Continuous Optimization: Assignment 2

Due on May 7, 2024

Honglu Ma **Hiroyasu Akada** **Mathivathana Ayyappan**

Exercise 1

(i)

The gradient of $\varphi(x)$ is given by

$$\begin{aligned}\nabla\varphi(x) &= \frac{1}{2}(A + A^\top)x - b \\ &= \frac{1}{2}(2A)x - b && A \text{ is symetric} \\ &= Ax - b\end{aligned}$$

Similarly, the Hessian of $\varphi(x)$ is given by

$$\begin{aligned}\nabla^2\varphi(x) &= (D(\nabla\varphi(x)))^\top \\ &= (A^\top)^\top = A\end{aligned}$$

Suppose $\nabla\varphi(x^\star) = 0$ and $\nabla^2\varphi(x^\star)$ is positive definite, then Theorem 6.9 shows that $x^\star$ is a local minimum of $\varphi(x)$. To meet the condition, we only need to show that $\nabla\varphi(x^\star) = 0$ because A is positive definite which meets the second part of the condition. Thus we can find the minimizer by solving $Ax = b$.

(ii)

The steepest descent direction is given by

$$\begin{aligned}d^{(k)} &= -\nabla\varphi(x^{(k)}) \\ &= -Ax^{(k)} + b = r^{(k)}\end{aligned}$$

(iii)

Let $g_k(\tau) = x^{(k)} + \tau r^{(k)}$ be the function that gives $x^{(k+1)}$ given τ at time step k. We can rewrite the objective function as $\min\limits_{\tau>0} \varphi(g_k(\tau))$ and to find the minimizer, we can solve $\frac{\partial\varphi}{\partial\tau} = 0$.

$$\begin{aligned}\frac{\partial\varphi}{\partial\tau} &= \frac{\partial g_k}{\partial\tau} \cdot \frac{\partial\varphi}{\partial g_k} \\ &= (b - Ax^{(k)})^\top(Ag_k - b) \\ &= -b^\top b + (Ax^{(k)})^\top b + b^\top Ag_k - (Ax^{(k)})^\top Ag_k \\ &= -b^\top b + (Ax^{(k)})^\top b + b^\top A(x^{(k)} + \tau(b - Ax^{(k)})) - (Ax^{(k)})^\top A(x^{(k)} + \tau(b - Ax^{(k)})) \\ &= -b^\top b + 2(Ax^{(k)})^\top b - (Ax^{(k)})^\top Ax^{(k)} + \tau(b - Ax^{(k)})^\top A(b - Ax^{(k)}) \\ &= -(b - Ax^{(k)})^\top(b - Ax^{(k)}) + \tau(b - Ax^{(k)})^\top A(b - Ax^{(k)}) = 0 \\ &\Rightarrow \tau = \frac{(b - Ax^{(k)})^\top(b - Ax^{(k)})}{(b - Ax^{(k)})^\top A(b - Ax^{(k)})}\end{aligned}$$

Exercise 3

(a)

To show that f has a global minimizer at (a, a^2) we need to show that $f(a, a^2) \leq f(x_1, x_2)$ for all (x_1, x_2).

$$f(a, a^2) - f(x_1, x_2) = 0 - (a - x_1)^2 - b(x_2 - x_1^2)^2 \leq 0$$

(b)