Continuous Optimization: Assignment 12

Due on July 16, 2024

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

Exercise 1

1

(i) Not convex. Consider $p = 3$ we have $f(x) = x^3$ and $f''(x) = 6x \leq 0$ when $x \leq 0$.
(ii) Convex. $f''(x) = x^{-2} \geq 0$
(iii) Convex. $f''(x) = \alpha^2 e^{\alpha x} \geq 0$
(iv) Convex. $f''(x) = \frac{1}{(1-x)} \geq 0$ when $x \in (0, 1)$

2

(i) Convex.
Take $x, y \in C$.
Consider $||(1-\lambda)x + \lambda y||_2 \overset{Cauchy-Schwarz}{\leq} ||(1-\lambda)x||_2 + ||\lambda y||_2 = (1-\lambda)||x||_2 + \lambda||y||_2 \leq 1.$
(ii) Not convex. Consider $x = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, $y = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\lambda = \frac{1}{2}$.
Clearly $||(1-\lambda)x + \lambda y||_2 = \frac{\sqrt{2}}{2} \neq 1$.
(iii) Not convex. Consider $x = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $y = \begin{pmatrix} -1 \\ -1 \end{pmatrix}$ and $\lambda = \frac{1}{2}$.
Clearly $||(1-\lambda)x + \lambda y||_\infty = || \begin{pmatrix} 0 \\ 0 \end{pmatrix} ||_\infty = 0 \neq 1$.
(iv) Convex. Similar to (i).

3

4

$$T_C(\begin{pmatrix} 1 & 1 \end{pmatrix}^\top) = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2 \mid x_1 \leq 1, x_2 \leq 1 \right\}$$
$$N_C(\begin{pmatrix} 1 & 1 \end{pmatrix}^\top) = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2 \mid x_1 \geq 1, x_2 \geq 1 \right\}$$

Exercise 2

(A)

In steepest descent method, we have $x^{(k+1)} = x^{(k)} - \alpha \nabla f(x^{(k)})$. We can approximate $f(x^{(k+1)})$ with its second order Taylor expansion around $x^{(k)}$.

$$\begin{aligned} f(x^{(k+1)}) &\approx f(x^{(k)}) + \nabla f(x^{(k)})^\top (x^{(k+1)} - x^{(k)}) + \frac{1}{2}(x^{(k+1)} - x^{(k)})^\top \nabla^2 f(x^{(k)})(x^{(k+1)} - x^{(k)}) \\ &= f(x^{(k)}) - \alpha \nabla f(x^{(k)})^\top \nabla f(x^{(k)}) + \frac{\alpha^2}{2} \nabla f(x^{(k)})^\top \nabla^2 f(x^{(k)}) \nabla f(x^{(k)}) =: g(\alpha) \end{aligned}$$

Exercise 2 continued on next page...

We can find the optimal α by solving $\frac{dg}{d\alpha} = 0$.

$$\begin{aligned} \frac{dg}{d\alpha} &= -\nabla f(x^{(k)})^\top \nabla f(x^{(k)}) + \alpha \nabla f(x^{(k)})^\top \nabla^2 f(x^{(k)}) \nabla f(x^{(k)}) = 0 \\ \nabla f(x^{(k)})^\top \nabla f(x^{(k)}) &= \nabla f(x^{(k)})^\top (\alpha \nabla^2 f(x^{(k)})) \nabla f(x^{(k)}) \end{aligned}$$

By choosing $\alpha = (\nabla^2 f(x^{(k)}))^{-1}$, we can cancel the affect of rescaling of the Hessian matrix. Thus we arrive at the Newton's method: $x^{(k+1)} = x^{(k)} - (\nabla^2 f(x^{(k)}))^{-1} \nabla f(x^{(k)})$.

(B)

Step 0 Let $x^{(0)} := \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

$$\begin{aligned} d^{(0)} = r^{(0)} &= -\nabla f(x^{(0)}) \\ &= -b - Qx^{(0)} \\ &= -\begin{pmatrix} 1 \\ -1 \end{pmatrix} - \begin{pmatrix} 4 & 3 \\ 3 & 6 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\ &= \begin{pmatrix} -1 \\ 1 \end{pmatrix} \end{aligned}$$

$$\begin{aligned} \tau_0 &= \frac{\langle r^{(0)}, r^{(0)} \rangle}{\langle d^{(0)}, Qd^{(0)} \rangle} \\ &= \frac{1}{2} \end{aligned}$$

Step 1

$$\begin{aligned} x^{(1)} &= x^{(0)} + \tau_0 d^{(0)} \\ &= \begin{pmatrix} 0 \\ 0 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} -1 \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} -\frac{1}{2} \\ \frac{1}{2} \end{pmatrix} \end{aligned}$$

$$\begin{aligned} r^{(1)} &= r^{(0)} + \tau_0 Q d^{(0)} \\ &= \begin{pmatrix} -1 \\ 1 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} 4 & 3 \\ 3 & 6 \end{pmatrix} \begin{pmatrix} -1 \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} \frac{3}{2} \\ -\frac{5}{2} \end{pmatrix} \end{aligned}$$

$$\begin{aligned} \beta_1 &= \frac{\langle r^{(1)}, r^{(1)} \rangle}{\langle r^{(0)}, r^{(0)} \rangle} \\ &= \frac{17}{4} \end{aligned}$$

$$\begin{aligned} d^{(1)} &= -r^{(1)} + \beta_1 d^{(0)} \\ &= \begin{pmatrix} \frac{11}{4} \\ -\frac{7}{4} \end{pmatrix} \end{aligned}$$