TMA4180
Optimisation I
Spring 2018

Norwegian University of Science and Technology
Department of Mathematical Sciences

Solutions to exercise set 9

$\boxed{1}$ Consider a constrained minimization problem with continuously differentiable objective function and constraints. Which of the following statements are true?

a) x^* is a (point of) global minimum $\implies$ x^* is a KKT point.

b) x^* is a local minimum and CQ holds $\implies$ x^* is a KKT point.

c) x^* is a KKT point and CQ holds $\implies$ x^* is a local minimum.

d) x^* is a global minimum and the problem is convex $\implies$ x^* is a KKT point.

e) x^* is a KKT point and the problem is convex $\implies$ x^* is a global minimum.

Solution: a), d): no, one needs CQ. See for example the convex problem $x \to \min!$, subject to $-x^2 - y^2 \geq 0$.

b), e): yes; KKT is necessary assuming CQ, and is sufficient assuming convexity

c): no; see e.g. $-x^2 \to \min!$ subject to $-1 \leq x \leq 1$. Then $x = 0$ is a KKT point and is not a local minimum (in fact a global maximum).

$\boxed{2}$ Consider the constrained optimization problem

$$x^2 + y^2 \to \min \qquad \text{such that} \qquad \begin{cases} x + y \geq 1, \\ y \leq 2, \\ y^2 \geq x. \end{cases}$$

a) Formulate the KKT-conditions for this optimization problem.

Solution: We begin by stating the problem in standard form, writing $\mathbf{x} = [x, y]^T$:

$$\min_{\mathbf{x}\in\mathbb{R}^2} f(\mathbf{x}) \quad \text{s.t.} \quad c_i(\mathbf{x}) \geq 0, \quad i = 1, 2, 3,$$

where

$$\begin{aligned} f(\mathbf{x}) &= x^2 + y^2, \\ c_1(\mathbf{x}) &= x + y - 1, \\ c_2(\mathbf{x}) &= 2 - y, \\ c_3(\mathbf{x}) &= y^2 - x. \end{aligned}$$

The KKT conditions can now be stated as follows:

$$2x^* - \lambda_1^* + \lambda_3^* = 0 \tag{1a}$$
$$2y^* - \lambda_1^* + \lambda_2^* - 2y^*\lambda_3^* = 0 \tag{1b}$$
$$x^* + y^* - 1 \geq 0 \tag{1c}$$
$$2 - y^* \geq 0 \tag{1d}$$
$$y^{*2} - x^* \geq 0 \tag{1e}$$
$$\lambda_i^* \geq 0, \quad i = 1, 2, 3 \tag{1f}$$
$$\lambda_1^*(x^* + y^* - 1) = 0 \tag{1g}$$
$$\lambda_2^*(2 - y^*) = 0 \tag{1h}$$
$$\lambda_3^*(y^{*2} - x^*) = 0. \tag{1i}$$

b) Find all KKT points for this optimization problem.

Solution: The feasible set is sketched in Figure 1.

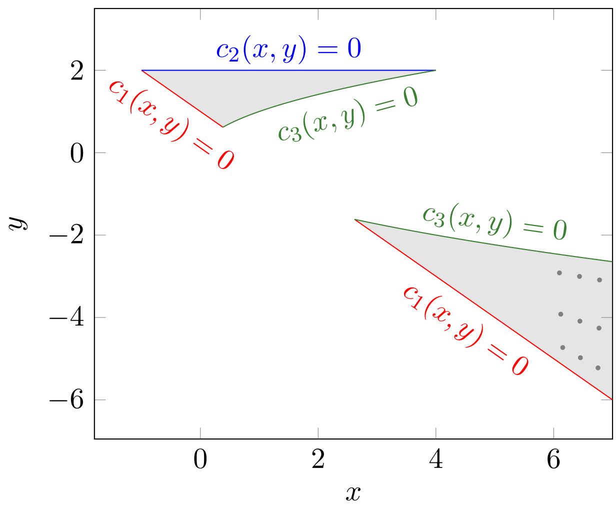

Figure 1: Feasible set. Note: The lower "triangle" extends further toward infinity.

We will find all KKT points by systematically considering all possible active sets of constraints. Remember that a constraint c_i is active at a point $\mathbf{x}$ if $c_i(\mathbf{x}) = 0$. Also, the LICQ conditions are satisfied at every point we consider here; with one active constraint, the LICQ conditions hold trivially, and in the cases with two constraints it is not hard to check that the LICQ conditions do hold. Observe that if $\mathbf{x}^* = [x^*, y^*]^T$ is a KKT point, then from (1a) and (1b) we have:

$$x^* = \frac{\lambda_1^* - \lambda_3^*}{2}, \quad y^* = \frac{\lambda_1^* - \lambda_2^*}{2(1 - \lambda_3^*)}.$$

From here on, we will drop the asterisk in the notation and write x for x^*, etc.

First, suppose that the active set is empty, i.e. neither of (1c)-(1e) are equalities. This corresponds to the interior of the domain. Then, by (1g)-(1i), we have $\lambda_1 = \lambda_2 = \lambda_3 = 0$, and so $x = y = 0$. But this point is not feasible, since it violates condition (1c). Thus, with the active set empty, there are no KKT points.

Next, we consider the case when the active set contains one index, i.e. exactly one of (1c)-(1e) is an equality. This corresponds to the boundaries of the domain, excepting the corner points. If (1c) is active, then $\lambda_2 = \lambda_3 = 0$ while $\lambda_1 \geq 0$. We get

$$x = \frac{\lambda_1}{2}, \quad y = \frac{\lambda_1}{2},$$

and inserting this into (1c) (which is now an equality), we get the condition

$$\frac{\lambda_1}{2} + \frac{\lambda_1}{2} - 1 = 0 \Rightarrow \lambda_1 = 1,$$

giving us the point $(x, y) = (\frac{1}{2}, \frac{1}{2})$. But this point violates condition (1e), so $(\frac{1}{2}, \frac{1}{2})$ is not a KKT point.

If (1d) is active, then $\lambda_1 = \lambda_3 = 0$ while $\lambda_2 \geq 0$, so

$$x = 0, \quad y = -\frac{\lambda_2}{2}.$$

Inserting this into the equality (1d), we get

$$2 + \frac{\lambda_2}{2} = 0 \Rightarrow \lambda_2 = -4.$$

Since the Lagrange multiplier is negative, KKT conditions are not satisfied at this point.

If (1e) is active, then $\lambda_1 = \lambda_2 = 0$ while $\lambda_3 \geq 0$, so

$$x = -\frac{\lambda_3}{2}, \quad y = 0.$$

Inserting this into the equality (1e), we get

$$\frac{\lambda_3}{2} = 0 \Rightarrow \lambda_3 = 0.$$

This gives the candidate point $(0, 0)$, which is not feasible since it violates (1c), and thereby is not a KKT point.

Having considered all possible active sets of one index, we now turn to the cases with two indices, i.e. exactly two of (1c)-(1e) are equalities. This corresponds to the corner points of the domain. First, if (1c) and (1d) are both active, then $\lambda_3 = 0$ while $\lambda_1, \lambda_2 \geq 0$. This gives us

$$x = \frac{\lambda_1}{2}, \quad y = \frac{\lambda_1 - \lambda_2}{2}.$$

Plugging this into equalities (1c) and (1d) yields:

$$\begin{aligned} \frac{\lambda_1}{2} + \frac{\lambda_1 - \lambda_2}{2} - 1 &= 0 \\ 2 - \frac{\lambda_1 - \lambda_2}{2} &= 0, \end{aligned}$$

with solutions $\lambda_1 = -2$ and $\lambda_2 = -6$. Since the multipliers are negative, this is not a KKT point.

Next, if (1c) and (1e) are both active, then $\lambda_2 = 0$ while $\lambda_1, \lambda_3 \geq 0$, which means

$$x = \frac{\lambda_1 - \lambda_3}{2}, \quad y = \frac{\lambda_1}{2(1-\lambda_3)}.$$

Plugging this into equalities (1c) and (1e) yields:

$$\frac{\lambda_1 - \lambda_3}{2} + \frac{\lambda_1}{2(1-\lambda_3)} - 1 = 0$$
$$\frac{\lambda_1^2}{4(1-\lambda_3)^2} - \frac{\lambda_1 - \lambda_3}{2} = 0.$$

Solving this set of equations yields $\lambda_1 = 5 \pm \frac{9}{\sqrt{5}}$ and $\lambda_3 = 2 \pm \frac{4}{\sqrt{5}}$, thereby giving the candidate points $(x, y) = (\frac{1}{2}(3 \pm \sqrt{5}), \frac{1}{2}(-1 \mp \sqrt{5}))$ which both satisfy the KKT conditions. Since $\lambda_1, \lambda_3 \geq 0$, these points are minimizer candidates. Note: This result can be arrived upon by the easier approach of first finding the points (x, y) where c_1 and c_3 are both active, then working out what λ_1 and λ_3 are.

Finally, we check the case where (1d) and (1e) are both active, i.e. $\lambda_1 = 0$ while $\lambda_2, \lambda_3 \geq 0$. This gives us

$$x = -\frac{\lambda_3}{2}, \quad y = -\frac{\lambda_2}{2(1-\lambda_3)}.$$

Plugging this into equalities (1d) and (1e) yields:

$$2 + \frac{\lambda_2}{2(1-\lambda_3)} = 0$$
$$\frac{\lambda_2^2}{4(1-\lambda_3)^2} + \frac{\lambda_3}{2} = 0,$$

which can be solved to find $\lambda_2 = -28$ and $\lambda_3 = -8$. Since the multipliers are negative, this is not a KKT point.

Concerning the case with all constraints active, we may conclude that no KKT point exists; all three constraint functions cannot be active at the same point. The investigation is summarized in the table below.

Point	λ_1	λ_2	λ_3	KKT?
(0,2)	0	-4	0	No
$(\frac{1}{2}(3+\sqrt{5}), \frac{1}{2}(-1-\sqrt{5}))$	$5+\frac{9}{\sqrt{5}}$	0	$2+\frac{4}{\sqrt{5}}$	Yes
$(\frac{1}{2}(3-\sqrt{5}), \frac{1}{2}(-1+\sqrt{5}))$	$5-\frac{9}{\sqrt{5}}$	0	$2-\frac{4}{\sqrt{5}}$	Yes
(-1,2)	-2	-6	0	No
(4,2)	0	-28	-8	No

c) Find all local and global minima for this optimization problem.

Solution: To determine whether the KKT points that are minimizer candidates are in fact local minimizers, we check the second order sufficient conditions from

Theorem 12.6 in N&W, i.e. whether

$$w^T \nabla^2_{xx} \mathcal{L}(x, \lambda) w > 0 \ \forall w \in \mathcal{C}(x, \lambda), w \neq 0, \tag{2}$$

where, $\mathcal{C}(x, \lambda)$ is the critical cone at x, given by (12.53) in N&W.

For both candidates, i.e. $(\frac{1}{2}(3 \pm \sqrt{5}), \frac{1}{2}(-1 \mp \sqrt{5}))$, we have that the critical cone is simply given as $\mathcal{C}(x, \lambda) = \{0\}$. This is because any $w \in \mathcal{C}(x, \lambda)$ must be orthogonal to the $\nabla c_i(x)$ for which $\lambda_i > 0$, of which there are two for each point. Since the LICQ conditions hold at both points, these two vectors are linearly independent and thus span $\mathbb{R}^2$. The only vector orthogonal to $\mathbb{R}^2$ is the zero vector. Thereby, the only vector in $\mathcal{C}(x, \lambda)$ is the zero vector for these points, and thus condition (2) holds. We can conclude that $(\frac{1}{2}(3 \pm \sqrt{5}), \frac{1}{2}(-1 \mp \sqrt{5}))$ are strict local minimizers.

We note that $f(\frac{1}{2}(3 - \sqrt{5}), \frac{1}{2}(-1 + \sqrt{5})) < f(\frac{1}{2}(3 + \sqrt{5}), \frac{1}{2}(-1 - \sqrt{5}))$ and $f(\mathbf{x}) \to \infty$ in the unbounded region of the feasible domain. This means that $(\frac{1}{2}(3 - \sqrt{5}), \frac{1}{2}(-1 + \sqrt{5}))$ is a global minimizer and $(\frac{1}{2}(3 + \sqrt{5}), \frac{1}{2}(-1 - \sqrt{5}))$ is a local minimizer.

3 Consider a *linear programming problem*

$$c^\top x \to \min \qquad \text{subject to} \qquad Ax \geq b,$$

where $c, x \in \mathbb{R}^n$, $b \in \mathbb{R}^m$, and $A \in \mathbb{R}^{m \times n}$.

a) Show that for each feasible point x the radial cone $R_\Omega(x)$ is the same as the cone of linearized feasible directions $F_\Omega(x)$. Conclude that $T_\Omega(x) = F_\Omega(x)$, i.e., CQ holds.

Solution: We will denote the rows of the matrix A by a_i. $d \in R_\Omega(x) \iff x + \epsilon d \in \Omega, \forall \text{small } \epsilon > 0 \iff a_i(x + \epsilon d) - b_i = a_i x - b_i + \epsilon a_i d \geq 0, \forall \text{small } \epsilon > 0, \forall i = 1, \ldots, m \iff a_i d \geq 0, \forall i \in \mathcal{A}(x) \iff d \in F_\Omega(x)$, because a_i is the (transposed) gradient of the ith linear constraint $a_i x - b_i \geq 0$.
Thus $R_\Omega(x) = F_\Omega(x)$. Owing to the inclusion $R_\Omega(x) \subseteq T_\Omega(x) \subseteq F_\Omega(x)$, which is always valid, we must have $T_\Omega(x) = F_\Omega(x)$.

b) State the KKT optimality conditions for this problem. Show that at every KKT point x we have the equality $c^\top x = b^\top \lambda$, where $\lambda \in \mathbb{R}^m_+$ is a vector of Lagrange multipliers.

Solution: Since CQ hold, for a feasible point x to be a local minimum the following conditions must hold:

$$\begin{aligned} Ax &\geq b, \\ c &= \sum_{i=1}^m \lambda_i a_i^\top = A^\top \lambda, \\ \lambda &\geq 0, \\ \lambda_i(a_i x - b_i) &= 0, \quad i = 1, \ldots, m. \end{aligned}$$

Note that the last condition holds iff $\lambda^\top(Ax-b)=0$, because $\lambda \geq 0$ and $Ax-b\geq 0$, owing to the feasibility of x. As a result, we have

$$b^\top \lambda = x^\top A^\top \lambda = x^\top c,$$

where we used the fact that $c = A^\top \lambda$.

$\boxed{4}$ Consider the constrained optimization problem

$$x \to \min \qquad \text{such that} \qquad \begin{cases} y \geq x^4, \\ y \leq x^3. \end{cases}$$

Find all KKT points and local minima for this optimization problem.

Solution: We begin by stating the problem in standard form, writing $\mathbf{x} = [x,y]^T$:

$$\min_{\mathbf{x}\in\mathbb{R}^2} f(\mathbf{x}) \quad \text{s.t.} \quad c_i(\mathbf{x}) \geq 0, \quad i=1,2$$

where

$$\begin{aligned} f(\mathbf{x}) &= x, \\ c_1(\mathbf{x}) &= y - x^4 \\ c_2(\mathbf{x}) &= x^3 - y \end{aligned}$$

The KKT conditions for this problem can be stated as follows:

$$\begin{aligned} 1 + 4x^3\lambda_1 - 3x^2\lambda_2 &= 0 && \text{(3a)} \\ -\lambda_1 + \lambda_2 &= 0 && \text{(3b)} \\ y - x^4 &\geq 0 && \text{(3c)} \\ x^3 - y &\geq 0 && \text{(3d)} \\ \lambda_i \geq 0,& \quad i=1,2 && \text{(3e)} \\ \lambda_1(y-x^4) &= 0 && \text{(3f)} \\ \lambda_2(x^3-y) &= 0. && \text{(3g)} \end{aligned}$$

Now, we can take a shortcut; from (3b), we see that $\lambda_1 = \lambda_2$, and from (3a) we see that there cannot exist any KKT point for which $\lambda_1 = \lambda_2 = 0$. Therefore, the cases with no active constraints ($\lambda_1 = \lambda_2 = 0$) and one active constraint ($\lambda_1 = 0$ or $\lambda_2 = 0$) cannot produce KKT points. We are left with considering the case where both constraints are active, i.e. the corner points (0,0) and (1,1).

In the point (1,1), we find (by (3a) and (3b)) that $\lambda_1 = \lambda_2 = -1$, and therefore this is not a KKT point.

The last point is (0,0), for which we cannot write the gradient of f at (0,0) (which is $[1,0]^T$) as a non-negative linear combination of the gradients of the constraints, and which therefore is not a KKT point. This does not, however, mean that it is not a minimizer. Applying common sense, it is clearly a local minimum, as no other points with $x=0$ are feasible, and $x=0$ is the lowest possible value of the objective function.

$\boxed{5}$ Consider the constrained optimization problem

$$xy \to \min \qquad \text{such that} \qquad \begin{cases} y \geq x, \\ y^4 \leq x^3. \end{cases}$$

a) Find all KKT points and local minima for this optimization problem.

Solution: We begin, as usual, by stating the problem in standard form, writing $\mathbf{x} = [x, y]^T$:

$$\min_{\mathbf{x} \in \mathbb{R}^2} f(\mathbf{x}) \quad \text{s.t.} \quad c_i(\mathbf{x}) \geq 0, \quad i = 1, 2$$

where

$$\begin{aligned} f(\mathbf{x}) &= xy, \\ c_1(\mathbf{x}) &= y - x \\ c_2(\mathbf{x}) &= x^3 - y^4 \end{aligned}$$

The KKT conditions are:

$$y + \lambda_1 - 3x^2\lambda_2 = 0 \tag{4a}$$
$$x - \lambda_1 + 4y^3\lambda_2 = 0 \tag{4b}$$
$$y - x \geq 0 \tag{4c}$$
$$x^3 - y^4 \geq 0 \tag{4d}$$
$$\lambda_i \geq 0, \quad i = 1, 2 \tag{4e}$$
$$\lambda_1(y - x) = 0 \tag{4f}$$
$$\lambda_2(x^3 - y^4) = 0. \tag{4g}$$

Now, we can check the different cases of active constraints to find KKT points. First, with no active constraints, i.e. $\lambda_1 = \lambda_2 = 0$, we get the point (0,0). In fact, this is a point with both constraints active; it just so happens that $\lambda_1 = \lambda_2 = 0$ here. One can check that the LICQ does not hold here, but it is still a KKT point because the gradient of f at (0,0) (which is 0) can be written as a non-negative linear combination of the gradients of the constraints. Thereby, we conclude that there exist no KKT points with no active constraints, but that (0,0) is a KKT point. Note that without CQ KKT conditions no longer *have to* hold at points of local minima, but they *may* still hold!

Next, we check with one active constraint. First, with $\lambda_1 \geq 0, \lambda_2 = 0$, we have from (4a) and (4b) that $x = \lambda_1$ and $y = -\lambda_1$. Inserting into the equality (4c) yields $\lambda_1 = 0$, and therefore $(x, y) = (0, 0)$ again, which has been discussed already.

With $\lambda_2 \geq 0, \lambda_1 = 0$, equations (4a), (4b) and (4d) become

$$\begin{aligned} y - 3x^2\lambda_2 &= 0, \\ x + 4y^3\lambda_2 &= 0, \\ x^3 &= y^4. \end{aligned}$$

Multiplying the first of these by x, the second by y, applying the third and adding the two first gives

$$y^4\lambda_2 = 0.$$

Any solution of this leads to the point (0,0), which we have already found to be a KKT point.

Finally, we check the case with two active constraints, for which there are two points; (0,0), which is already considered, and (1,1). In the point (1,1), we find (by (4a) and (4b)) that $\lambda_1 = -7$ and $\lambda_2 = -2$ and is therefore this is not a KKT point.
Thus the only KKT point is $(0,0)$. This also happens to be the only point, where CQ do not hold. Since the continuous function must attain its minimum on a non-empty bounded and closed set, and KKT conditions are necessary for optimality at all other feasible points, the point $(0,0)$ must the the point of global, hence also local minimum. Indeed: all other feasible points in our domain have positive coordinates, and therefore the minimum of xy occurrs at $(x,y) = (0,0)$.

b) Compute the critical cone at $(0,0)$, and show that there exist directions d contained in the critical cone for which $d^\top \nabla^2 \mathcal{L}((0,0);\lambda^*)d < 0$.

Solution: To find the critical cone C at $(0,0)$, we use the definition given by equation (12.53) in N&W page 330. First, we find the gradients of the constraints at this point:

$$\nabla c_1(0,0) = \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \qquad \nabla c_2(0,0) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Since $\lambda_1 = \lambda_2 = 0$ at this point, we have that $d \in C(0,0)$ if and only if $\nabla c_1(0,0)^T d \geq 0$ and $\nabla c_2(0,0)^T d \geq 0$. The latter condition clearly holds for all d, and so we find that

$$\begin{aligned} C(0,0) &= \{d = (d_1,d_2) \in \mathbb{R}^2 : \nabla c_1(0,0)^T d \geq 0\} \\ &= \{d = (d_1,d_2) \in \mathbb{R}^2 : d_2 \geq d_1\}. \end{aligned}$$

Next, we find that the Hessian of the Lagrangian at $(0,0)$ with Lagrange multipliers $\lambda^* = (\lambda_1,\lambda_2) = (0,0)$ is given by

$$\nabla^2 \mathcal{L}((0,0);(\lambda_1,\lambda_2)) = \nabla^2 \mathcal{L}((0,0);(0,0)) = \nabla^2 f(0,0) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

Thus, $d^T \nabla^2 \mathcal{L}((0,0);(\lambda_1,\lambda_2))d = 2d_1 d_2$. There are clearly directions in the critical cone for which this is negative; one can choose $d_2 > 0$ and $d_1 < 0$.

c) Show that $d^\top \nabla^2 \mathcal{L}((0,0);\lambda^*)d \geq 0$ for all vectors d contained in the tangent cone to the feasible set at $(0,0)$.

Solution: We can find the tangent cone to the feasible set at $(0,0)$ by looking at the limiting vectors along the lines $c_1(\mathbf{x}) = 0$ and $c_2(\mathbf{x}) = 0$ as $\mathbf{x} \to 0$. Traveling toward (0,0) along $c_1(\mathbf{x}) = 0$ we take, for example,

$$z_k = \begin{bmatrix} 1/k \\ 1/k \end{bmatrix}, \quad t_k = ||z_k|| = \sqrt{2}/k,$$

and find the limiting direction

$$d = \lim_{k\to\infty} \frac{z_k}{t_k} = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}.$$

Along $c_2(\mathbf{x}) = 0$, we take

$$z_k = \begin{bmatrix} 1/k \\ 1/k^{3/4} \end{bmatrix}, \quad t_k = ||z_k|| = \frac{\sqrt{\sqrt{k}+1}}{k},$$

and find the limiting direction

$$d = \lim_{k\to\infty} \frac{z_k}{t_k} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

The tangent cone at (0,0) contains all vectors between these limiting cases, which can be shown to be:

$$T(0,0) = \{d \in \mathbb{R}^2 : d_1 \geq 0 \text{ and } d_2 \geq d_1\}.$$

It is then easy to see that $d^T \nabla^2 \mathcal{L}((0,0);(\lambda_1,\lambda_2))d = 2d_1d_2 \geq 0$ for all d in the tangent cone at $(0,0)$.