Module:Optimization 1
–Tutorial 05 –

Algorithms for Optimization without constraind

In all the following exercises, A denotes a symmetric positive definite $N \times N$ matrix with eigenvalues $0 < \lambda_1 \leq \ldots \leq \lambda_N$, and $b \in \mathbb{R}^N$. For $v \in \mathbb{R}^N$, let $J(v) = \frac{1}{2}(Av, v) - (b, v)$. Let u be the point of minimum of J in $\mathbb{R}^N$.

Exercise 1 (Fixed Step-Size Gradient Algorithm)

1. Convergence of Fixed Step-Size Gradient Algorithm

Show that the fixed step-size gradient algorithm

$$u^{n+1} = u^n - \rho \nabla J(u_n)$$

converges for $0 < \rho < \frac{2}{\lambda_N}$.

2. Maximum Convergence Rate

Show that the maximum convergence rate is achieved for $\rho = \frac{2}{\lambda_1 + \lambda_N}$ and specify its value.

Exercise 2 (Optimal Step-Size Gradient Algorithm)

Consider the optimal step-size gradient algorithm:

$$u^{n+1} = u^n + \rho^n d^n,$$

where $d^n = -\nabla J(u^n)$, and ρ^n minimizes $\inf_{\rho \in \mathbb{R}} J(u^n + \rho d^n)$.

1. Show that for all n, $(d_{n+1}, d_n) = 0$. Calculate ρ_n.
2. Let $e^n = u^n - u$. Show that:

$$\|e^{n+1}\|_A^2 = \left(1 - (Ad^n, d^n)\left(\frac{(d^n, d^n)}{(e^n, d^n)}\right)\frac{(e^n, Ad^n)}{(e^n, d^n)}\right)\|e^n\|_A^2,$$

where $\|v\|_A^2 = (Av, v)$. Deduce that:

$$\|e^{n+1}\|_A \leq \left(\lambda_N - \frac{\lambda_1}{\lambda_N + \lambda_1}\right)\|e^n\|_A.$$

Exercise 3 (Conjugate Gradient Algorithm)

This involves the optimal step-size gradient descent algorithm:

$$u^{n+1} = u^n + \rho^n d^n \quad (2)$$

with

$$\begin{aligned} d^0 &= -\nabla J(u^0), \\ d^n &= -\nabla J(u^n)) + \frac{\|\nabla J(u^n))\|^2}{\|\nabla J(u^{n+1}))\|^2}, \\ \rho^n &= -(\nabla J(u^n), d^n)^n, d^n)(Ad^n, d^n) \end{aligned}$$

Recall that the principle of this method is as follows:

$$\text{starting from } u^0 \in \mathbb{R}^n, \text{ a sequence } (u^n) \text{ is constructed such that:}$$
$$u^{n+1} \in u^n + G^n \quad \text{and} \quad J(u^{n+1}) = \inf_{v \in u^n + G^n} J(v) \quad (3)$$

where $G^n = \text{vect}(\nabla J(u^0), \ldots, \nabla J(u^n))$.

1. Verify that
$$J(v) - J(u) = \frac{1}{2}\|v-u\|_A^2 \quad \forall v \in \mathbb{R}^N. \quad (4)$$

2. Show that the space G^n is generated by $\{\nabla J(u^0), A\nabla J(u^0), \ldots, A^n \nabla J(u^0)\}$.

3. For $k \geq 0$, set $e^k = u^k - u$.

   (a) For $v \in u^n + G^n$, show that there exists a polynomial Q of degree n such that $v = u^0 + Q(A)\nabla J(u^0)$.

   (b) Using 3 and (**??**), demonstrate that
$$\|e^{k+1}\|_A = \min_{P \in P_{n+1}} \|P(A)e^0\|_A. \quad (5)$$

4. Show that
$$\|P(A)e^0\|_A^2 \leq \|e^0\|_A^2 \max_i P^2(\lambda_i) \quad \forall P \in P_k. \quad (6)$$

   Conclude that
$$\|e_k\|_A \leq 2\left(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right)^n \|e^0\|_A, \quad \text{where } \kappa = \frac{\lambda_N}{\lambda_1} \quad (7)$$

5. Hint. $\min_{p \in P_k} \|p\|_{L^1([a,b])} = \frac{1}{T_k\left(\frac{a+b}{a-b}\right)}, where T_k$ denotes the Chebyshev polynomial of degree k. Moreover, we have
$$T_k\left(\frac{b}{a}\right) = T_k\left(\frac{a+b}{a-b}\right) = T_k\left(\frac{\frac{b}{a}-1}{\frac{b}{a}+1}\right) > \frac{1}{2}\left(\frac{\sqrt{\frac{b}{a}}+1}{\sqrt{\frac{b}{a}}-1}\right)^k, \text{for } b > a.$$