Continuous Optimization: Assignment 9

Due on June 25, 2024

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

Exercise 1

The problem of finding the clostest point to another point can be formulated as

$$\min_{x \in \mathbb{R}^3} ||x-p||^2 \quad \text{s.t.} \quad ||x||^2 = 4$$

while the problem of finding the farthest point can be formulated as

$$\max_{x \in \mathbb{R}^3} ||x-p||^2 \quad \text{s.t.} \quad ||x||^2 = 4$$

which is equivalent to

$$\min_{x \in \mathbb{R}^3} -||x-p||^2 \quad \text{s.t.} \quad ||x||^2 = 4$$

where $p = (2,4,2)^\top$

Closest Point

We want the constraint levelset to be tangent to the curve of the objective function at the optimal point i.e.

where $f(x) = ||x-p||^2$ and $g(x) = ||x||^2$ satisfying $g(x) = 4$.
After calculating the gradients, we have

$$\begin{aligned} 2(x-p) &= 2\lambda x \\ x - p &= \lambda x \\ x &= \frac{1}{1-\lambda} p \end{aligned}$$

x still has to satisfy the constraint $||x||^2 = 4$.

$$\begin{aligned} \frac{1}{(1-\lambda)^2}||p||^2 &= 4 \\ \frac{24}{(1-\lambda)^2} &= 4 \\ \frac{6}{(1-\lambda)^2} &= 1 \\ 1-\lambda &= \pm\sqrt{6} \\ \lambda &= 1 \pm \sqrt{6} \end{aligned}$$

We have two solutions for x:

$$x = \pm\frac{1}{\sqrt{6}} p$$

The Langrange multiplier method gives us only the stationary points and we have to determine the minimum by checking which solution results in a smaller objective function value.

$$||x-p||^2 = \frac{1+\sqrt{6}}{-\sqrt{6}}||p||^2 \text{ or } \frac{1-\sqrt{6}}{\sqrt{6}}||p||^2$$

Apparently, the latter is smaller.

Thus, the closest point is $-\frac{1}{\sqrt{6}}\begin{pmatrix}2\\4\\2\end{pmatrix}$.

Exercise 1 continued on next page...

Farthest Point

Similarly, we have

$$\begin{aligned} 2(x-p) &= -2\lambda x \\ x-p &= -\lambda x \\ x &= \frac{1}{1+\lambda}p \end{aligned}$$

with constraint

$$\begin{aligned} \left|\left|\frac{1}{1+\lambda}p\right|\right|^2 &= 4 \\ \frac{24}{(1+\lambda)^2} &= 4 \\ \lambda &= -1 \pm \sqrt{6} \end{aligned}$$

We have two solutions for x:

$$x = \pm\frac{1}{\sqrt{6}}p$$

Which is the same as the ones we calculated for the closest point. Thus, the farthest point is $\frac{1}{\sqrt{6}}\begin{pmatrix}2\\4\\2\end{pmatrix}$.

Exercise 2

The minimization problem is given by

$$\min_{x\in\mathbb{R}^n} ||x||^2 \quad \text{s.t.} \quad Ax = b$$

It is equivalent to

$$\min_{x\in\mathbb{R}^n} \frac{1}{2}||x||^2 \quad \text{s.t.} \quad Ax = b$$

Let $f(x) = \frac{1}{2}||x||^2$ and $c_i(x) = a_i^\top x$ where a_i is the i-th row of A.
The constraint $Ax = b$ can be rewritten as m smaller constraints: $c_i(x) = b_i$ for $i = 1, \ldots, m$.
Using the Lagrange multiplier method, we compose such equation:

$$\begin{aligned} \nabla f(x) &= \sum_{i=1}^{m} \lambda_i \nabla c_i(x) \\ x &= \sum_{i=1}^{m} \lambda_i a_i \\ x &= A^\top \lambda \end{aligned}$$

where λ_i is the Lagrange multiplier for the i-th constraint and λ is a column vector consists of all multipliers.
We also have the constraint level sets:

$$\begin{aligned} Ax &= b \\ AA^\top\lambda &= b \\ \lambda &= (AA^\top)^{-1}b \\ \Rightarrow x &= A^\top(AA^\top)^{-1}b \end{aligned}$$

Exercise 3

We have the following optimization problem:

$$\min_{x \in \mathbb{R}^n} f(x) \quad \text{s.t.} \quad g(x) = 1$$

with $f(x) = c^\top x + \frac{1}{2\tau}||x - \bar{x}||^2$ and $g(x) = \sum_{i=1}^n x_i$
where x_i is the i-th element of x.
Using the Lagrange multiplier method, we have

$$\begin{aligned} \nabla f(x) &= \lambda \nabla g(x) \\ c + \frac{1}{\tau}(x - \bar{x}) &= \lambda \mathbf{1} && \text{observe that } \nabla g(x) = \mathbf{1} \text{ which is a column vector of 1s} \\ x &= \tau(\lambda \mathbf{1} - c) + \bar{x} \end{aligned}$$

Plug x back into the constraint $g(x) = 1$, we have

$$\begin{aligned} g(\begin{pmatrix} \tau(\lambda - c_i) + \bar{x}_i \\ \vdots \\ \tau(\lambda - c_n) + \bar{x}_n \end{pmatrix}) &= 1 \\ \sum_{i=1}^n \tau(\lambda - c_i) + \bar{x}_i &= 1 \\ n\tau\lambda - \tau \sum_{i=1}^n c_i + \sum_{i=1}^n \bar{x}_i &= 1 \\ n\tau\lambda - \tau g(c) + g(\bar{x}) &= 1 \\ \lambda &= \end{aligned}$$

Now that λ is determined, we can calculate x with the formula we derived earlier

$$\begin{aligned} x &= \tau(\frac{1 + \tau g(c) - g(\bar{x})}{n\tau}\mathbf{1} - c) + \bar{x} \\ &= (\frac{1 + \tau g(c) - g(\bar{x})}{n}\mathbf{1} - c) + \bar{x} \end{aligned}$$

Exercise 4

(a)

(b)

(c)

We know that

$$\lambda_1 = \min_{x \in \mathbb{R}^n} f(x) \quad \text{s.t.} \quad c_1(x) = 1$$

Exercise 4 continued on next page...

where $f(x) = \langle x, Qx \rangle$ and $c_1(x) = ||x||^2$.
Apply the Lagrange multiplier method, we have

$$\begin{aligned} \nabla f(x) &= \tilde{\lambda}_1 \nabla c_1(x) \\ Qx &= \tilde{\lambda}_1 x \end{aligned}$$

with Lagrange multiplier $\tilde{\lambda}_1$ and constraint $c_1(x) = 1$.
One can interpret this as $\tilde{\lambda}_1$ being the eigenvalue of Q and x being the eigenvector.

Exercise 5

(a)

We can find the projection of $\bar{x}$ onto set $C = \{x \in \mathbb{R}^n \mid \langle a, x \rangle \leq b\}$ by solving a constrained optimization problem:

$$\min_{x \in \mathbb{R}^n} ||x - \bar{x}||^2 \quad \text{s.t.} \quad \langle a, x \rangle = b$$

We have the following two conditions:

1. $\hat{x} \in C$
2. $\hat{x} \notin C$

By Fermat's Rule, we know that $-\nabla f(\hat{x}) = N_C(\hat{x})$.
If $\hat{x} \in C$, then $N_C(\hat{x}) = \{0\}$. We have

$$\begin{aligned} -\nabla f(\hat{x}) &= 0 \\ \bar{x} - \hat{x} &= 0 \\ \hat{x} &= \bar{x} \end{aligned}$$

If $\hat{x} \notin C$, then $N_C(\hat{x}) = \{\hat{x} + t\frac{a}{||a||} \mid t \geq 0\}$. We have

$$\begin{aligned} -\nabla f(\hat{x}) &= \hat{x} + t\frac{a}{||a||} \\ \bar{x} - \hat{x} &= \hat{x} + t\frac{a}{||a||} \\ \hat{x} &= \frac{1}{2}(\bar{x} - t\frac{a}{||a||}) \end{aligned}$$

We also know that $\hat{x} \in C$ i.e.

$$\begin{aligned} \langle a, \hat{x} \rangle &= b \\ \langle a, \frac{1}{2}(\bar{x} - t\frac{a}{||a||}) \rangle &= b \\ \langle a, \bar{x} \rangle - t||a|| &= 2b \\ t &= \frac{\langle a, \bar{x} \rangle - 2b}{||a||} \end{aligned}$$

Plug t back into the formula of $\hat{x}$, we have

$$\hat{x} = \bar{x} - \frac{\langle a, \bar{x} \rangle}{||a||^2} a$$

In both condition $\hat{x}$ is expressed as a function of $\bar{x}$ which shows that the projection is singleton.