Prof. Dr. Kathrin Flaßkamp
M.Sc. Matthias Hoffmann
Systems Modeling and Simulation
Saarland University

Optimization – Exercise 3 – WS 21/22

KKT conditions and Multiobjective Optimization

Exercise 3.1 – For Preparation: Constrained optimization problems and KKT conditions

Consider the optimization problem

$$\begin{aligned} \min \quad & \frac{1}{2}\left(x_1^2 + x_2^2\right) \\ \text{s.t.} \quad & x_1 + x_2 + 2 \leq 0. \end{aligned}$$

Solve the OP

a) graphically,

b) analytically using the KKT conditions,

c) numerically in MATLAB using `fmincon`.

Exercise 3.2 – For Preparation: Relation of scalarization methods

In the lecture we presented you the following four scalarization methods:

- weighted-sum method
- reference-point method
- ε-constraint method
- Pascoletti-Serafini scalarization

Some of these methods are related to one another. In particular, show that:

a) the ε-constraint method is a special case of Pascoletti-Serafini scalarization, if direction $\boldsymbol{d}$ and point $\boldsymbol{s}$ are chosen correctly.

b) the norm within the reference point method can be chosen such that the method corresponds to weighted-sum method.
Hint: Take a look at slide 14 of lecture 8.

You are free to come up with formal mathematical reasoning or intuitive graphical arguments.

Solution of 3.1

See live script.

Solution of 3.2 – a

The Pascoletti-Serafini scalarization is given by

$$\min_{x\in\mathbb{R}^n,\, l\in\mathbb{R}} \; -l \tag{1}$$

$$\text{s.t. } dl + s \geq F(x). \tag{2}$$

By setting d to $-e_k^\intercal$, with e_k the k-th unit vector, and $s = [\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_{k-1}, 0]^\intercal$, (2) can be given as

$$\begin{aligned} \text{s.t. } & f_i(x) \leq \varepsilon_i \quad \forall i \in [1, k-1] \\ & -l \geq f_k(x). \end{aligned} \tag{3}$$

Since $-l$ is minimized, $f_k(x)$ is its lower bound. Assuming equality, we can plug (3) into (1) and get the epsilon-constraint method

$$\begin{aligned} \min_{x\in\mathbb{R}^n} \; & f_k(x) \\ \text{s.t. } & f_i(x) \leq \varepsilon_i \quad \forall i \in [1, k-1]. \end{aligned}$$

Solution of 3.2 – b

Show, that the RPP

$$\min_{x\in\mathbb{R}^n}\left(\sum_{i=1}^{k}\left(T_i - f_i(x)\right)^p\right)^{\frac{1}{p}}$$

and WSP

$$\min_{x\in\mathbb{R}^n}\left(\sum_{i=1}^{k} w_i f_i(x)\right)$$

are related. For this, we choose a visual explanation starting from the implicit line equation and the level sets of p-norms. First, let's visualize how the level sets (points, where the equations have the same constant value) of general p-norms look like, they are shown in Figure 1. You can see, that these norms describe rhombuses ($p = 1$), circles ($p = 2$), and

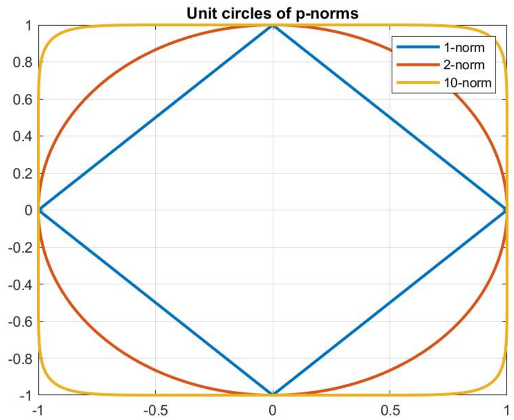

Figure 1: The unit circles (all points with a distance of 1) of p-norms.

squircles (else) around the origin (or around T in the RPP). With increasing constant values, the area enclosed by these shapes also increases.

Now, we also visualize the line equations $w_1x + w_2y = 0.5$ in Figure 2.

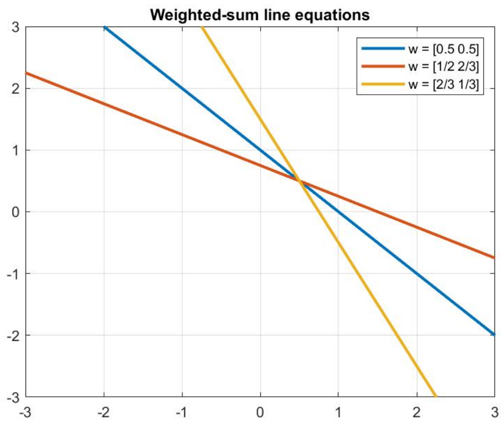

Figure 2: The lines resulting from interpreting the weighted sum as implicit line equations.

We see that in the 1st quadrant both the implicit line equation for $w = [0.5\ 0.5]$ and the 1-norm produce the same image.

Using the more general weighted p-norms

$$J = \left(\sum_{i=1}^{k} w_i |f_i|^p \right)^{1/p}$$

we can visualize the level sets of weighted 1-norms in Figure 3 and compare these to Fig. 2.

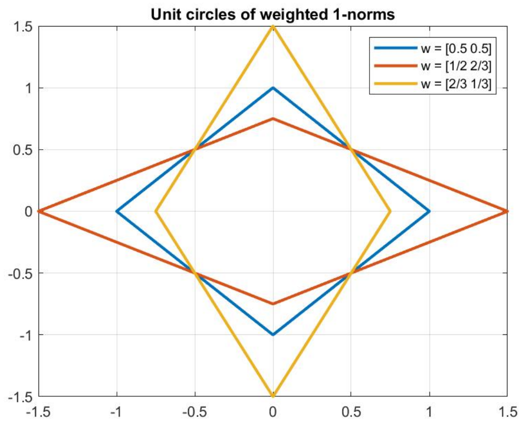

Figure 3: The unit circles of weighted 1-norms by different weightings.

In fact, if we restrict the equations to positive x- and y-values they are (obviously) the same. If we want to abide by this idea, we can assume the reference point to be equal to the utopian point, because then every difference $F(x) - T$, with $T_i = \min_x f_i(x)$ and x feasible, is greater than 0. Thus, with the Pareto front shifted such that the utopian point coincides with the origin, the whole Pareto front lies in the 1st quadrant and both results are equal.

To conclude: Under the assumption, that we use weighted 1-norms in the RPP and select the reference point to the utopian point, both methods will yield the same results.

For the interested: This also hints to other ideas like exponentiating the cost functions in the weighted sum method

$$J = \sum_{i=1}^{k} w_i f_i^p,$$

to get the so called *weighted exponential sum method.* This method is typically not working super well, as exponentiating typically increases the number of terms, thus making building gradients harder. The main difference between the methods from a) and the methods from b) is that while a) shoots rays, b) tries to make functions in implicit form tangent to the Pareto front. Ideally, the Inf-norm, also called max-norm, would be used, as the level sets have sharp edges that can find all Pareto optimal points in any concave region. As this is hard to compute lots of other functions exerting similar knees have been explored.