TD 5 – Optimization

Mathematics of data

16/10/24

Exercise 1. Consider the least squares problem:

$$\min_{x \in \mathbb{R}^d} f(x) := \frac{1}{2} \|Ax - b\|^2 ,$$

where $A \in \mathbb{R}^{n \times p}$ is the data matrix and $b \in \mathbb{R}^n$ is the vector of labels, and $\|c\| := \sqrt{\sum c_i^2}$ is the Euclidean norm.

1. Assume that $n < p$ and that $AA^\top$ is invertible. We define $A^+ := A^\top (AA^\top)^{-1}$ the pseudo-inverse of A. Check that $x^* := A^+ b$ is a solution of the least squares problem. What is the set of all solutions of this problem?

The aim of this exercise is to determine to which solution gradient descent converges. Note that we cannot apply the method seen in TD 4 ex 2, as here the conditioning of the problem is equal to 0.

2. Denote $x^0, x^1, \dots, x^k, \dots$ the iterates of gradient descent with step-size α, starting from $x^0 = 0$. Demonstrate that there exists $u^k \in \mathbb{R}^n$ such that $x^k = A^\top u^k$ for all $k \geq 0$. What is the recursion satisfied by u^k?
3. Assume that $\alpha \leq \frac{1}{\lambda_{\max}(AA^\top)}$. Show that $\lim_{k \to +\infty} u^k = (AA^\top)^{-1} b$.
4. What is the limit of x^k? Show that this is the vector with the smallest norm among all minimizers of f.

Exercise 2.

1. Show that the set of minimizers of a convex function is a convex set.
2. Let $y \in \mathbb{R}$ and
$$f \colon x \in \mathbb{R} \mapsto \frac{1}{2}(x - y)^2 \qquad \text{and} \qquad g \colon (a, b) \in \mathbb{R}^2 \mapsto \frac{1}{2}(ab - y)^2.$$
   (a) What is the set of minimizers of f? Of g?
   (b) If f convex, coercive? Same question for g.
3. (a) Compute the gradient and Hessian of f and g.
   (b) Assume that y is positive. Give the equation of the region where g is locally convex, and draw it.
4. Let $Y \in \mathbb{R}^n$. Define
$$f \colon X \in \mathbb{R}^n \mapsto \frac{1}{2} \|X - Y\|^2 \qquad \text{and} \qquad g \colon (A, B) \in \mathbb{R}^n \times \mathbb{R}^n \mapsto \frac{1}{2}(AB - Y)^2.$$
   (a) What is the set of minimizers of f? Of g? If f convex, coercive? Same question for g.
   (b) Compute the gradient of f and g (represented as a matrix).

Exercise 3. Let $b \in \mathbb{R}^d$ and denote $\mathbf{1}$ the d-dimensional vector with all coordinates equal to 1. What is the solution (i.e. the minimizer) of

$$\min_{x \in \mathbb{R}} \|x\mathbf{1} - b\|$$

1. when $\|y\| := \sqrt{\sum_{i=1}^d y_i^2}$ (ℓ_2 norm)?
2. when $\|y\| := \max_{i=1}^d |y_i|$ (ℓ_∞ norm)?
3. when $\|y\| := \sum_{i=1}^d |y_i|$ (ℓ_1 norm)?

If you have finished all the exercises, you can move on to the TP5 on github.com/vcastin/teaching