ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Institute for Operations Research
ETH Zurich HG G21-22

Kevin Zemmer kevin.zemmer@ifor.math.ethz.ch
Christoph Glanzer christoph.glanzer@ifor.math.ethz.ch

IFOR
Institute for Operations Research
Department of Mathematics

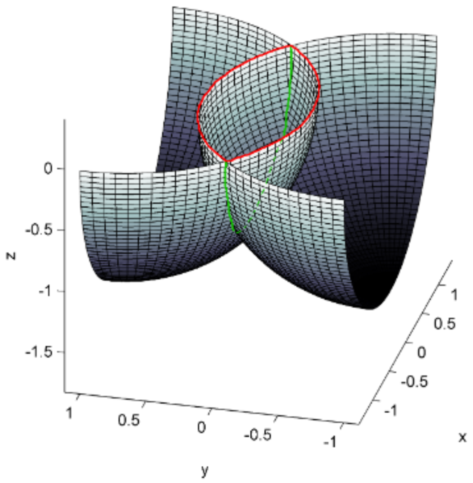
Mathematical Optimization — Assignment 8

https://moodle-app2.let.ethz.ch/course/view.php?id=3610

Exercise 1: Subgradients

$$\begin{aligned} f_1(x,y) &= -\sqrt{2-x^2-(y-1)^2}, & \mathrm{dom} f_1 &= \{(x,y) \mid x^2+(y-1)^2 \le 2\}, \\ f_2(x,y) &= -\sqrt{2-x^2-(y+1)^2}, & \mathrm{dom} f_2 &= \{(x,y) \mid x^2+(y+1)^2 \le 2\}, \\ f(x,y) &= \max\{f_1(x,y), f_2(x,y)\}, & \mathrm{dom} f &= \mathrm{dom} f_1 \cap \mathrm{dom} f_2 \end{aligned}$$

Let us first visualize the situation with a picture ($\mathrm{dom} f$ is the area delimited by the red curve, projected onto the xy-plane):



We recall a rule for computing subgradients of the maximum of convex functions seen in the lecture:

$$f(x) := \max_i f_i(x) \Longrightarrow \partial f(x_0) = \mathrm{conv}\{\partial f_i(x_0) \mid f_i(x_0) = f(x_0)\}, \forall x_0 \in \bigcap_i \mathrm{dom}(f_i).$$

- The subgradient of f at the boundary points $\{(x,y) \in \mathrm{dom} f \mid f(x,y) = 0\}$ (red line in the picture above) is empty, since any supporting hyperplane of f at those points should be vertical, while the subgradient is constructed using the non-vertical supporting hyperplanes of $\mathrm{epi}(f)$.
- The subgradient of f at the points $\{(x,y) \in \mathrm{dom} f \mid y \neq 0, f(x,y) \neq 0\}$ is a singleton, since the function there is differentiable (it consists of f_1 or f_2). The subdifferential of f is then equal to the differential of f_1 resp. f_2 (obtained by computing $\partial f_i(x,y)/\partial x$ and $\partial f_i(x,y)/\partial y$):

  If $y < 0$ (i.e. $f = f_1$): $\nabla f(x,y) = \frac{1}{\sqrt{2-x^2-(y-1)^2}} \binom{x}{y-1}$.

  If $y > 0$ (i.e. $f = f_2$): $\nabla f(x,y) = \frac{1}{\sqrt{2-x^2-(y+1)^2}} \binom{x}{y+1}$.
- The subgradient of f at the points $\{(x,0) \in \mathrm{dom} f \mid f(x,y) \neq 0\} = \{(x,0) \mid -1 < x < 1\}$ (i.e. where f is given by both f_1 and f_2, the green line in the picture above) is, according to the rule above,

the convex hull of the (sub)gradients of f_1 and f_2:

$$\begin{aligned}\partial f(x,0) = & \ \operatorname{conv}\left\{\frac{1}{\sqrt{1-x^2}}\begin{pmatrix}x\\-1\end{pmatrix}, \frac{1}{\sqrt{1-x^2}}\begin{pmatrix}x\\+1\end{pmatrix}\right\}\\ = & \left\{\lambda\frac{1}{\sqrt{1-x^2}}\begin{pmatrix}x\\-1\end{pmatrix} + (1-\lambda)\frac{1}{\sqrt{1-x^2}}\begin{pmatrix}x\\+1\end{pmatrix}\,\Big|\,\lambda\in[0,1]\right\}\\ = & \left\{\frac{1}{\sqrt{1-x^2}}\begin{pmatrix}x\\t\end{pmatrix}\,\Big|\,t\in[-1,+1]\right\}.\end{aligned}$$

Exercise 2: Separation of Convex Sets

(a) We consider only the case of strong separation, the case of separation is similar, but simpler.

Proof. To prove the convexity of $C-D$, let $x-y$, $z-w\in C-D$, where $x,z\in C$, $y,w\in D$ and $0\le\lambda\le 1$. Then

$(1-\lambda)(x-y)+\lambda(z-w) = (1-\lambda)x+\lambda z-(1-\lambda)y+\lambda w\in C-D$

by convexity of C and D. Thus $C-D$ is convex.

(i) $\Rightarrow$ *(ii)*

Let $S=\{x:\alpha\le u\cdot x\le\beta\}$, $\alpha<\beta$, be a set which strongly separates C and D, say $C\subseteq\{x: u\cdot x\le\alpha\}$ and $D\subseteq\{y:u\cdot y\ge\beta\}$. Then $C-D\subseteq\{x-y:u\cdot(x-y)\le\alpha-\beta\}$. Thus $C-D$ and $\{0\}$ are separated by the set $\{z:\alpha-\beta\le u\cdot z\le 0\}$.

(ii) $\Rightarrow$ *(i)*

Let the set $S=\{z:-\gamma\le u\cdot z\le 0\}$, $\gamma>0$, separate $\{0\}$ and $C-D$. Then $u\cdot(x-y)\le-\gamma$, i.e. $u\cdot x+\gamma\le u\cdot y$ for all $x\in C$ and $y\in D$. Let $\alpha = sup\{u\cdot x: x\in C\}$. Then the set $\{z:\alpha\le u\cdot z\le\alpha+\gamma\}$ separates C and D.

□

(b) *Proof.* Choose $p\in C, q\in D$ having minimum distance. Let $u=q-p(\neq 0)$. Then the set $\{x:u\cdot p\le u\cdot x\le u\cdot q\}$ separates C and D.

□

Exercise 3: Equality constrained least squares

Recall the KKT-conditions from the lecture for $\min\{f(x)\mid g(x)\le b\}$ and $x^*\in\operatorname{dom}(f)$:

There exist $\lambda_i^*\ge 0$ for all $i=1,\dots,m$ such that

- $\lambda_i^*=0$ for all $i\notin I(x^*)=\{j\colon g_i(x^*)=b_i\}$
- $\nabla f(x^*)+\sum_{i=1}^m\lambda_i^*\nabla g_i(x^*)=0.$

In this example, these conditions are

$$2A^T(Ax^*-b)+G^T\nu^*=0 \qquad Gx^*=h.$$