Assignment 2

Convex Optimization SS 2018

Patrick Knöbelreiter, knoebelreiter@icg.tugraz.at

23.04.2018

Submission: Upload your solution as a PDF and your implementation as PY files to the TeachCenter. You can also use your scanned hand-written manuscript.

Deadline: May 14th, 2018 at 23:55h.

Subgradients

Determine the subgradients of the following functions:

1. $$f(x) = \begin{cases} 0 & \text{for } x \in C \\ +\infty & \text{for } x \notin C, \end{cases} \qquad \text{where } C = [-1, 2]$$

2. $$f(x) = \begin{cases} 0 & \text{for } |x| \leq 1 \\ |x| - 1 & \text{for } 1 < |x| \leq 2 \\ +\infty & \text{for } |x| > 2 \end{cases}$$

3. $$f(x_1, x_2) = 2|x_1| + 3|x_2|$$

4. $$f(x) = \begin{cases} -x & \text{for } x < 0 \\ x^2 & \text{for } x \geq 0 \end{cases}$$

Convex conjugate

1. The aim of this exercise is to graphically construct the convex conjugate f^* of the following function:

$$f(x) = \begin{cases} \infty & \text{for } x \leq -2 \\ -x + 1 & \text{for } -2 < x \leq 0 \\ \frac{8}{27}x^2 + 1 & \text{for } 0 < x \leq \frac{3}{2} \\ x + \frac{1}{6} & \text{for } x > \frac{3}{2} \end{cases}$$

Check your result by constructing the bi-conjugate f^{**} which gives the original function again as $f(x)$ is a convex function.

In the top row of Figure 1 the function $f(x)$ is plotted on the left, on the right there is space for your solution. It suffices if you find the corresponding function values of $f^*(y)$ for the points marked in red in the original function. For easier construction the y-axis on the right is inverted. Below there is space to draw f^{**}.

Example to calculate y :

ad 1) $k = -\frac{5}{2} = -2.5 \Rightarrow y$

ad 2) $k = -\frac{1}{1} = -1 \Rightarrow y$

usw.

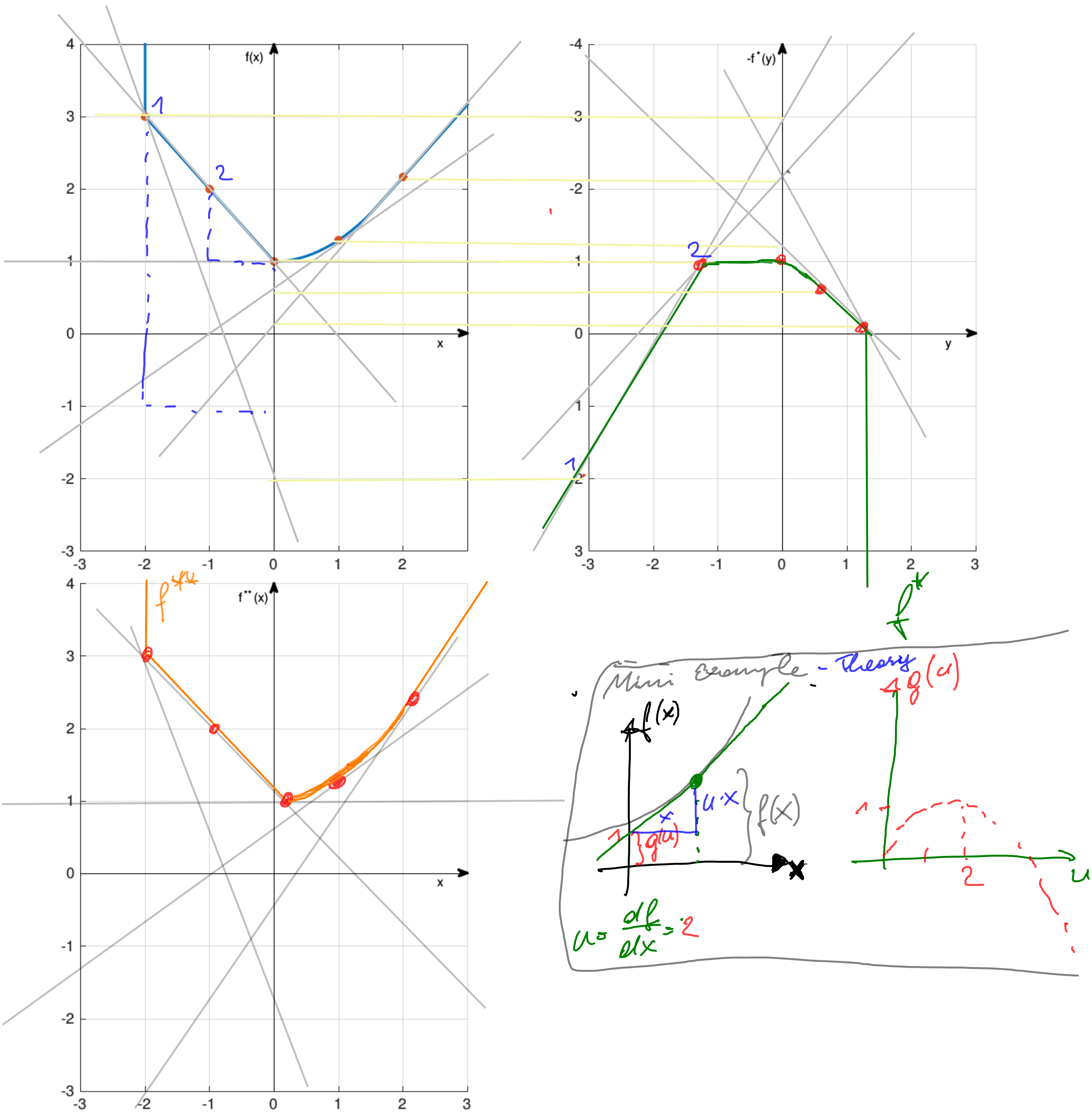


Figure 1: Graphical construction of $f^*(y)$ and $f^{**}(x)$.

$f = f^{**}$ ... convex $\square$

Hints for the reconstruction Draw a tangent at a point x_0 to $f(x)$. The tangent line is of the form $g(x) = kx + d$. The slope k of the tangent gives us the value for y in the new coorinate system, and $-d$, the negative intersection with the original y-axis, gives you the function value $f^*(y)$. This relationship can be derived if you calculate the conjugate of $g(x)$.

2. Which of these functions are convex? Calculate f^* and f^{**} analytically.

   a)
   $$f(x) = x\log(x) \quad \text{for } x > 0$$

   b)
   $$f(x) = \begin{cases} -\frac{5}{2} + 2|x| & \text{for } |x| > 1 \\ -|\frac{x}{2}| & \text{for } |x| \leq 1 \end{cases}$$

Subgradient method

The subgradient method is a method to find a minimizer of a convex, possibly non-smooth function $f(x)$. One iteration is given by

$$x^{k+1} = x^k - t^k g^k,$$

where t^k is a sequence of positive step sizes, $g^k \in \partial f(x^k)$ is any subgradient of f at x^k for $k \geq 0$. Consider the function

$$f(x) = \max_i \{a_i^T x + b\}$$

with

$$a_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad a_2 = \begin{pmatrix} -\frac{1}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} \end{pmatrix}, \quad a_3 = \begin{pmatrix} -\frac{1}{\sqrt{5}} \\ -\frac{2}{\sqrt{5}} \end{pmatrix}, \quad b = -3$$

and complete the following tasks

1. Compute the subdifferential $\partial f(x)$ analytically. Hint: Use the maximum of smooth convex functions.
2. Compute the optimal point x^* and $f^* = f(x^*)$ analytically.
3. Implement the subgradient method with Python and `numpy`.
   a) Execute your algorithm using a *constant step size*
   b) Execute your algorithm using the *diminishing step size*. Choose appropriate parameters β and γ.
   c) Execute your algorithm using the *dynamic step size*.
4. Plot the level sets of the function using `matplotlib`. You can use `plt.contour()` for this. Additionaly plot the iterates of the algorithm using the different step-sizes.
5. Plot the convergence of $f(x) - f^*$. Which step size selection works best?
6. Try to experiment with the dynamic step size when f^* is not known.
7. Experiment with different choices of the subgradient $g^k \in \partial f(x^k)$

---

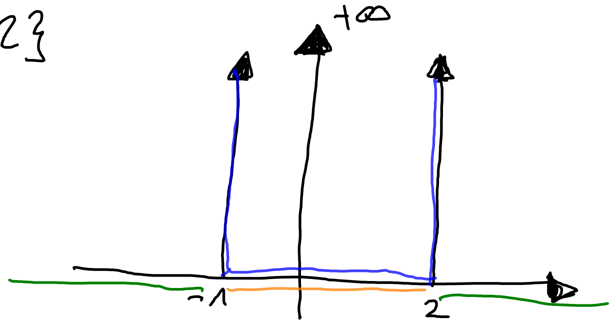
1) Subgradients

$$f(x) = \begin{cases} 0 & \text{for } x \in C \\ +\infty & \text{for } x \notin C \end{cases}, \text{ where } C = [-1, 2]$$

0 for $x \notin C$, $x \in \mathbb{R} \setminus \{-1, 2\}$

$[0, \infty]$ for $x = -1$

$[-\infty, 0]$ for $x = 2$



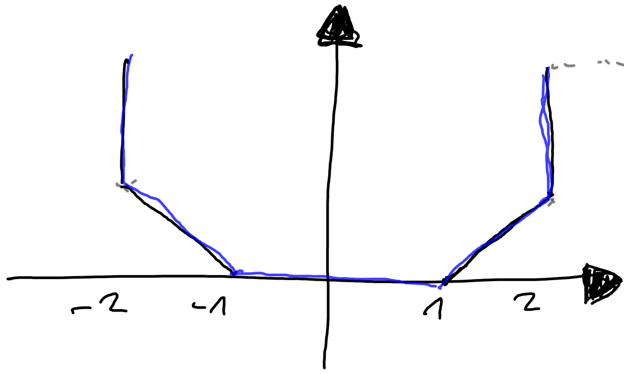
Auxilary Calculation: $\infty = C$

$\frac{\partial 0}{\partial x} = 0$ for $x \in \,]-1, 2[$

$$\frac{\partial c}{\partial x} = 0 \quad \text{for } x \in \mathbb{R} \setminus [-1,2]$$

$$\partial f(x) = \begin{cases} 0 & \text{for } x \in \mathbb{R} \setminus \{-1,2\} \\ [0,\infty] & \text{for } x = 2 \\ [-\infty,0] & \text{for } x = -1 \end{cases}$$

2) $$f(x) = \begin{cases} 0 & |x| \le 1 \\ |x|-1 & 1 < |x| \le 2 \\ +\infty & |x| > 2 \end{cases}$$

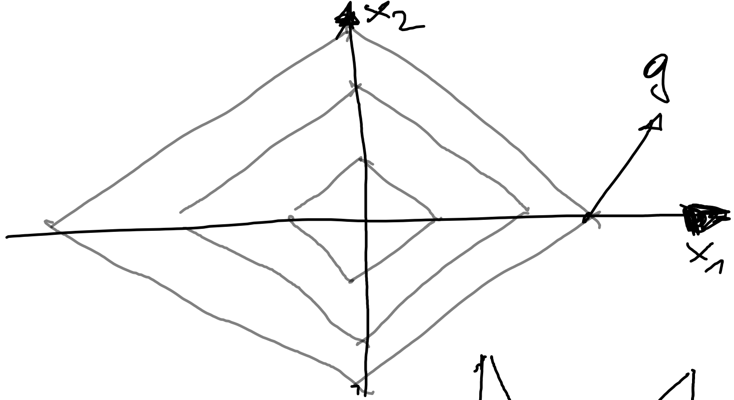


Rewrite $f(x)$ into cases:

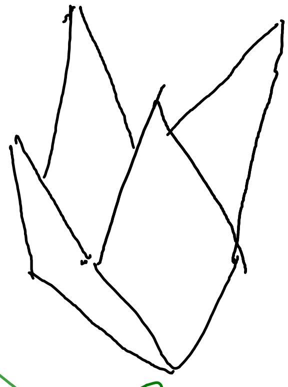
$$f(x) = \begin{cases} 0 & \text{for } x \ge 0,\ x \le 1 \text{ and } x \le 0,\ x \ge -1 \\ x-1 & \text{for } x > 1,\ x \le 1 \\ -x-1 & \text{for } x < -1,\ x \ge -2 \\ \infty & \text{for } x > 2 \text{ and } x < -2 \end{cases}$$

$$\partial f(x) = \begin{cases} [-1,0] & x = -1 \\ [0,1] & x = 1 \\ [-\infty,-1] & x = -2 \\ [1,\infty] & x = 2 \\ 0 & |x| < 1 \\ 0 & |x| < 2 \\ 1 & 1 \le x \le 2 \\ -1 & -2 \le x \le -1 \end{cases}$$

3) $f(x_1,x_2) = 2|x_1| + 3|x_2|$



$$|x| = \begin{cases} -1 & \text{if } x<0 \\ [-1,1] & \text{if } x=0 \\ 1 & \text{if } x>0 \end{cases}$$

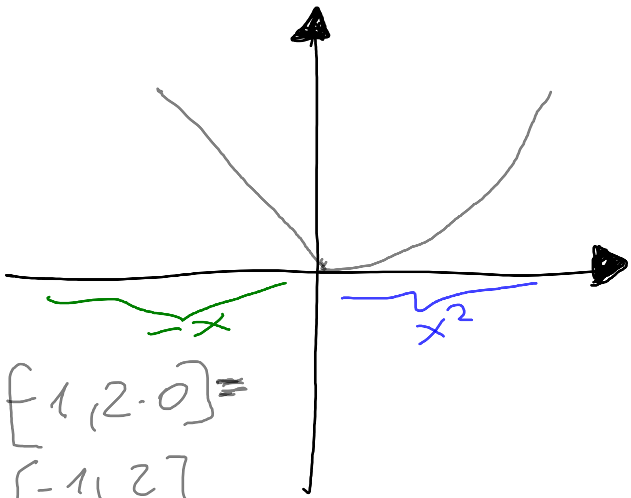


We investigate all kinked points:

$$\partial f \begin{cases} \left([-2,2]\times[-3,3]\right)^T & x_1 = x_2 = 0 \\ (2,3)^T & x_1, x_2 > 0 \\ (-2,3)^T & x_1<0, x_2>0 \\ (-2,-3)^T & x_1, x_2 < 0 \\ (2,-3)^T & x_1>0, x_2=0 \\ ([-2,2],3)^T & x_1=0, x_2>0 \\ (2,[-3,3])^T & x_1>0, x_2=0 \\ ([-2,2],-3)^T & x_1=0, x_2<0 \\ (-2,[-3,3])^T & x_1<0, x_2=0 \end{cases}$$

(column headings: ∂x_1 ∂x_2)

4) $f(x) = \begin{cases} -x & \text{for } x<0 \\ x^2 & \text{for } x \geq 0 \end{cases}$



$$\partial f \begin{cases} \cancel{x} < 0 & -1 \\ \cancel{x} > 0 & 2x \\ x = 0 & [-1, 2x] \Rightarrow [-1, 2\cdot 0] = \\ & = [-1, 2] \end{cases}$$

Convex Conjugate:

2.a) $f(x) = x \log x \quad \text{for } x > 0$

$$f^*(y) = \sup_{x \in \text{dom} f} (x^T y - x \log x)$$

$$= y - \log(x) - 1 = 0$$

$$= y - 1 = \log(x)$$

$$= \underline{\underline{x = e^{y-1}}}$$

$$f^*(y) = e^{y-1} y - e^{y-1} \log e^{y-1}$$

$$= e^{y-1} y - e^{y-1} (y-1)$$

$$= \cancel{e^{y-1} y} - \cancel{e^{y-1}} + e^{y-1} = e^{y-1}$$

$$f^{**}(x) = \sup_{y \in \text{dom} f^*} (x^T y - f^*(y))$$

$$= x^T - e^{y-1} = 0$$

$$x = e^{y-1}$$

$$\log(x) = y - 1$$

$$y = \log(x) + 1$$

$$f^{**}(x) = x^T y - e^{y-1}$$

$$= x^T(\log(x)+1) - e^{(\log(x)+1)-1}$$

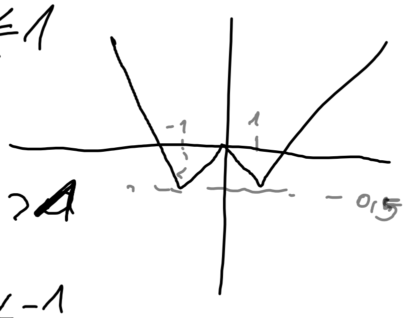
$$= x^T \log(x) + \cancel{x} - \cancel{x}$$

$$= x^T \log(x)$$

$f(x) = f^{**}(x) \quad \ldots \text{ convex}$

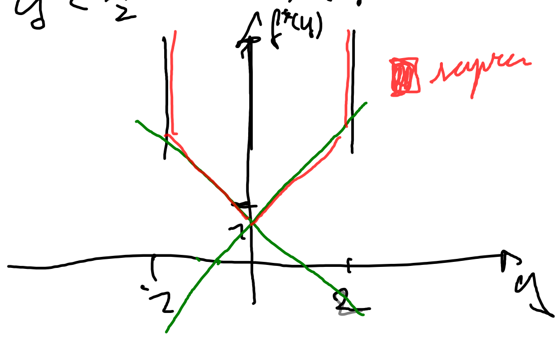
b)

$$f(x) = \begin{cases} -\frac{5}{2} + 2|x| & \text{for } |x| > 1 \\ -\left|\frac{x}{2}\right| & \text{for } |x| \le 1 \end{cases}$$



$$f^*(y) = \sup_{x \in \operatorname{dom} f} \quad \begin{matrix} xy - \left(-\frac{5}{2} + 2x\right) & x > 1 \\ xy - \left(-\frac{5}{2} - 2x\right) & x < -1 \\ xy - \left(-\frac{x}{2}\right) & x > 0, x \le 1 \\ xy - \left(+\frac{x}{2}\right) & x < 0, x > -1 \\ \cancel{xy - \left(-\frac{x}{2}\right)} = 0 & x = 0 \end{matrix}$$

$x > 1$	$x \to \infty$	$, y > 2$	$x(y-2) + \frac{5}{2}$
$x < -1$	$x \to -\infty$	$, y < -2$	$x(y+2) + \frac{5}{2}$
$x > 0, x \le 1$	$x \to 1$	$, y > \frac{1}{2}$	$x(y + \frac{1}{2})$
$x < 0, x > -1$	$x \to -1$	$, y < \frac{1}{2}$	$x(y - \frac{1}{2})$
0	$x = 0$		



$$f^*(y) = \begin{cases} \infty & |y| > 2 \\ |y| + \frac{1}{2} & |y| \le 2 \end{cases}$$

$$f^{**}(x) = \sup_{y \in \operatorname{dom} f^*} \{x^T y - f^*(y)\}$$

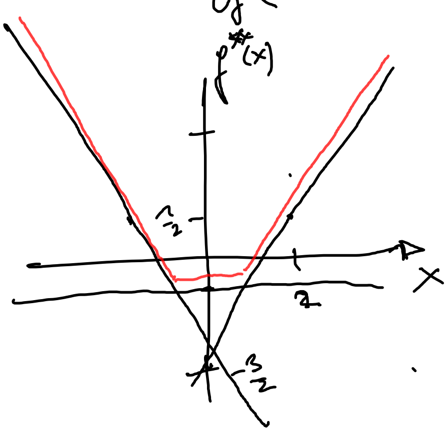
$$\begin{bmatrix} x^T y - (y + \frac{1}{2}) \\ x^T y - (-y + \frac{1}{2}) \\ -0{,}5 \end{bmatrix} \quad \begin{bmatrix} y > 0, y \le 2 \\ y < 0, y \ge -2 \\ y = 0 \end{bmatrix}$$

$y > 0, y \le 2$: $y \to 2$, $x = 2(x-1) - \frac{1}{2}$

$y < 0, y \ge -2$: $y \to -2$, $x = -2(x+1) - \frac{1}{2}$

$y(x-1) - \frac{1}{2}$

$y(x+1) - \frac{1}{2}$



$f^{**} \neq f$ ... nicht konvex!

Subgradient Method

1) $f(x) = \max\{(a_1^T x + b), \max\{a_2^T x + b, a_3 x + b\}\}$

$$\partial f(x) \begin{cases} a_1 & f = f_1 \\ a_2 & f = f_2 \\ a_3 & f = f_3 \\ \left(\left[-\frac{1}{\sqrt{5}}, 1\right], \left[-\frac{2}{\sqrt{5}}, 0\right]\right)^T & f = f_1 \ \&\& \ f = f_2 \\ \left(\left[-\frac{1}{\sqrt{5}}, 1\right], \left[-\frac{2}{\sqrt{5}}, 0\right]\right)^T & f = f_1 \ \&\& \ f = f_3 \\ \left(\left[-\frac{1}{\sqrt{5}}, \left(\frac{2}{\sqrt{5}}, \frac{2}{\sqrt{5}}\right)\right]\right)^T & f = f_2 \ \&\& \ f = f_3 \\ \left(\left[-\frac{1}{\sqrt{5}}, 1\right], \left[-\frac{2}{\sqrt{5}}, \frac{2}{\sqrt{5}}\right]\right)^T & f = f_1 \ \&\& \ f = f_2 \ \&\& \ f = f_3 \end{cases}$$

2) x^* and $f^* = f(x^*)$

The optimal point x^* of a subgradient $\partial f(x)$ is that case, which contains $(0,0)$.

$x^* = (0,0)^T$

$f^* = f(x^*) =$

$f^1 = f^2 = f^3$ $\qquad \underbrace{a_i^T \cdot \underbrace{x}_{0}}_{0} + b = b$

$$a_1^T x + b = a_2^T x + b = a_3^T x + b$$

$$(a_{11}\ a_{12}) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = (a_{21}\ a_{22}) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

$$a_{11} x_1 + a_{12} \cdot x_2 = a_{21} x_1 + a_{22} x_2$$

$$a_{11} x_1 - a_{21} x_1 = a_{12} x_2 + a_{22} x_2$$

$$x_1 (a_{11} - a_{21}) = x_2 (a_{12} + a_{22})$$

$$x_1 = \frac{x_2 (a_{12} + a_{22})}{a_{11} - a_{21}}$$

$\underline{a_2 x + b = a_3 x + b:}$

$$x_1^* = \frac{x_2 \left(-\frac{2}{\sqrt{5}} - \frac{2}{\sqrt{5}}\right)}{\underbrace{-\frac{1}{\sqrt{5}} + \frac{1}{\sqrt{5}}}_{0}} = \underline{\underline{0}}$$

$$\underline{\underline{x_2^* = 0}}$$

$\underline{a_1^T x + b = a_2^T x + b:}$ $x_1^* = \frac{x_2 \left(0 + \frac{2}{\sqrt{5}}\right)}{1 + \frac{1}{\sqrt{5}}} = \underline{\underline{\frac{2 x_2}{1 + \sqrt{5}}}}$

$$\underline{\underline{x_2^* = \frac{(1+\sqrt{5}) x_1}{2}}}$$

$\underline{a_1^T x + b = a_3^T x + b =}$ $x_1^* = \frac{x_2 \left(0 - \frac{2}{\sqrt{5}}\right)}{1 + \frac{1}{\sqrt{5}}} = \underline{\underline{-\frac{2 x_2}{1 + \sqrt{5}}}}$

$$\underline{\underline{x_2^* = \frac{(1+\sqrt{5}) x_1}{-2}}}$$

The optimal point $x^* = (0, 0)^T$.

$$f^* = \left((a_{11} \; a_{12}) \begin{pmatrix} 0 \\ 0 \end{pmatrix} - 3 \right) = \underline{\underline{-3}}$$

$\boxed{\text{→ Plots are in another pdf!}}$