Continuous Optimization **Srping 2023**

Homework 4: Constrained Optimization

Lecturer: Aurelien Lucchi, Student: Julian Bopp

Homework 1 (Constrained problem):

Consider the following 2-dimensional problem

$$\begin{aligned} \min f(x,y) &\coloneqq x(1-y^2) \\ \text{s.t. } x^2+y^2 &= 1. \end{aligned}$$

1. Write the stationary and primal feasibility conditions.

   Stationary conditions: Define the lagrangian

   $$\begin{aligned} L(x,y,v) &= f(x,y) + vl(x,y) \\ &= x(1-y^2) + v(x^2+y^2-1), \end{aligned}$$

   where $l(x,y) \coloneqq x^2+y^2-1$ encodes the equality constraint. Find the derivative's w.r.t x and y

   $$\begin{aligned} \partial_x L(x,y,v) &= 1-y^2+2vx \\ \partial_y L(x,y,v) &= -2yx+2vy. \end{aligned}$$

   Now the stationary conditions are

   $$\begin{aligned} 1-y^2+2vx &= 0 \\ -2yx+2vy &= 0. \end{aligned}$$

   Primal feasibility condition: The primal feasibily condition is that the constraint is satisfied, i.e. that

   $$l(x,y) \coloneqq x^2+y^2-1=0$$

2. Derive the optimal solution (x^*, y^*).

   Solving the system of two equations from the stationary conditons we get

   $$\left\{(-\frac{1}{2v},0),(v,1-2v^2)\right\}$$

   as a set of possible solutions. Combining this with the primal feasibility condition yields

   $$\{(-1,0),(1,0),(0,-1),(0,1)\}$$

   as a set of possible solutions. We see that

   $$\begin{aligned} f(-1,0) &= -1 \\ f(1,0) &= 1 \\ f(0,-1) &= 0 \\ f(0,1) &= 0, \end{aligned}$$

   and conclude that $(x^*,y^*) \coloneqq (-1,0)$ solves the constrained problem.

Homework 2 (KKT problem with two constraints):

Consider the following 3-dimensional problem

$$\begin{aligned} \min f(x,y,z) &\coloneqq x+y+z \\ \text{s.t. } x^2-y^2 = 1 &\text{ and } 2x+z-1=0. \end{aligned}$$

1. Write the stationary and primal feasibility conditions.

   Stationary conditions: Define the lagrangian

$$\begin{aligned} L(x, y, z, v_1, v_2) &= f(x, y, z) + v_1 l_1(x, y, z) + v_2 l_2(x, y, z) \\ &= x + y + z + v_1(x^2 - y^2 - 1) + v_2(2x + z + 1), \end{aligned}$$

   where $l_1(x, y, z) := x^2 - y^2 - 1$ and $l_2(x, y, z) := 2x + z + 1$ encode the equality constraints. Find the derivative's w.r.t x and y and z

$$\begin{aligned} \partial_x L(x, y, z, v_1, v_2) &= 1 + 2v_1 x + 2v_2 \\ \partial_y L(x, y, z, v_1, v_2) &= 1 - 2v_1 y \\ \partial_z L(x, y, z, v_1, v_2) &= 1 + v_2. \end{aligned}$$

   Now the stationary conditions are

$$\begin{aligned} 1 + 2v_1 x + 2v_2 &= 0 \\ 1 - 2v_1 y &= 0 \\ 1 + v_2 &= 0. \end{aligned}$$

   Primal feasibility conditions: The primal feasibily conditions are that the equality constraints are satisfied, i.e. that

$$\begin{aligned} l_1(x, y, z) &:= x^2 - y^2 - 1 = 0 \\ l_2(x, y, z) &:= 2x + z - 1 = 0 \end{aligned}$$

2. Derive all the optimal solutions.

   From the the 3rd equality of the stationary conditions we get $v_2 = -1$ and with that we get $x = y = \frac{1}{2v_1}$. Combining this with the first primal feasibility condition we geometry

$$\left(\frac{1}{2v_1}\right)^2 - \left(\frac{1}{2v_1}\right)^2 - 1 = 0 \implies -1 = 0,$$

   which is a contradiction. Therefore no optimal solution exists.

3. Can you comment on the results?

   There exists no solution because f is not bounded from below on the constraints. To see this let $x = k \geq 1$. To satisfy the first equality constraint we pick $y = -\sqrt{k^2 - 1}$ and to satisfy the second equality constraint we let $z = 1 - 2k$. We then get $f(x, y, z) = -k - \sqrt{k^2 - 1} + 1$, which goes to $-\infty$, as k goes to ∞.

Homework 3 (Projection onto hyperplane):

Consider the projection of a vector onto a hyperplane identified by the equation $\mathbf{Ax} = \mathbf{b}$, i.e.

$$\mathbf{x} = \mathrm{Proj}_{\mathbf{Ax}=\mathbf{b}}(\mathbf{y}) = \underset{\mathbf{x}:\mathbf{Ax}=\mathbf{b}}{\mathrm{argmin}} \frac{1}{2} \|\mathbf{y} - \mathbf{x}\|^2 . \tag{1}$$

1. Write down the Lagrangian corresponding to the constrained problem defined in Eq. (1).

   We want to minimize the function $f(\mathbf{x}) = \frac{1}{2} \|\mathbf{y} - \mathbf{x}\|^2$ given the equality constraint $\mathbf{Ax} = \mathbf{b}$. Let $\mathbf{A} \in \mathbb{R}^{m \times n}$, therefore $\mathbf{x} \in \mathbb{R}^n$ and $\mathbf{b} \in \mathbb{R}^m$. We can encode the equality constraint $\mathbf{Ax} = \mathbf{b}$ as a vector valued function

$$l : \mathbb{R}^n \to \mathbb{R}^m, \; x \mapsto Ax - b.$$

   Now let $\mathbf{v} = (v_1, \ldots, v_m)^\top$ and define the lagrangian

$$\begin{aligned} L(\mathbf{x}, \mathbf{v}) &= f(\mathbf{x}) + \mathbf{v}^\top l(\mathbf{x}) \\ &= \frac{1}{2} \|y - x\|^2 + \mathbf{v}^\top (\mathbf{Ax} - \mathbf{b}). \end{aligned}$$

2. Calculate the optimal value of $\mathbf{x}$ (using the KKT conditions). Show that

$$\mathbf{x} = \mathbf{P}\mathbf{y} + \mathbf{A}^\top(\mathbf{A}\mathbf{A}^\top)^{-1}\mathbf{b},$$

where $\mathbf{P} := \mathbf{I} - \mathbf{A}^\top(\mathbf{A}\mathbf{A}^{\top^{-1}}\mathbf{A})$ is a projection matrix.

Stationary conditions: Find the derivative of the lagrangian w.r.t $\mathbf{x}$

$$\partial_{\mathbf{x}} L(\mathbf{x}, \mathbf{v}) = \mathbf{x} - \mathbf{y} + A^\top v.$$

Therefore the stationary condition is $\mathbf{x} - \mathbf{y} + \mathbf{A}^\top\mathbf{v} = \mathbf{0}$, i.e. $\mathbf{x} = \mathbf{y} - \mathbf{A}^\top\mathbf{v}$.

Primal feasibility conditions The primal feasibility conditions are our equality constraints $\mathbf{A}\mathbf{x} = \mathbf{b}$. Using the result obtained from the stationary conditions we get

$$\begin{aligned} &\mathbf{A}(\mathbf{y} - \mathbf{A}^\top\mathbf{v}) = \mathbf{b} \\ \implies &\mathbf{A}\mathbf{y} - \mathbf{A}\mathbf{A}^\top\mathbf{v} = \mathbf{b} \\ \implies &\mathbf{A}\mathbf{A}^\top\mathbf{v} = \mathbf{A}\mathbf{y} - \mathbf{b} \\ \implies &\mathbf{v} = (\mathbf{A}\mathbf{A}^\top)^{-1}(\mathbf{A}\mathbf{y} - \mathbf{b}) \end{aligned}$$

Plugging this into our expression for $\mathbf{x}$ we get

$$\begin{aligned} \mathbf{x} &= \mathbf{y} - \mathbf{A}^\top(\mathbf{A}\mathbf{A}^\top)^{-1}(\mathbf{A}\mathbf{y} - \mathbf{b}) \\ &= \underbrace{(\mathbf{I} - \mathbf{A}^\top(\mathbf{A}\mathbf{A}^\top)^{-1}\mathbf{A})}_{\mathbf{P}:=}\mathbf{y} + \mathbf{A}^\top(\mathbf{A}\mathbf{A}^\top)^{-1}\mathbf{b}, \end{aligned}$$

which is what was to be shown.

Homework 4 (Normal cones):

Consider the following two sets:

$$\Omega_\infty := \left\{\mathbf{x} \in \mathbb{R}^d : \|\mathbf{x}\|_\infty \le 1\right\},$$

and

$$\Omega_2 := \left\{\mathbf{x} \in \mathbb{R}^d : \|\mathbf{x}\|_2 \le 1\right\},$$

1. Show that Ω_∞ and Ω_2 are non-empty, convex and closed.

These results can be shown independent of the norm used. We use Ω and $\|\cdot\|$ for both sets and norms.

Non-empty: $\|\mathbf{0}\| = 0$ and therefore $\mathbf{0} \in \Omega$.

Convex: Let $\mathbf{a}, \mathbf{b} \in \Omega$ and $\lambda \in (0, 1)$. Ω is convex if $\lambda\mathbf{a} + (1 - \lambda)\mathbf{b} \in \Omega$.

$$\begin{aligned} \|\lambda\mathbf{a} + (1 - \lambda)\mathbf{b}\| &\le \lambda\|\mathbf{a}\| + (1 - \lambda)\|\mathbf{b}\| \\ &\le \lambda + 1 - \lambda \\ &= 1. \end{aligned}$$

Therefore $\lambda\mathbf{a} + (1 - \lambda)\mathbf{b} \in \Omega$ and Ω is convex. For the first inequality we used the triangle inequality and for the second inequality we used the fact that $\mathbf{a}, \mathbf{b} \in \Omega$ and therefore have norm less than 1.

Closed: Let $\Omega^c = \mathbb{R}^d \setminus \Omega$. To show that Ω is closed we show that Ω^c is open. To show that a set is open we must show that for every element in the set there exists an open ball that is completely contained in the set. Let $\mathbf{x} \in \Omega^c$, this implies $\|\mathbf{x}\| > 1$. Let $\epsilon = \|\mathbf{x}\| - 1$ and define the open ball $B(\mathbf{x}, \epsilon)$ with center $\mathbf{x}$ and radius ϵ. Let $\mathbf{y} \in B(\mathbf{x}, \epsilon)$. Then by the definition of an open ball we have $\|\mathbf{x} - \mathbf{y}\|_2 < \epsilon$ (here I use the euclidean norm because I assume that we want to show that Ω_∞, Ω_2 are both open in $(\mathbb{R}^d, \|\cdot\|_2)$. Now we show that $\mathbf{y} \in \Omega^c$. First, by the triangle inequality

$$\|\mathbf{x}\| = \|\mathbf{x} - \mathbf{y} + \mathbf{y}\| \le \|\mathbf{y}\| + \|\mathbf{x} - \mathbf{y}\|.$$

From this we get

$$\begin{aligned} \|\mathbf{y}\| &\ge \|\mathbf{x}\| - \|\mathbf{x} - \mathbf{y}\| \\ &\overset{(*)}{\ge} \|\mathbf{x}\| - \|\mathbf{x} - \mathbf{y}\|_2 \\ &> 1 + \epsilon - \epsilon \\ &= 1. \end{aligned}$$

($*$) This inequality is trivial if $\|\cdot\| = \|\cdot\|_2$, but if $\|\cdot\| = \|\cdot\|_\infty$ we use the result that $\|\mathbf{x}\|_\infty \leq \|\mathbf{x}\|_2 \ \forall \mathbf{x} \in \mathbb{R}^d$. This shows that for every $\mathbf{x} \in \Omega^c$ there exists an $\epsilon > 0$ such that the open ball centered at $\mathbf{x}$ with radius ϵ is contained in Ω^c. Therefore Ω^c is an open set and hence Ω is a closed set.

2. Determine the normal cones of Ω_∞ and Ω_2 for $d = 2$ at the point $\mathbf{x} = (1, 0)^\top$.

From the lecture we know that $N_{\Omega_2} = \left\{ \begin{pmatrix} \lambda \\ 0 \end{pmatrix} : \lambda > 0 \right\}$, independent of the point. Therefore

$$N_{\Omega_2}(\mathbf{x}) = \left\{ \begin{pmatrix} \lambda \\ 0 \end{pmatrix} : \lambda > 0 \right\}$$

For $N_{\Omega_\infty}(\mathbf{x})$ we use the definition

$$N_{\Omega_\infty}(\mathbf{x}) = \left\{ \mathbf{w} : \langle \mathbf{w}, \mathbf{y} - \begin{pmatrix} 1 \\ 0 \end{pmatrix} \rangle \leq 0, \ \forall \mathbf{y} \in \Omega_\infty \right\}$$

This means that a point $\mathbf{w} = (w_1, w_2)^\top$ in $N_{\Omega_\infty}(\mathbf{x})$ has to satisfy the following inequality for any $\mathbf{y} = (y_1, y_2)$ in Ω_∞

$$w_1(y_1 - 1) + w_2 y_2 \leq 0.$$

In particular for $\mathbf{y}_+ = (1, 1)$ and $\mathbf{y}_- = (1, -1)$ we get

$$w_2 \leq 0, \ w_2 \geq 0 \implies w_2 = 0$$

Therefore the constraints on w_1 are

$$w_1(y_1 - 1) \leq 0. \tag{2}$$

Since $\mathbf{y} \in \Omega_\infty$ we know that $y_1 \in [0, 1]$ and therefore $(y_1 - 1) \leq 0$. From this and Eq. (2) we conclude that $w_1 \geq 0$ and therefore

$$N_{\Omega_\infty}(\mathbf{x}) = \mathbb{R}_0^+ \times \{0\}.$$