k) - f(\boldsymbol{x}_{k+1})] + 2t\Big[f(\boldsymbol{x}^*) - f(\boldsymbol{x}_k) - \frac{m}{2}\|\boldsymbol{x}_k - \boldsymbol{x}^*\|^2\Big] \\
&= (1-mt)\|\boldsymbol{x}_k - \boldsymbol{x}^*\|^2 + 2t[f(\boldsymbol{x}^*) - f(\boldsymbol{x}_{k+1})] \\
&\le (1-mt)\|\boldsymbol{x}_k - \boldsymbol{x}^*\|^2
\end{aligned}$$

Thus we have $\|\boldsymbol{x}_{k+1} - \boldsymbol{x}^*\| \le \sqrt{1-mt}\|\boldsymbol{x}_k - \boldsymbol{x}^*\| + tE$.

(c) Apparently, $\|\boldsymbol{x}_0 - \boldsymbol{x}^*\| \le q^0\|\boldsymbol{x}_0 - \boldsymbol{x}^*\| + \dfrac{1-q^0}{1-q}tE$.

If $\|\boldsymbol{x}_k - \boldsymbol{x}^*\| \le q^k\|\boldsymbol{x}_0 - \boldsymbol{x}^*\| + \dfrac{1-q^k}{1-q}tE$, then:

$$\begin{aligned}
\|\boldsymbol{x}_k - \boldsymbol{x}^*\| &\le q\|\boldsymbol{x}_k - \boldsymbol{x}^*\| + tE \\
&\le q \times \left(q^k\|\boldsymbol{x}_0 - \boldsymbol{x}^*\| + \frac{1-q^k}{1-q}tE\right) + tE \\
&= q^{k+1}\|\boldsymbol{x}_0 - \boldsymbol{x}^*\| + \frac{1-q^{k+1}}{1-q}t
\end{aligned}$$

By induction, the inequality "$\|\boldsymbol{x}_k - \boldsymbol{x}^*\| \le q^k\|\boldsymbol{x}_0 - \boldsymbol{x}^*\| + \dfrac{1-q^k}{1-q}tE$" holds.

(d)

$$\begin{aligned}
\limsup_{k\to\infty}\|\boldsymbol{x}_k - \boldsymbol{x}^*\| &\le \limsup_{k\to\infty}\left(q^k\|\boldsymbol{x}_0 - \boldsymbol{x}^*\| + \frac{1-q^k}{1-q}tE]\right) \\
&= \frac{tE}{1-q} \\
&= \frac{tE}{1-\sqrt{1-mt}} \\
&\le \frac{tE}{1-(1-mt/2)} \\
&= \frac{2E}{m}
\end{aligned}$$