Homework accompanying the lecture "Basics in Applied Mathematics"

Homework 14

Hand in: Tuesday, 04.02.2025, after the lecture in the mailbox at the Math Institut
(Don't forget to put your name on your homework.
Please hand in your solutions in groups of two.)

Exercise 1 (Heavy-Ball method for Quadratic Programming (QP); 8 points)

In this exercise, we will study the convergence rate of the Heavy-Ball method applied to some Quadratic Programming (QP) problem:

$$\min_x \; f(x) \coloneqq \frac{1}{2} x^\top Q x - c^\top x, \tag{1}$$

with $\mu I_n \preceq Q \preceq L I_n$ for some values $0 < \mu < L$.

We recall the Heavy-Ball method definition:

$$x_{k+1} = x_k - \alpha \nabla f(x_k) + \beta(x_k - x_{k-1}), \tag{2}$$

where we set $x_{-1} = x_0$.

We assume that α and β are chosen such that the following condition holds:

$$\frac{\left(1-\sqrt{\beta}\right)^2}{\mu} \leq \alpha \leq \frac{\left(1+\sqrt{\beta}\right)^2}{L} \quad \text{and} \quad \beta \in (0,1) \tag{3}$$

a) Let $x^\star$ be the solution of (1), and define the sequence $\tilde{x}_k \coloneqq x_k - x^\star$. Show that for some symmetric matrix A (that you need to find), the following holds:

$$\tilde{x}_{k+1} = A \tilde{x}_k - \beta \tilde{x}_{k-1} \tag{4}$$

b) Let $a_1, \dots, a_n$ be the eigenvalues of A, and let $v_1, \dots, v_n$ be a corresponding orthonormal basis of eigenvectors, i.e.:

$$A v_j = a_j v_j \quad \text{for all } i = j, \dots, n.$$

Show that for all $j = 1, \dots, n$, we have $|a_j| < 2\sqrt{\beta}$.

c) Show that:

$$\tilde{x}_k = \sum_{j=1}^{n} \tilde{x}_{j,k} v_j \tag{5}$$

where $\tilde{x}_{j,k} \in \mathbb{R}$ are scalars.

In addition, express $\tilde{x}_{j,k+1}$ as a function of $\tilde{x}_{j,k}$ and $\tilde{x}_{j,k-1}$.

d) For $j \in \{1, \dots, n\}$ and $k \in \mathbb{N}$, prove that:

$$z_j \tilde{x}_{j,k} - \beta \tilde{x}_{j,k-1} = (z_j)^k w_j \tag{6}$$

where $z_j \in \mathbb{C}$ is a complex number that verifies $z_j^2 - a_j z_j + \beta = 0$ and $w_j \in \mathbb{C}$ is to be found.

Hint: Show that $z_j \tilde{x}_{j,k} - \beta \tilde{x}_{j,k-1}$ is a geometric sequence.

e) Show that:

$$|\tilde{x}_{j,k}| \leq \left(\sqrt{\beta}\right)^k c_j \, |\tilde{x}_{0,j}| \tag{7}$$

where c_j is a constant that you need to find.

Hint 1: After ensuring that $\mathrm{Im}(z_j) \neq 0$, express $\tilde{x}_{j,k}$ in terms of $\mathrm{Im}(z_j \tilde{x}_{j,k} - \beta \tilde{x}_{j,k-1})$

($\mathrm{Im}(z)$ denotes the imaginary part of a $z \in \mathbb{C}$, not the image of a function!).

Hint 2: Show that $|z_j| = \sqrt{\beta}$.

f) Deduce that:

$$\|x_k - x^\star\|^2 \leq \beta^k C \|x_0 - x^\star\|^2 \tag{8}$$

for some constant C that you need to find.

g) Conclude the following:

$$f(x_k) - f(x^\star) \leq \beta^k \tilde{C} \left(f(x_0) - f(x^\star)\right) \tag{9}$$

for some constant $\tilde{C}$ that you need to find.

Hint: Show that for all x:

$$\frac{\mu}{2}\|x - x^\star\|^2 \leq f(x) - f(x^\star) \leq \frac{L}{2}\|x - x^\star\|^2$$

h) Now assume that α and β are chosen as follows:

$$\alpha = \left(\frac{2}{\sqrt{L} + \sqrt{\mu}}\right)^2 \qquad \beta = \frac{\sqrt{L} - \sqrt{\mu}}{\sqrt{L} + \sqrt{\mu}}. \tag{10}$$

Show that the assumption (3) holds for this choice of α and β.

Exercise 2 (A function that is difficult to optimize; 4 points)

In this exericise, we consider the following optimization problem:

$$\min_{x \in \mathbb{R}^n} \; f(x) := \frac{1}{2}\left((x^{[1]} - 1)^2 + \sum_{i=1}^{n-1}\left(x^{[i+1]} - x^{[i]}\right)^2\right) \tag{11}$$

where $x^{[i]}$ denotes the indices of x (to not mistake it with the indices of the opimization algorithm).
We consider an optimization algorithm of the following form:

$$x_{k+1} = x_k + \sum_{j=1}^{k} \nu_{k,j} \nabla f(x_{k-j}) \tag{12}$$

for some values of $\nu_{k,j}$ (that might depend on the previous iterations).

a) Prove the the algorithms listed below fall into the general form (12).

- Gradient Descent
- Heavy-Ball method
- Conjugate Gradient

b) Let us assume that $x_0 = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$. Then, prove the following property for all $k \leq n$:

$$x_k^{[i]} = 0 \quad \text{for all } i > k. \tag{13}$$

c) Show that the problem (11) has a unique solution $x^\star$ (that needs to be found).

Prove the following inequality holds for all $k \leq n$:

$$\|x_k - x^\star\| \geq \sqrt{1 - \frac{k}{n}} \|x_0 - x^\star\| \tag{14}$$

d) To solve a convex QP, among algorithms of the form of (12), which algorithm is the fastest for finding the exact solution (in the worst-case scenario)?

Exercise 3 (Programming exercise; 6 points)

Open the jupyter notebook, and fill in the missing parts of the code.