Norwegian University of Science and Technology
Department of Mathematical Sciences

TMA4180 Optimization Theory
Spring 2013

Exercise set 1

Tutorial: Thursday 24.01 16:15-17:00 in Kjl 4.

1 Consider the problem

$$\min_{(x,y)\in\Omega\subset\mathbb{R}^2}\{x^2+4y^2-2x-8y\}$$

where Ω is defined by the conditions (*constraints*)

$$x, y \geq 0,$$
$$5x+2y \leq 4.$$

(Here and later: $x \geq 0$ for a *vector* in $\mathbb{R}^n$ means $x_i \geq 0$ for $i = 1, 2, \ldots, n$. For matrices, $A \geq 0$ means *positive semidefinite*.)

Prove that Ω is convex and f is strictly convex. Do we have a solution and is it unique? Solve the problem in some way or another.

Solution: The constraints limit our domain Ω, to a triangle in the first quadrant with corners $(0,0)$, $(4/5,0)$, and $(0,2)$. The problem is immediately solvable by inspection and simple sketches. However, Ω is convex and f is strictly convex, since

$$f(x,y) = -\begin{bmatrix}2 & 8\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix} + \begin{bmatrix}x & y\end{bmatrix}\begin{bmatrix}1 & 0\\0 & 4\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix}.$$

If we find a (local) minimum, this will be the unique global minimum! The unconstrained minimum follows from

$$\frac{\partial f}{\partial x} = 2x-2 = 0, \qquad \frac{\partial f}{\partial y} = 8y-8 = 0,$$

that is, $\mathbf{x}_m = (1,1)$, but this is not in Ω. All minima have to be on the boundaries, and the line segment from $(0,2)$ to $(4/5,0)$ is closest to $\mathbf{x}_m$. We express the function in x only and find

$$f\left(x, \frac{4-5x}{2}\right) = x^2 + 4\left(\frac{4-5x}{2}\right)^2 - 2x - 8\left(\frac{4-5x}{2}\right) = 26x^2 - 22x,$$

with a minimum at $x_0 = 22/52 = 11/26$, and $y_0 = \frac{4-5x_0}{2} = 49/52$.

We expect that the negative gradient at (x_0, y_0) will point in the general direction of $\mathbf{x}_m$, and (x_0, y_0) will be the minimum if $\nabla f(x_0, y_0)$ is parallel to the normal of the line. This is easily seen to be true with

$$\nabla f(x_0,y_0) = \begin{bmatrix}2x_0-2\\8y_0-8\end{bmatrix} = -\frac{3}{13}\begin{bmatrix}5\\2\end{bmatrix}.$$

Thus, (x_0, y_0) is our unique global minimum.

2 **a)** Let $f\colon \mathbb{R}^n \to \mathbb{R}$ be a differentiable function where $\nabla^2 f(x)$ exists and is positive semi-definite, i.e. $\nabla^2 f(x) \geq 0$ at all points. Show that f is convex (*Hint*: Apply the theorem in the Basic Mathematical Tools note about the graph of functions being above their tangent planes).

What can you say if $\nabla^2 f(x)$ is positive definite, that is, $\nabla^2 f(x) > 0$?

Solution: We use Taylor's theorem:

$$f(x) = f(x_0) + \nabla f(x_0)^{\mathrm{T}}(x - x_0) + \tfrac{1}{2}(x - x_0)^{\mathrm{T}}\nabla^2 f(x_\theta)(x - x_0).$$

But $T(x) = f(x_0) + \nabla f(x_0)^{\mathrm{T}}(x - x_0)$ *is the tangent plane in* x_0*, and the last term* $\frac{1}{2}(x - x_0)^{\mathrm{T}}\nabla^2 f(x_\theta)(x - x_0) \geq 0$ *by assumption. Hence* f *is above all tangent planes. If* $\nabla^2 f(x_\theta) > 0$*, the function will be* strictly *convex.*

b) If $0 \neq a \in \mathbb{R}^n$ and $\beta \in \mathbb{R}$, the set S defined by

$$S = \{x \mid a^{\mathrm{T}}x \geq \beta\}$$

defines what is called a half-space. Show that it is convex by applying the definition for a convex set. Also show that the set

$$\{x \mid Ax \geq b, x \in \mathbb{R}^n, b \in \mathbb{R}^m, A \in \mathbb{R}^{m\times n}\}$$

is convex. (*Hint*: Apply the result about the intersection of convex sets.)

Solution: The equation for a plane in $\mathbb{R}^n$ *through* x_0*, having a normal vector* a*, is exactly similar to the equation we know from* $\mathbb{R}^3$*:*

$$P = \{x \mid a^{\mathrm{T}}(x - x_0) = 0\} = \{x \mid a^{\mathrm{T}}x = b\},$$

where $b = a^{\mathrm{T}}x_0$ *(this is called a* hyperplane *when* $n > 3$*). A hyperplane splits the space in two half-spaces*

$$\begin{aligned} P^+ &= \{x \mid a^{\mathrm{T}}x \geq b\}, \\ P^- &= \{x \mid a^{\mathrm{T}}x \leq b\}, \end{aligned}$$

having the plane defined by $a^{\mathrm{T}}x = b$ *in common. Both are easily seen to be convex: For* P^-*, with* $x, y \in P^-$ *and* $x_\theta = \theta x + (1-\theta)y$*, we have*

$$a^{\mathrm{T}}x_\theta = a^{\mathrm{T}}(\theta x + (1-\theta)y) = \theta a^{\mathrm{T}}x + (1-\theta)a^{\mathrm{T}}y \leq \theta b + (1-\theta)b = b.$$

We now observe that

$$\begin{aligned} \{x \mid Ax \geq b, x \in \mathbb{R}^n, b \in \mathbb{R}^m\} &= \{x \mid a_j^{\mathrm{T}}x \geq b_j, x \in \mathbb{R}^n, j = 1, \ldots, m\} \\ &= \bigcap_{j=1}^{m}\{x \mid a_j^{\mathrm{T}}x \geq b_j, x \in \mathbb{R}^n\}, \end{aligned}$$

and intersections of convex sets are convex.

c) Determine all minima of the function

$$f(x, y) = x^2 + y^2 - 2yx - 2y + 2x + 5$$

when (x, y) are free to vary over $\mathbb{R}^2$.

Solution: This is an unconstrained problems where

$$\nabla f(\mathbf{x}) = \begin{bmatrix} 2x - 2y + 2 \\ 2y - 2x - 2 \end{bmatrix},$$

and

$$\nabla^2 f(\mathbf{x}) = \begin{bmatrix} 2 & -2 \\ -2 & 2 \end{bmatrix} \geq 0.$$

The minima are along the line $y = x + 1$ since the function is convex and the points on the line solve $\nabla f(\mathbf{x}) = 0$.

d) Sketch how the result of **b)** and **c)** could be used to find all solutions of

$$\min f(x, y),$$

when (x, y) are restricted to satisfy

$$A \begin{bmatrix} x \\ y \end{bmatrix} \geq b, \tag{1}$$

where $f(x, y)$ is the function in **c)**, $A \in \mathbb{R}^{m\times 2}$, $b \in \mathbb{R}^m$, and we know that one of the solutions in **c)** satisfies the inequality (1).
(Show that if we apply what we know from **c)**, all 2-dimensional inequalities in (1) reduce to simple one-dimensional inequalities).

Solution: The constraints are all of the form

$$a_{j1}x + a_{j2}y \geq b_j, \quad j = 1, \ldots, m, \tag{2}$$

and defines a convex set in the plane. The solution set is the intersection of this set and the line $y = x + 1$. If the equation for the line is introduced into (**??**), *we obtain a set of simple one-dimensional inequalities*

$$(a_{j1} + a_{j2})x \geq b_j - a_{j2}, \quad j = 1, \ldots, m,$$

which may be inspected in turn.

3 A very common optimization problem is the following:

$$\min_{x\in\mathbb{R}^n_+} f(x),$$
$$\mathbb{R}^n_+ = \{x \mid x_i \geq 0, i = 1, \ldots, n\}.$$

Here, $\mathbb{R}^n_+$ is the *non-negative cone* in $\mathbb{R}^n$, and the problem is therefore to find the minimum for x-es where all components are non-negative. We assume that ∇f exists in all points.

Consider the following line-search algorithm: If we are at $x = (x_1, \ldots, x_n)^{\mathrm{T}}$, the search direction, $p = (p_1, \ldots, p_n)^{\mathrm{T}}$, is selected as follows:

$$p_i = \begin{cases} -\frac{\partial f}{\partial x_i}, & \text{if } x_i > 0 \text{ or } \frac{\partial f}{\partial x_i} < 0, \\ 0, & \text{if } x_i = 0 \text{ and } \frac{\partial f}{\partial x_i} \geq 0. \end{cases}$$

a) What are the first-order conditions for a minimum point for this problem? (*Hint*: First-order conditions for a minimum are most easily explained in term of the directional derivatives, which have to be non-negative for all feasible directions d, $\delta f(x,d) = \nabla f(x)^{\mathrm{T}} d \geq 0$. Use this to find the conditions on $\frac{\partial f}{\partial x_i}$, both for interior points and the boundary points.)
Solution: Necessary first-order conditions defined in terms of the directional derivative, which has to be non-negative in all feasible directions d; in this case,

$$\delta f(x,d) = \nabla f(x)^{\mathrm{T}} d \geq 0.$$

Thus, at an interior point *where d is free to vary,*

$$\nabla f(x) = 0.$$

For a boundary point $x = \{x_i\}_{i=1}^n$, *the feasible directions must point* into Ω, *and hence the components d_i of d, must be non-negative whenever $x_i = 0$. This implies that*

$$\frac{\partial f}{\partial x_i}(x) = \begin{cases} 0, & x_i > 0, \\ \textit{non-negative}, & x_i = 0. \end{cases}$$

b) Show that $p = 0$ at a point satisfying the first-order conditions. (*Hint*: Apply the result from **a)**.)
*Solution: This follows from **a)** and the definition of p.*

c) Show that any $p \neq 0$ is a *descent direction.* (The direction d is a descent direction if $\delta f(x,d) < 0$.)
Solution: The definition of p gives immediately that

$$\nabla f(x)^{\mathrm{T}} p = \sum_{\substack{1 \leq i \leq n \\ p_i \neq 0}} \frac{\partial f}{\partial x_i} \cdot \left(- \frac{\partial f}{\partial x_i} \right) = - \sum_{i=1}^{n} |p_i|^2 = -\|p\|^2 < 0,$$

and p is really a descent direction.