Convex Optimization and Modeling	Jun.-Prof. Matthias Hein
Solution of Exercise Sheet 4 - 05.05.2010	

Exercise 7 - Equivalent convex optimization problems

a. **(2 Points)** Exercise 4.5. Derive the equivalence of problem a) and b). Note, that the first problem is unconstrained but has a non-differentiable objective, whereas the second problem has a differentiable objective but has additional constraints. Thus both methods have to be solved by quite different optimization techniques. It can happen that one problem can be solved much more efficient than the other one. Recognizing equivalent problems is a key ability in convex optimization.

b. **(5 Points)** Exercise 4.11. Note again, that the minimization of the norm is an unconstrained but non-smooth problem, whereas the the equivalent optimization problems have a smooth objective but require additional constraints.

Solution:

a. We fix x and minimize with respect to w_k (note that we can minimize w_k componentwise),

$$\frac{\partial}{\partial w_k}\Big[\sum_{i=1}^{m}\Big(\frac{(\langle a_i, x\rangle - b_i)^2}{w_i+1} + M^2 w_i\Big)\Big] = -\frac{(\langle a_k, x\rangle - b_k)^2}{(w_k+1)^2} + M^2.$$

Solving for the extremal point yields

$$w_k = \frac{|\langle a_k, x\rangle - b_k|}{M} - 1.$$

Now, we have the constraint $w_k \geq 0$. We note that $w_k < 0$ if $|\langle a_k, x\rangle - b_k| < M$. However, the derivative is strictly positive if $|\langle a_k, x\rangle - b_k| < M$ and thus the minima is attained at $w_k = 0$ under the constraint $w_k \geq 0$. In total we have,

$$w_k = \begin{cases} \frac{|\langle a_k, x\rangle - b_k|}{M} - 1, & \text{if } |\langle a_k, x\rangle - b_k| \geq M, \\ 0, & \text{otherwise.} \end{cases}$$

The objective becomes

$$\frac{(\langle a_k, x\rangle - b_k)^2}{w_k+1} + M^2 w_k = \begin{cases} 2M|\langle a_k, x\rangle - b_k| - M^2, & \text{if } |\langle a_k, x\rangle - b_k| \geq M, \\ (\langle a_k, x\rangle - b_k)^2, & \text{if } |\langle a_k, x\rangle - b_k| < M. \end{cases}$$

and thus we get back the objective of the robust least squares problem Thus both problems are equivalent since the optimization over individual variables leads to equivalent problems.

b. In the following we always have $x \in \mathbb{R}^n$, $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$. Moreover, $\preceq, \succeq$ denotes componentwise inequalities,

- Minimizing $\|Ax - b\|_\infty$ is equivalent to the LP

$$\begin{aligned} &\min t \\ \text{subject to: } &Ax - b \preceq t\,\mathbf{1}, \\ &Ax - b \succeq -t\,\mathbf{1}, \end{aligned}$$

  where $t \in \mathbb{R}$ and $\mathbf{1}$ is a vector of ones.

- Minimizing $\|Ax-b\|_1$ is equivalent to the LP

$$\begin{aligned} \min \ & \sum_{i=1}^{m} t_i \\ \text{subject to: } & Ax-b \preceq t, \\ & Ax-b \succeq -t, \ , \end{aligned}$$

where now $t \in \mathbb{R}^m$,

- Minimizing $\|Ax-b\|_1$ subject to $\|x\|_\infty \leq 1$ is equivalent to the LP

$$\begin{aligned} \min \ & \sum_{i=1}^{m} t_i \\ \text{subject to: } & Ax-b \preceq t, \\ & Ax-b \succeq -t, \ , \\ & -\mathbf{1} \preceq x \preceq \mathbf{1}, \end{aligned}$$

where $t \in \mathbb{R}^m$.

- Minimizing $\|x\|_1$ subject to $\|Ax-b\|_\infty \leq 1$ is equivalent to the LP

$$\begin{aligned} \min \ & \sum_{i=1}^{n} t_i \\ \text{subject to: } & Ax-b \preceq \mathbf{1}, \\ & Ax-b \succeq -\mathbf{1}, \ , \\ & -t \preceq x \preceq t, \end{aligned}$$

where $t \in \mathbb{R}^n$.
Another equivalent formulation can be found by decomposing $x = x^+ - x^-$,

$$\begin{aligned} \min \ & \sum_{i=1}^{n} [x_i^+ + x_i^-] \\ \text{subject to: } & Ax^+ - Ax^- - b \preceq \mathbf{1}, \\ & Ax^+ - Ax^- - b \succeq -\mathbf{1}, \ , \\ & x^+ \succeq 0, \ x^- \succeq 0. \end{aligned}$$

- Minimizing $\|Ax-b\|_1 + \|x\|_\infty$ can be formulated as the LP

$$\begin{aligned} \min \ & \sum_{i=1}^{m} t_i + s \\ \text{subject to: } & Ax-b \preceq t, \\ & Ax-b \succeq -t, \ , \\ & -s\,\mathbf{1} \preceq x \preceq s\,\mathbf{1} \end{aligned}$$

where now $t \in \mathbb{R}^m$ and $s \in \mathbb{R}$.

Exercise 8 - Solving simple convex optimziation problems

a. **(5 Points)** Exercise 4.1. Use the optimality conditions given in the lecture and/or provide reasonable arguments e.g. using the optimality results for linear programming derived in the second lecture.

b. **(3 Points)** Exercise 4.8. a) and d). It often happens that one can rewrite an optimization problem and solve the minimization over some of the variables in closed form. Therefore it is important to know the solution of simple optimization problems.

Solution:

- We have the following optimization problem,

$$\begin{aligned} &\min\ f_0(x_1, x_2) \\ \text{subject to: } & 2x_1 + x_2 \geq 1, \\ & x_1 + 3x_2 \geq 1, \\ & x_1 \geq 0, \quad x_2 \geq 0. \end{aligned}$$

  The extremal points of the feasible set[1] are given as

$$u_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad u_2 = \begin{pmatrix} \frac{2}{5} \\ \frac{1}{5} \end{pmatrix}, \quad u_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

  The second one is the intersection of the two hyperplanes,

$$x_1 + 3x_2 - 1 = 2x_1 + x_2 - 1 \quad \Longrightarrow \quad 2x_1 + x_2 = 1 \quad \Longrightarrow \quad x_2 = \frac{1}{2}x_1.$$

  Plugging this condition again into the hyperplane, we attain the second extremal point.

  - $f_0(x_1, x_2) = x_1 + x_2$. This is a linear program and thus the optimal value (if it exists) is attained at one of the extremal points, which we can easily check

$$f_0(1, 0) = 1, \quad f_0\Big(\frac{2}{5}, \frac{1}{5}\Big) = \frac{3}{5}, \quad f_0(0, 1) = 1.$$

    Since obviously the minimum is attained, we have $x^* = \left(\frac{2}{5}, \frac{1}{5}\right)$ and $p^* = \frac{3}{5}$.
  - $f_0(x_1, x_2) = -(x_1 + x_2)$. The problem is unbounded from below, $p^* = -\infty$.
  - $f_0(x_1, x_2) = x_1$. With the constraint we have $f_0(x_1, x_2) \geq 0$ and the minimum 0 is attained at the set $\{(0, x_2) \,|\, x_2 \geq 1\}$.
  - $f_0(x_1, x_2) = \max\{x_1, x_2\}$. The level set $\{x \,|\, f_0(x) = c\}$ of the max-function is the set $\{(c, x_2) \,|\, 0 \leq x_2 \leq c\} \bigcup \{(x_1, c) \,|\, 0 \leq x_1 \leq c\}$. From the sketch of the feasible set it is obvious that thus the minimum is attained on the diagonal ($x_1 = x_2$). The "minimal" diagonal point contained in the set, lies on the first hyperplane $3x_1 = 1 \Rightarrow x^* = \left(\frac{1}{3}, \frac{1}{3}\right)$. This point is feasible since the second hyperplane has the value $\frac{4}{3} \geq 1$. One can check that the point on the diagonal on the second hyperplane is not feasible.
  - $f_0(x_1, x_2) = x_1^2 + 9x_2^2$. We know that x_0 is optimal given that,

$$\langle \nabla f, x - x_0 \rangle \geq 0, \quad \forall\, x \in X,$$

    where X is the feasible set. If the optimum lies in the interior of X the gradient has to vanish. We have for the gradient

$$\nabla f = (2x_1, 18x_2),$$

    We first note that the gradient vanishes only at $x = (0, 0)$ which is not feasible. Thus the optimum is attained at the boundary of X. Thus the optimum is either attained

[1] In the extended value formulation we could add $(\infty, 0)$ and $(0, \infty)$ as extremal points. Then the feasible set would be the convex hull of its extremal points.

at an extremal point or if the optimum lies at a point of the boundary which is not extremal, then the gradient ∇f has to be orthogonal to the boundary. First we check the function value of f_0 at the extremal points:

$$f_0(1,0) = 1, \quad f_0\Big(2/5, 1/5\Big) = 4/25 + 9/25 = 13/25, \quad f_0(0,1) = 9.$$

The feasible set is the intersection of four half-spaces which have (inward-pointing) normals,

$$n_1 = (0,1), \quad n_2 = (2,1), \quad n_3 = (1,3), \quad n_4 = (1,0).$$

We compute for each normal vector the point for which the gradient would be proportional to the normal vector,

$$\alpha \geq 0, \qquad u_1 = \Big(0, \frac{\alpha}{18}\Big), \quad u_2 = \Big(\alpha, \frac{\alpha}{18}\Big), \quad u_3 = \Big(\frac{\alpha}{2}, \frac{\alpha}{6}\Big), \quad u_4 = \Big(\frac{\alpha}{2}, 0\Big).$$

For u_1 and u_4 the resulting point $(0,0)$ which fulfills $x_2 = 0$ resp. $x_1 = 0$ is not feasible. For u_2 the point lies on the hyperplane $2x_1 + x_2 = 1$ given that $\alpha = \frac{18}{37}$. However, the resulting point $x_1 = \frac{18}{37}$ and $x_2 = \frac{1}{37}$ does not fulfill the second constraint $x_1 + 3x_2 \geq 1$ and thus is also not feasible. For u_3 the point lies on the hyperplane $x_1 + 3x_3 = 1$ for $\alpha = 1$ so that the candidate is $x^* = (\frac{1}{2}, \frac{1}{6})$. This point fulfills also all other constraints, in particular

$$2x_1^* + x_2^* = \frac{7}{6} \geq 1.$$

The function value is $f_0(x^*) = \frac{1}{4} + 9\frac{1}{36} = \frac{1}{2}$. Note, that this is also smaller than the function value at all extremal points (even though this is not necessary to check as the global optimum is for this problem unique since the objective is strictly convex).

- – We distinguish three cases. First, $Ax = b$ has no solution, that is b is not in the range of A. In this case the problem is infeasible, $p^* = \infty$. Second, the problem is feasible and c is orthogonal to the null space of A. Any feasible vector x can be written as, $x = x_0 + v$, where x_0 is a solution of $Ax = b$ and $v \in \ker(A)$. The objective function is then constant on the feasible set

  $$\langle c, x\rangle = \langle c, x_0\rangle + \langle c, v\rangle = \langle c, x_0\rangle\,.$$

  Thus the optimal value is $p^* = \langle c, x_0\rangle$. Third, the problem is feasible but c is not orthogonal to the null space of A. In this case it is obvious that $p^* = -\infty$ as $\langle c, v\rangle$ can be made arbitrarily small.

  – We have,

  $$\langle c, x\rangle \geq \min_i c_i \sum_{i=1}^n x_i.$$

  Moreover, the minimum is attained if we use

  $$x_i^* = \begin{cases} 1 & \text{if } i = \arg\min_j c_j, \\ 0 & \text{else.} \end{cases}$$

  Note, that the optimal solution is not unique if the minimum of c is not unique. The optimal value is $p^* = \min_i c_i$. If $\sum_i x_i \leq 1$ then the solution is $p^* = \min\{0, \min_i c_i\}$.