TD4

Friday, November 13, 2020 1:55 PM

UNIVERSITE DE GENEVE
Faculté des sciences

Section de Mathématiques

Optimization with Application I
Exercise Sheet 4 - Discussed on 13.11.2020

Exercise 1. Let X be an $N \times P$ matrix with $P > N$ of rank N and $y \in \mathbb{R}^N$. Using a Lagrange multiplier $\lambda \in \mathbb{R}^N$ show that

$$\min_{\alpha \in \mathbb{R}^P} ||\alpha||_2^2 \text{ s.t. } y = X\alpha$$

has a closed form expression. Comment on the dimension of the matrix that needs to be inverted.

Exercise 2. Consider the problem of minimizing

$$f(x, y) = x + y$$

under the constraints

$$h_1(x, y) = (x-1)^2 + y^2 - 1 = 0$$

and

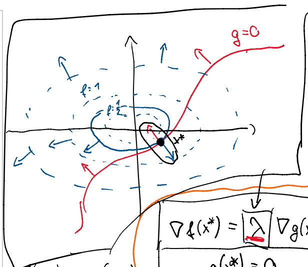
$$h_2(x, y) = (x-2)^2 + y^2 - 4 = 0.$$

Draw the level lines of the cost function and the constraint sets. Identify the unique feasible solution.

Show that this point is not a regular point and that there exists no Lagrange multipliers for this problem.

Exercise 3. Use the Lagrange multiplier theorem to solve $\min f(x)$ subject to $h(x) = 0$ with

1. $f(x) = ||x||_2^2, h(x) = \sum_{p=1}^P x_p - 1$
2. $f(x) = \sum_{p=1}^P x_p, h(x) = ||x||_2^2 - 1$.



$f(x,y)$, $g(x,y)$, $f, g \in C^1$

$\arg\min_{x,y} f(x,y)$ s.t. $\boxed{g(x,y) = 0}$

$\nabla \mathcal{L} = 0$

$\lambda \in \mathbb{R}$

$$\boxed{\mathcal{L}(x,y,\lambda) = f(x,y) - \lambda g(x,y)}$$

$$\rightarrow \nabla \mathcal{L}(x,y,\lambda) = \begin{pmatrix} \nabla f - \lambda \nabla g \\ g(x,y) \end{pmatrix} \begin{matrix} \}\mathbb{R}^2 \\ \}\mathbb{R} \end{matrix} \rightarrow = 0$$

$$\boxed{\begin{matrix} \nabla f(x^*) = \boxed{\lambda} \nabla g(x^*) \\ g(x^*) = 0 \end{matrix}} \Leftarrow \exists \lambda \text{ s.t.}$$

$\Leftrightarrow \nabla \mathcal{L}(x,y,\lambda) = 0$

Thm: $f: \mathbb{R}^n \to \mathbb{R}$, $g: \mathbb{R}^n \to \mathbb{R}^c$, $f, g \in C^1$

Let $x^* \in \mathbb{R}^n$ be optimal solution to $[\min_x f(x) \text{ s.t. } g(x) = 0]$

and $\operatorname{rank}(Dg(x^*)) = \boxed{c < n}$

Then: $\exists! \lambda^* \in \mathbb{R}^c$ s.t. $\boxed{Df(x^*) = \lambda^{*T} Dg(x^*)}$

$\rightarrow$ necessary condition for x^* (equivalent to the gradient test $\boxed{\nabla f(x) = 0}$)

$\operatorname{rank}(D(g(x^*))) < n$ ✓

$\# 2 = c$

Ex 1: $f(\alpha) = ||\alpha||_2^2$, $g(\alpha) = X\alpha - y$, $X \in \mathbb{R}^{N \times P}$, $N < P$, $\operatorname{rank}(X) = N$

$\mathcal{L}(\alpha, \lambda) = ||\alpha||_2^2 - \lambda^T (X\alpha - y)$ $\in \mathbb{R}$

$\nabla_\alpha \mathcal{L}(\alpha, \lambda) = 2\alpha - (\lambda^T X)^T = 2\alpha - X^T \lambda = 0$

$\nabla_\lambda \mathcal{L}(\alpha, \lambda) = -X\alpha + y$

P

N

$\nabla \mathcal{L}(\alpha, \lambda) \stackrel{!}{=} 0$

$\underbrace{2X\alpha}_{=2y} - XX^T\lambda = 0$

$2y - XX^T\lambda = 0 \Rightarrow \boxed{\lambda = 2(XX^T)^{-1} y}$

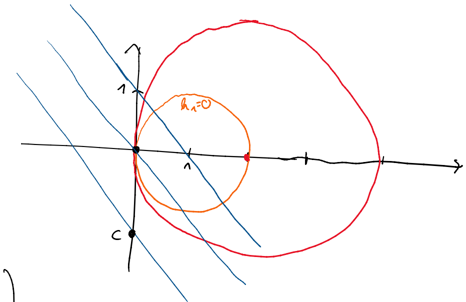
$\boxed{\nabla \mathcal{L} = 0} \Rightarrow \boxed{\alpha = X^T \underbrace{(XX^T)^{-1}}_{} y}$ $\rightarrow$ satisfies $\nabla \mathcal{L}(\alpha, \lambda) = 0$ $\rightarrow$ necessary condition satisfied

$N \times N$ of rank N $\Rightarrow (XX^T)^{-1}$ exists

Ex 2: $f(x,y) = x + y$

$h_1(x,y) = (x-1)^2 + y^2 - 1 = 0$ $\rightarrow B(1,1)$

$h_2(x,y) = (x-2)^2 + y^2 - 4 = 0$ $\rightarrow B(2,2)$



$x^2 + y^2 - 1 = 0$

$f(x,y) = c$

$\Leftrightarrow y = -x + c$

$\mathcal{L}(x,y,\lambda_1,\lambda_2) = f(x,y) - (\lambda_1, \lambda_2) \begin{pmatrix} h_1(x,y) \\ h_2(x,y) \end{pmatrix}$

$$\nabla \mathcal{L}(x,y,\lambda_1,\lambda_2) = \begin{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} - \begin{pmatrix} 2(x-1) & 2(x-2) \\ 2y & 2y \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} \\ h_1(x,y) \\ h_2(x,y) \end{pmatrix}$$

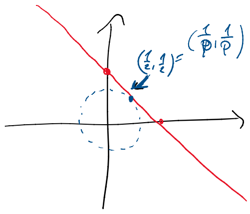
$\rightarrow \begin{pmatrix} 1 \\ 1 \end{pmatrix} - \begin{pmatrix} -2 & -4 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = \begin{pmatrix} 1 + 2\lambda_1 + 2\lambda_2 \\ 1 \end{pmatrix} \neq 0 \quad \forall \lambda_1, \lambda_2$

$= 0$

$\Rightarrow (x,y) = 0$

Ex 3: $f(x) = ||x||_2^2$, $g(x) = \sum_{p=1}^P x_p - 1$

$$\nabla \mathcal{L}(x, \lambda) = \begin{pmatrix} 2x - \lambda \cdot \mathbb{1} \\ \Sigma x_p - 1 \end{pmatrix} \begin{matrix} \}\mathbb{R}^P \\ \}\mathbb{R} \end{matrix} = 0$$



$\Rightarrow x = \frac{\lambda}{2} \mathbb{1}$, $\Sigma x_p = 1$

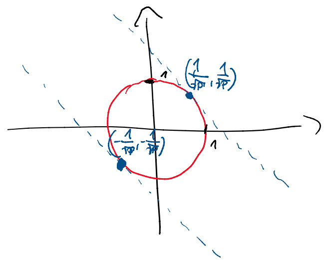
$\Rightarrow x_p = \frac{1}{P} \quad \forall p \leq P$

$\Rightarrow x = \frac{1}{P} \mathbb{1}$

$f(x) = \sum_{p=1}^P x_p$, $g(x) = ||x||_2^2 - 1$

$f(x) = \sum_{p=1}^{P} x_p$, $g(x) = \|x\|_2^2 - 1$

$$\nabla \mathcal{L}(x,\lambda) = \begin{pmatrix} \mathbb{1} - \lambda 2x \\ \|x\|_2^2 - 1 \end{pmatrix} = 0$$



$\Rightarrow x = \underbrace{\frac{1}{2\lambda}} \mathbb{1}$ $\leftarrow$, $\|x\|_2^2 = 1$

$$\sum_{p=1}^{P} \left(\frac{1}{2\lambda}\right)^2 = P \left(\frac{1}{2\lambda}\right)^2 = 1$$

$$\Leftrightarrow \lambda^2 = \frac{P}{4} \Rightarrow \underbrace{\lambda} = \pm \frac{\sqrt{P}}{2}$$

$\Rightarrow x = \pm \frac{1}{\sqrt{P}} \cdot \mathbb{1}$

$f(\pm \frac{1}{\sqrt{P}} \mathbb{1}) = \pm \frac{P}{\sqrt{P}} = \pm \sqrt{P}$ $\rightsquigarrow$ two points satisfy necessary condition.

$\rightsquigarrow x^* = -\frac{1}{\sqrt{P}} \cdot \mathbb{1}$