Exercise #7

February 28, 2023

Problem 1.

Consider the constrained optimization problem

$$\min_{(x,y)} -x^2 - (y-1)^2 \qquad \text{such that} \qquad \begin{cases} y \geq Cx^2, \\ y \leq 2, \end{cases}$$

where $C > 0$ is some positive parameter.

a) Show that the point $(0,0)$ is a KKT point for all parameters $C > 0$ and that the LICQ is satisfied at $(0,0)$.

b) Formulate the second order necessary and sufficient optimality conditions for the point $(0,0)$. For which parameters C are these conditions satisfied? For which parameters C is the point $(0,0)$ a local minimum?

Solution.

a) Introducing

$$f(x,y) = -x^2 - (y-1)^2, \quad c_1(x,y) = y - Cx^2, \qquad \text{and} \quad c_2(x,y) = 2 - y,$$

the minimisation problem becomes

$$\min_{x,y} f(x,y) \qquad \text{subject to} \qquad c_1(x,y) \geq 0 \text{ and } c_2(x,y) \geq 0.$$

Let also

$$\mathcal{L}(x,y,\lambda_1,\lambda_2) = f(x,y) - \lambda_1 c_1(x,y) - \lambda_2 c_2(x,y)$$

be the Lagrangian, with multipliers λ_1 and λ_2.

Focusing on the KKT conditions, it is clear that $(0,0)$ is feasible. Moreover, from the complementarity conditions

$$\lambda_1 c_1(0,0) = 0 \qquad \text{and} \qquad \lambda_2 c_2(0,0) = 0,$$

we require $\lambda_2 = 0$ because c_2 is inactive. (Note: c_1 is active, so the first condition holds.) Computing

$$\nabla_{x,y}\mathcal{L}(x,y,\lambda_1,0) = \nabla f(x,y) - \lambda_1 \nabla c_1(x,y) = \begin{bmatrix} -2x(1-\lambda_1 C) \\ -2(y-1) - \lambda_1 \end{bmatrix}$$

and demanding that this gradient vanishes at $(x,y) = (0,0)$, then give $\lambda_1 = 2$, with no restriction on C. Hence, $(0,0)$ is a KKT point for all $C > 0$. Additionally, since c_1 is the only active constraint and $\nabla c_1(0,0) = (0,1) \neq 0$, it follows that the LICQ is satisfied as well.

b) Let C be the critical cone at $(0,0)$ with Lagrange multipliers $(\lambda_1, \lambda_2) = (2,0)$. Then $(0,0)$, which satisfies the KKT conditions, is a local minimiser of the constrained problem only if (*necessary condition*) the Hessian

$$\nabla^2_{(x,y)}\mathcal{L}(0,0,2,0) = \begin{bmatrix} -2(1-2C) & 0 \\ 0 & -2 \end{bmatrix}$$

of the Lagrangian is positive semi-definite on C, that is,

$$w^\top \nabla^2_{(x,y)}\mathcal{L}(0,0,2,0)w \geq 0 \quad \text{for all} \quad w \in C.$$

If (*sufficient condition*), however, this Hessian is positive definite on C, then $(0,0)$ is a (strict) local minimiser.

Since c_1 is the only active constraint and $\lambda_1 > 0$, we find that

$$C = \{(w_1, w_2) \in \mathbb{R}^2 : \nabla c_1(0,0)^\top w = 0\} = \{(w_1, 0) \in \mathbb{R}^2 : w_1 \in \mathbb{R}\}.$$

Therefore, with $w = (w_1, 0) \in C$,

$$w^\top \nabla^2_{(x,y)}\mathcal{L}(0,0,2,0)w = -2(1-2C)w_1^2,$$

which is nonnegative if and only if $C \geq 1/2$, and strictly positive for all $w \in C \setminus \{0\}$ if and only if $C > 1/2$. Thus $(0,0)$ is a (strict) local minimum whenever $C > 1/2$, but cannot be a minimiser if $0 < C < 1/2$. It remains to examine $C = 1/2$. To this end, we consider, for example, points (x,y) approaching $(0,0)$ along $c_1(x,y) = 0$, that is, points for which $y = x^2/2 \to 0$. This yields

$$f\left(x, \tfrac{1}{2}x^2\right) = -x^2 - \left(\tfrac{1}{2}x^2 - 1\right)^2 = -\tfrac{1}{4}x^4 - 1,$$

which is strictly less than $f(0,0) = -1$ for all $x \neq 0$. In particular, $(0,0)$ is not a local minimiser when $C = 1/2$.

Problem 2.

Consider the constrained optimisation problem

$$\min_{(x,y)} \frac{1}{2}(x^2 + y^2) \qquad \text{subject to } xy = 1.$$

a) Find (by whatever means) the solutions of this problem. In addition, find the values of the corresponding Lagrange multipliers.

b) Formulate the unconstrained optimisation problem that results from the application of the quadratic penalty method with parameter $\mu > 0$. Solve these problems for all possible parameters μ and verify that the solutions converge to the solutions of the constrained optimization problem as $\mu \to \infty$.

c) Formulate the augmented Lagrangian for this constrained optimization problem and find (for all possible parameters $\lambda \in \mathbb{R}$ and $\mu > 0$) the global solutions of this (unconstrained) optimization problem. For which parameters does one recover the solution of the original constrained problem?

Solution.

a) One strategy: let $f(x,y) = \frac{1}{2}(x^2 + y^2)$ and $c(x,y) = xy - 1$. By completing the square, we get that

$$f(x,y) = \tfrac{1}{2}(x-y)^2 + xy = \tfrac{1}{2}(x-y)^2 + 1,$$

whose global minimisers evidently satisfy $x = y$. And from the constraint $xy = 1$, this gives solutions $(-1,-1)$ and $(1,1)$. Furthermore, at optima, ∇f must be parallel to ∇c, or, $\nabla f = \lambda \nabla c$ for some Lagrange multiplier $\lambda \in \mathbb{R}$. Since $\nabla f(-1,-1) = (-1,-1)$ and $\nabla c(-1,-1) = (-1,-1)$, this gives $\lambda = 1$. At $(1,1)$, we similarly find a corresponding $\lambda = 1$.

(Another option is to set up and solve the KKT conditions, plus argue, for example, via second order sufficient conditions that these points are indeed minima.)

b) Constructively, the quadratic penalty method with parameter $\mu > 0$ seeks to minimise

$$Q(x, y; \mu) := f(x, y) + \frac{\mu}{2}c(x, y)^2 = \frac{1}{2}\left(x^2 + y^2\right) + \frac{\mu}{2}(xy - 1)^2$$

unconstrained over all $(x, y) \in \mathbb{R}^2$. Note that Q is smooth and coercive and thus admits a global minimum, which also must be a stationary point. Calculating

$$\nabla Q(x, y; \mu) = \begin{bmatrix} x + \mu(xy - 1)y \\ y + \mu(xy - 1)x \end{bmatrix},$$

we find that the first component of ∇Q vanishes whenever

$$x = \frac{\mu y}{1 + \mu y^2}.$$

Inserted into the second component of the equation $\nabla Q = 0$, this yields

$$y\left[1 - \frac{\mu^2}{(1 + \mu y^2)^2}\right] = 0. \qquad (\star)$$

If $y = 0$, then $x = 0$ also, so $(0, 0)$ is a stationary point. Examining the Hessian of Q at $(0, 0)$ shows that $\nabla^2 Q(0, 0; \mu)$ is positive definite when $\mu < 1$, and negative definite when $\mu > 1$. Thus $(0, 0)$ is a strict local minimiser when $\mu < 1$ and a strict local maximiser when $\mu > 1$. If $\mu = 1$, then

$$Q(x, y; 1) = \frac{1}{2}\left[(x - y)^2 + (xy)^2 + 1\right] \geq \frac{1}{2} = Q(0, 0),$$

with equality if and only if $x = y = 0$. As such, $(0, 0)$ is a strict local minimiser also for $\mu = 1$.

If $y \neq 0$, then $(\star)$ simplifies to

$$1 + \mu y^2 = \mu,$$

with solutions

$$y = \pm\sqrt{1 - \frac{1}{\mu}},$$

provided $\mu \geq 1$. This also gives

$$x = \frac{\mu y}{1 + \mu y^2} = \pm\sqrt{1 - \frac{1}{\mu}},$$

and it can be verified that these points (x, y) are minimisers. In total, $(0, 0)$ is the global minimiser of $Q(\cdot, \cdot; \mu)$ when $\mu \leq 1$, while the two points

$$(x, y) = \left(\pm\sqrt{1 - \frac{1}{\mu}}, \pm\sqrt{1 - \frac{1}{\mu}}\right)$$

minimise $Q(\cdot, \cdot; \mu)$ when $\mu > 1$. Finally, as $\mu \to \infty$, we find that (x, y) converges to the global minimisers $\pm(1, 1)$ of the original constrained problem.

c) The augmented Lagrangian for this problem is

$$L_A(x, y, \lambda, \mu) = \frac{1}{2}(x^2 + y^2) - \lambda(xy - 1) + \frac{\mu}{2}(xy - 1)^2,$$

which is coercive and lower semi-continuous such that a minimizer exists, and it has the gradient

$$\nabla L_A(x, y, \lambda, \mu) = \begin{bmatrix} x - \lambda y + \mu(xy^2 - y) \\ y - \lambda x + \mu(x^2 y - x) \end{bmatrix}.$$

After a similar computation to that in part b), we find

$$x = \frac{(\mu + \lambda)y}{1 + \mu y^2}$$

and the equation for y:

$$(1 + \mu y^2)^2 = (\lambda + \mu)^2.$$

In addition, we have the solution $(x, y) = (0, 0)$. We must be somewhat careful in finding y. First, we have

$$1 + \mu y^2 = \pm(\lambda + \mu),$$

but since the left hand side is positive, we must choose the right hand side positive as well. Therefore, we have

$$1 + \mu y^2 = |\lambda + \mu|$$

and thus

$$y^* = \pm\sqrt{\left|\frac{\lambda}{\mu} + 1\right| - \frac{1}{\mu}},$$

which exists if $|\lambda + \mu| \geq 1$. It can be checked that here, too, we have $x^* = y^*$. The points (x^*, y^*) are the global minimizers if $\lambda + \mu \geq 1$. Otherwise, $(0, 0)$ is the global minimizer. We see that the original solution is obtained when either $\lambda = 1$ or $\mu \to \infty$. The fact that (x^*, y^*) are the global minimizers if $\lambda + \mu \geq 1$ can seen by checking when $\mathcal{L}_A(x^*, y^*, \lambda, \mu) \leq \mathcal{L}_A(0, 0, \lambda, \mu)$. This leads (after some computation) to the condition

$$(\lambda + \mu - 1)(|\lambda + \mu| - 1) \geq \frac{1}{2}(|\lambda + \mu| - 1)^2.$$

Since (x^*, y^*) exist only if $|\lambda + \mu| \geq 1$, and if $|\lambda + \mu| = 1$ then $(x^*, y^*) = (0, 0)$, we can divide by $|\lambda + \mu| - 1$ to obtain the condition

$$(\lambda + \mu - 1) \geq \frac{1}{2}(|\lambda + \mu| - 1),$$

which holds if $\lambda + \mu \geq 1$ but not if $\lambda + \mu \leq -1$.

Problem 3.

Sketch the region $\Omega \subset \mathbb{R}^2$ defined by the inequalities

$$y \geq x^4 \qquad \text{and} \qquad y \leq x^3,$$

and compute the tangent cone and the set of linearized feasible directions for each point in Ω. For which points in Ω is the LICQ satisfied? (Note that this is same feasible region of the Problem 4.)

Solution.

Defining

$$c_1(x, y) = y - x^4 \quad \text{and} \quad c_2(x, y) = x^3 - y$$

gives $\Omega = \{(x, y) \in \mathbb{R}^2 : c_1(x, y) \geq 0 \text{ and } c_2(x, y) \geq 0\}$, which is shown in Figure 1.

Omitting details—the process is very similar to the Problem 1 of Exercise 6—we obtain that the LICQ condition holds at all feasible points except $(0, 0)$. Moreover, $T_\Omega(x, y) = \mathcal{F}(x, y)$ if (x, y) lies in the interior of Ω;

$$T_\Omega(x, y) = \mathcal{F}(x, y) = \left\{d \in \mathbb{R}^2 : d_2 \geq 4x^3 d_1\right\}$$

when only c_1 is active;

$$T_\Omega(x, y) = \mathcal{F}(x, y) = \left\{d \in \mathbb{R}^2 : 3x^2 d_1 \geq d_2\right\}$$

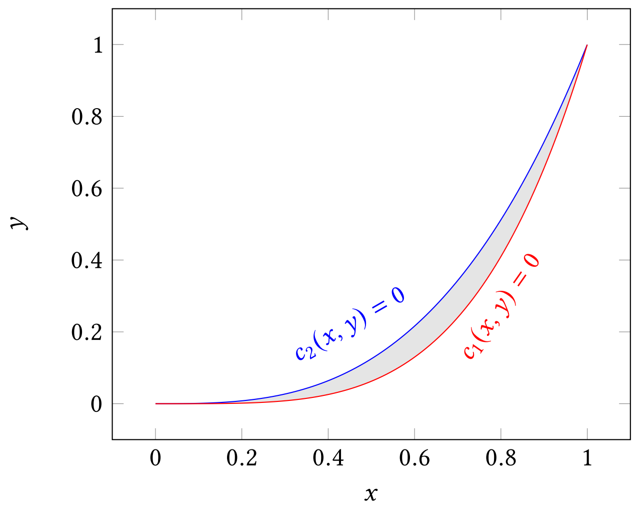

Figure 1: Region Ω in grey, with colors on the boundary specifying the active constraints.

when only c_2 is active;

$$T_\Omega(1,1) = \mathcal{F}(1,1) = \left\{d \in \mathbb{R}^2 : 3d_1 \geq d_2 \geq 4d_1\right\};$$

and

$$\mathcal{F}(0,0) = \left\{d \in \mathbb{R}^2 : d_2 = 0\right\} \quad \text{and} \quad T_\Omega(0,0) = \left\{d \in \mathbb{R}^2 : d_2 = 0 \text{ and } d_1 \geq 0\right\}.$$

Problem 4.

Consider the constrained optimization problem

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

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:

$$1 + 4x^3\lambda_1 - 3x^2\lambda_2 = 0 \tag{1a}$$

$$-\lambda_1 + \lambda_2 = 0 \tag{1b}$$

$$y - x^4 \geq 0 \tag{1c}$$

$$x^3 - y \geq 0 \tag{1d}$$

$$\lambda_i \geq 0, \quad i = 1, 2 \tag{1e}$$

$$\lambda_1(y - x^4) = 0 \tag{1f}$$

$$\lambda_2(x^3 - y) = 0. \tag{1g}$$

Now, we can take a shortcut; from (1b), we see that $\lambda_1 = \lambda_2$, and from (1a) 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 (1a) and (1b)) 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 (here we can simply write $(0,0)$ does not satisfy the condition 3a, therefore, $(0,0)$ cannot be 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.

Problem 5.

Consider the constrained optimisation problem

$$\min_{(x,y)}(x + y) \text{ such that } x^2 + y^2 \leq 1.$$

Formulate a logarithmic barrier method for the solution of this constrained optimisation problem and compute its solution for each parameter $\mu > 0$ in the barrier functional.

Solution.

Constructively, a logarithmic barrier approach may be written as

$$\min_{x,y,s}(x + y - \mu \log s) \text{ subject to } 1 - x^2 - y^2 - s = 0,$$

where $s(\geq 0)$ is the slack variable, and $\mu > 0$ is the barrier parameter which we intend to drive to 0. Introducing a Lagrange multiplier λ, the KKT conditions for this problems are

$$1 + 2x\lambda = 0,\ 1 + 2y\lambda = 0,\ -\frac{\mu}{s} + \lambda = 0 \text{ and } 1 - x^2 - y^2 - s = 0.$$

This gives first that

$$\lambda = \frac{\mu}{s} \text{ and } x = y = -\frac{s}{2\mu},$$

and inserted into the constraint equation, we find that

$$1 - \frac{s^2}{2\mu^2} - s = 0.$$

The relevant solution of this quadratic equation is $s = \mu(\sqrt{\mu^2 + 2} - \mu)$, and we end up with

$$x = y = -\frac{1}{2}(\sqrt{\mu^2 + 2} - \mu) \text{ and } \lambda = (\sqrt{\mu^2 + 2} - \mu)^{-1}.$$

Since the Hessian of the Lagrangian to this problem is positive definite, the found KKT point is the unique global minimizer of the logarithmic barrier formulation. Notably, as $\mu \to 0^+$, we recover the exact solution $x^* = y^* = -\frac{1}{\sqrt{2}}$, with $\lambda^* = \frac{1}{\sqrt{2}}$, of the original problem.