Norwegian University of Science and Technology
Department of Mathematical Sciences

TMA4180
Optimisation I
Spring 2020

Solutions to exercise set 6

1 Consider a constrained minimisation problem with continuously differentiable objective function and constraints. Which of the following statements are true?

(CQ . . . constraint qualification)

1. x^* is a global minimum $\implies$ x^* is a KKT point.
2. x^* is a local minimum and CQ holds $\implies$ x^* is a KKT point.
3. x^* is a KKT point and CQ holds $\implies$ x^* is a local minimum.
4. x^* is a global minimum and the problem is convex $\implies$ x^* is a KKT point.
5. x^* is a KKT point and the problem is convex $\implies$ x^* is a global minimum.

Solution:

1. False; one needs CQ.
2. True; KKT is necessary for local minima when CQ holds.
3. False; consider the minimization of $-x^2$ subject to $-1 \le x \le 1$, then $x^* = 0$ is a KKT point where CQ holds.
4. False; consider the convex problem of minimizing x, subject to $-x^2 - y^2 \ge 0$.
5. True; KKT is sufficient for global minima when the function is convex.

2 Assume that one wants to solve the optimisation problem

$$\max_x f(x) \qquad \text{such that} \qquad \begin{cases} c_i(x) = 0 & \text{for all } i \in \mathcal{E}, \\ c_i(x) \ge 0 & \text{for all } i \in \mathcal{I}. \end{cases}$$

How do the KKT conditions have to be modified such that one obtains (first order) necessary conditions for this maximisation problem?

Solution: Let

$$\mathcal{L}(x, \lambda) = f(x) - \sum_{i \in \mathcal{E} \cup \mathcal{I}} \lambda_i c_i(x)$$

be the Lagrangian associated with the maximisation problem. Since solving $\max_x f(x)$ is equivalent to solving $\min_x -f(x)$, we can state the KKT conditions for the minimisation problem. To this end, let

$$\widehat{\mathcal{L}}(x, \mu) = -f(x) - \sum_{i \in \mathcal{E} \cup \mathcal{I}} \mu_i c_i(x)$$

be the Lagrangian for the minimisation problem, so that the KKT conditions become

$$-\nabla f(x) - \sum_{i \in \mathcal{E} \cup \mathcal{I}} \mu_i \nabla c_i(x) = \nabla_x \widehat{\mathcal{L}}(x, \mu) = 0,$$

$$c_i(x) = 0 \quad \text{for all } i \in \mathcal{E},$$

$$c_i(x) \geq 0 \quad \text{for all } i \in \mathcal{I},$$

$$\mu_i \geq 0 \quad \text{for all } i \in \mathcal{I},$$

$$\mu_i c_i(x) = 0 \quad \text{for all } i \in \mathcal{E} \cup \mathcal{I}.$$

Since

$$\mathcal{L}(x, -\mu) = -\widehat{\mathcal{L}}(x, \mu) \quad \text{and} \quad \nabla_x \mathcal{L}(x, -\mu) = -\nabla_x \widehat{\mathcal{L}}(x, \mu),$$

we see that changing the signs of the Lagrange multipliers, that is, putting $\lambda = -\mu$, is the only modification in the KKT conditions for the maximisation problem.

3 Consider the constrained optimization problem

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

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:

$$2x^* - \lambda_1^* + \lambda_3^* = 0 \tag{1a}$$

$$2y^* - \lambda_1^* + \lambda_2^* - 2y^* \lambda_3^* = 0 \tag{1b}$$

$$x^* + y^* - 1 \geq 0 \tag{1c}$$

$$2 - y^* \geq 0 \tag{1d}$$

$$y^{*2} - x^* \geq 0 \tag{1e}$$

$$\lambda_i^* \geq 0, \quad i = 1, 2, 3 \tag{1f}$$

$$\lambda_1^*(x^* + y^* - 1) = 0 \tag{1g}$$

$$\lambda_2^*(2 - y^*) = 0 \tag{1h}$$

$$\lambda_3^*(y^{*2} - x^*) = 0. \tag{1i}$$

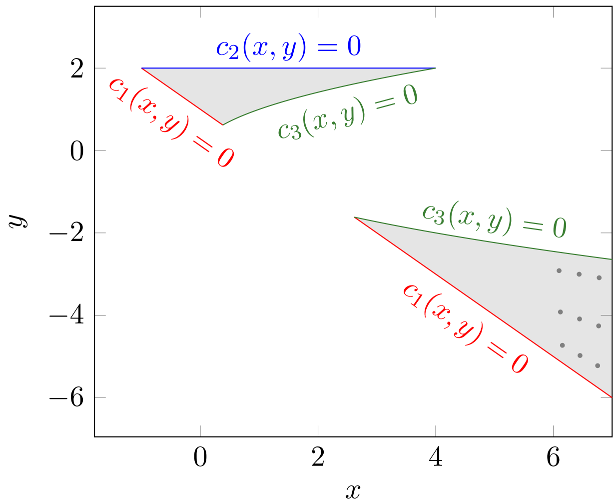

Figure 1: Feasible set. Note: The lower "triangle" extends further toward infinity.

b) Find all KKT points for this optimization problem.

Solution: The feasible set is sketched in Figure 1.
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:

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

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:

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

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

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

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.

4 (Problem 5, Exam 2014) Consider the following constrained optimization problem:

$$\min_{x,y} \frac{1}{2}(x^2+y^2), \quad \text{s.t. } x - y - 1 = 0 \tag{3}$$

a) Find the globally optimal solution (x^*, y^*) for (3) (graphically, if you like). Also find the value of the Lagrange multiplier λ^* associated with the constraint at the globally optimal solution.

Solution: The feasible set and level curves of $f(x,y) = \frac{1}{2}(x^2+y^2)$ are sketched in

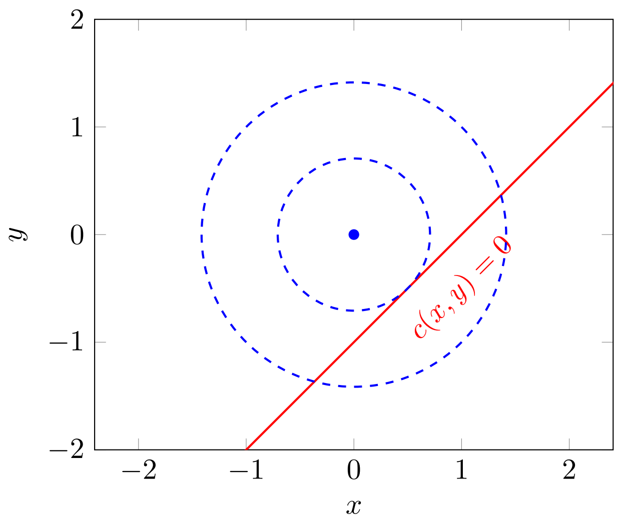

Figure 2: Feasible set and level curves

Figure 2.

To simplify notation, we write $\mathbf{x} = [x, y]^\top$.

We have

$$\mathcal{L}(\mathbf{x}, \lambda) = f(\mathbf{x}) - \lambda c(\mathbf{x}) = \frac{1}{2}(x^2+y^2) - \lambda(x - y - 1),$$

$$\nabla f(\mathbf{x}) = \begin{bmatrix} x \\ y \end{bmatrix}, \qquad \nabla c(\mathbf{x}) = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

and

$$\nabla_{\mathbf{x}} \mathcal{L}(\mathbf{x}, \lambda) = \begin{bmatrix} x + \lambda \\ y - \lambda \end{bmatrix}.$$

By solving the KKT equations

$$\begin{aligned} \nabla_{\mathbf{x}} \mathcal{L}(\mathbf{x}^*, \lambda^*) &= 0 \\ c(\mathbf{x}^*) &= 0, \end{aligned}$$

(or by inspection) we get that

$$x^* = \frac{1}{2}, y^* = -\frac{1}{2}, \lambda^* = \frac{1}{2}.$$

b) Formulate the unconstrained minimization problem corresponding to the application of the quadratic penalty method applied to (3). Solve the resulting unconstrained minimization problem for the penalty parameter $\mu = 2$.

Solution: The unconstrained minimization problem is

$$\min_{\mathbf{x}\in\mathbb{R}^2} Q(\mathbf{x};\mu)$$

where

$$\begin{aligned} Q(\mathbf{x};\mu) =& f(\mathbf{x}) + \frac{\mu}{2}c(\mathbf{x})^2 \\ =& \frac{1}{2}(x^2+y^2) + \frac{\mu}{2}(x-y-1)^2 \end{aligned}$$

$Q(\mathbf{x};\mu)$ is strictly convex for $\mu \geq 0$, and we obtain the unique minimum by solving $\nabla_{\mathbf{x}} Q(\mathbf{x};\mu) = 0$. For $\mu = 2$, we get:

$$\nabla_{\mathbf{x}} Q(\mathbf{x};\mu) = \begin{bmatrix} x + 2(x-y-1) \\ y - 2(x-y-1) \end{bmatrix} = \begin{bmatrix} 3x - 2y - 2 \\ -2x + 3y + 2 \end{bmatrix} = 0$$

with solution

$$x^\mu = \frac{2}{5}, \quad y^\mu = -\frac{2}{5}.$$

For part **c)**, we note that $\|\mathbf{x}^* - \mathbf{x}^\mu\| = \frac{\sqrt{2}}{10}$.

c) State the augmented Lagrangian penalty function corresponding to (3) and the Lagrange multiplier λ and penalty parameter $\mu > 0$. Find the unconstrained global minimum of the augmented Lagrangian corresponding to $\lambda = 0.5$, $\mu = 2$. Compare the accuracy of the obtained approximate solutions to (3) with those obtained in the previous task.

Solution: The augmented Lagrangian penalty function is

$$\begin{aligned} \mathcal{L}_A(\mathbf{x},\lambda;\mu) &= \mathcal{L}(\mathbf{x},\lambda) + \frac{\mu}{2}c(\mathbf{x})^2 \\ &= \frac{1}{2}(x^2+y^2) - \lambda(x-y-1) + \frac{\mu}{2}(x-y-1)^2 \end{aligned}$$

$\mathcal{L}_A$ is a sum of two strictly convex functions $f(\mathbf{x})$ and $\frac{\mu}{2}c(\mathbf{x})^2$, and an affine (thus convex) function $-\lambda(x-y-1)$. As a result $\mathcal{L}_A$ is a strictly convex function of $\mathbf{x}$, and we can find the unique minimum by setting the gradient equal to 0.

If we set $\lambda = \frac{1}{2}, \mu = 2$ and write out the equation $\nabla_{\mathbf{x}}\mathcal{L}_A(\mathbf{x},\lambda;\mu) = 0$, we get

$$\begin{aligned} x - \frac{1}{2} + 2(x-y-1) = 0 \\ y + \frac{1}{2} - 2(x-y-1) = 0 \end{aligned}$$

with solution $x^A = \frac{1}{2}, y^A = -\frac{1}{2}$, *exactly* equal to the exact solution (x^*, y^*).

We see that the quadratic penalty method gave us an error of norm $\frac{\sqrt{2}}{10}$, while with the same error parameter μ, the augmented Lagrangian method gave the exact solution to the constrained problem. (Since we used the exact λ^*. In a practical setting, we would need to estimate λ^* as well.)