TMA 4180 OPTIMIZATION THEORY

June 4, 2012

Solution with additional comments

(PRELIMINARY VERSION)

1 Problem

Let

$$f(x,y) = \frac{x^2}{2} - \frac{1}{12}x^4 + xy + y^2, \; (x,y) \in \mathbb{R}^2. \tag{1}$$

(a) Find the domain $D_c \subset \mathbb{R}^2$ where f is convex.

(b) Determine all local and global minima of f in the domain

$$D_2 = \left\{(x,y) \in \mathbb{R}^2; \; x \leq 2, \; y \in \mathbb{R}\right\}. \tag{2}$$

Solution:

The gradient of f is

$$\nabla f(x,y) = \left[-\frac{1}{3}x^3 + x + y, \; x + 2y\right], \tag{3}$$

and the Hessian

$$\nabla^2 f(x,y) = \begin{bmatrix} 1 - x^2 & 1 \\ 1 & 2 \end{bmatrix}. \tag{4}$$

(a) The function is convex in all points where $\nabla^2 f(x,y) \geq 0$. This occurs for

$$0 \leq 1 - x^2, \tag{5}$$

$$0 \leq 2, \tag{6}$$

$$0 \leq 2\left(1 - x^2\right) - 1. \tag{7}$$

The third and most restrictive inequality is valid for $x^2 \leq 1/2$, that is, $x \in \left[-1/\sqrt{2}, 1/\sqrt{2}\right]$. Thus,

$$D_c = \left\{(x,y) \in \mathbb{R}^2; \; x^2 \leq 1/2, \; y \in \mathbb{R}\right\}. \tag{8}$$

Alternatively, the smallest eigenvalue of $\nabla^2 f(x,y)$ is

$$\lambda_{\min} = \frac{3 - x^2}{2} - \frac{1}{2}\sqrt{x^4 + 2x^2 + 5}, \tag{9}$$

which is non-negative for $x^2 \leq 1/2$.

(b)

Stationary points are solutions of $\nabla f(x,y) = 0$, or

$$\begin{aligned} -\frac{1}{3}x^3 + x + y &= 0, \\ x + 2y &= 0. \end{aligned} \tag{10}$$

Solutions:

$$(0,0)\,,\ \left(-\sqrt{3/2},\sqrt{3/8}\right),\left(\sqrt{3/2},-\sqrt{3/8}\right). \tag{11}$$

Only $(0,0)$ is within D_c, and since it is not a boundary point, $\nabla^2 f(0,0)$ is positive definite and $(0,0)$ a local minimum. The other two stationary points, in D_2 but outside D_c, are saddle points.

We also need to check the boundary $x = 2$. Along the boundary

$$f(2,y) = y^2 + 2y + \frac{2}{3}, \tag{12}$$

with a minimum $y_0 = -1$. We then consider the gradient in $(2,-1)$:

$$\nabla f(2,-1) = \left[-\frac{5}{3},\ 0\right]. \tag{13}$$

Since $\nabla f(x,y)\cdot d > 0$ for all feasible directions *into* D_2, and y_0 is clearly a minimum along the boundary, we conclude that $(2,-1)$ is also a local minimum.

Obviously, there are no global minima for $x \le 2$, since $f(x,0) = \frac{x^2}{2} - \frac{1}{12}x^4 \to -\infty$ when $x \to -\infty$.

In conclusion, $(0,0)$ and $(2,-1)$ are *local* minima, but there are no *global* minima in D_2.

2 Problem

(a) Write down the Karush-Kuhn-Tucker (KKT) equations for the problem

$$\begin{aligned} &\min_{x\in\mathbb{R}^n} f(x) \\ &c_i(x) = 0, i \in \mathcal{E}, \\ &c_i(x) \ge 0, i \in \mathcal{I}, \end{aligned} \tag{14}$$

and explain what happens if $f(x)$ and $\{-c_i(x)\}_{i\in\mathcal{I}}$ are convex and all $c_i, i \in \mathcal{E}$, are linear.

(b) State the KKT equations for the LP problem

$$\begin{aligned} &\min_{x\in\mathbb{R}^2}\ c'x, \\ &Ax = b, \\ &0 \le x, \end{aligned} \tag{15}$$

when (π, s) are multipliers for equality and inequality constraints, respectively. Show that the corresponding dual problem

$$\begin{aligned} &\max_{\pi} b'\pi, \\ &A'\pi \le c, \end{aligned} \tag{16}$$

has identical KKT equations.

(c) Determine the minimum value of

$$2x_1 + 3x_2 + 3x_3 + 2x_4 \tag{17}$$

when

$$\begin{aligned} x_1 + 3x_2 + x_3 + 0 &= 4, && (18)\\ 2x_1 + x_2 + x_3 + x_4 &= 1, && (19)\\ x_1, x_2, x_3, x_4 &\geq 0. && (20) \end{aligned}$$

by solving the dual problem graphically.

Solution:

(a) We start by forming the Lagrange function, $L(x,\lambda) = f(x) - \lambda' c(x)$. Then the KKT equations are

$$\begin{aligned} (\nabla_x L)' = \nabla f - \sum_{i\in\{\mathcal{E},\mathcal{I}\}} \lambda_i \nabla c_i &= 0, \\ c_i(x) &= 0, i \in \mathcal{E}, \\ c_i(x) &\geq 0, i \in \mathcal{I}, && (21)\\ \lambda' c(x) &= 0,\ \ i \in \mathcal{I} \cup \mathcal{E}, && (22)\\ \lambda_i &\geq 0 \text{ for } i \in \mathcal{I}. && (23) \end{aligned}$$

If the objective function $f(x)$ and $\{-c_i(x)\}_{i\in\mathcal{I}}$ are convex, and all c_i, $i \in \mathcal{E}$, are linear, the Lagrangian will be convex if also $\lambda_i \geq 0$, $i \in \mathcal{I}$. A KKT-point $\{x^*, \lambda^*\}$, where $\lambda_i^* \geq 0$ for $i \in \mathcal{I}$, is then a global minimum since $L(x, \lambda^*)$ is convex, $f(x) \geq L(x, \lambda^*)$, and $f(x^*) = L(x^*, \lambda^*)$.

(b) In the present case, the Lagrange function is

$$L(x,\lambda) = c'x - \pi'(Ax - b) - s'x, \tag{24}$$

and the KKT equations are

$$\begin{aligned} (\nabla_x L)' &= c - A'\pi - s = 0, \\ Ax - b &= 0,\ \ 0 \leq x, \\ s'x &= 0, \\ s &\geq 0. && (25) \end{aligned}$$

Since $s = c - A'\pi$, this may be simplified to

$$\begin{aligned} c - A'\pi &\geq 0, \\ Ax - b &= 0, \\ x &\geq 0. && (26)\\ (c - A'\pi)' x &= 0, && (27) \end{aligned}$$

We write the dual problem as

$$\begin{aligned} &\min_{\pi} -b'\pi, \\ &0 \leq c - A'\pi && (28) \end{aligned}$$

and use x as the Lagrange multiplier. Then

$$\mathcal{L}(\pi, x) = -b'\pi - x'(c - A'\pi)$$

and

$$\begin{aligned} \nabla_\pi L(\pi, x)' &= -b + Ax = 0 \\ (c - A'\pi)' x &= 0, \\ c - A'\pi &\geq 0, \\ x &\geq 0, \end{aligned} \tag{29}$$

which are identical to the KKT-equations for the standard problem.

(c) The dual problem is

$$\begin{aligned} &\max_\pi b'\pi, \\ &A'\pi \leq c, \end{aligned} \tag{30}$$

that is,

$$\begin{bmatrix} 1 & 2 \\ 3 & 1 \\ 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \pi_1 \\ \pi_2 \end{bmatrix} \leq \begin{bmatrix} 2 \\ 3 \\ 3 \\ 2 \end{bmatrix}, \quad b = \begin{bmatrix} 4 \\ 1 \end{bmatrix}. \tag{31}$$

We follow the hint and make a drawing, see Fig. 1. *Unfortunately, the stated b-vector does not result in a bounded solution.* As shown in Fig. 1, the feasible domain is *unbounded* and consists of all points below the lowest of u_1, u_2 and u_4. Moving down along u_2 makes the objective as large as we want. This is turn means that *the primal problem has no feasible points,* and *the primal problem has no solution.*

This was really not the intention!

Again looking at the plot, we observe that for $b = (\alpha, 1)$ to have a solution including the crossing of u_1 and u_2, it is necessary that $1/2 \leq \alpha \leq 3$, and 4 is outside this interval. The solution will include the crossing of u_1 and u_4 if $0 \leq \alpha \leq 1/2$. This covers all possibilities.

3 Problem

Consider the following constrained optimization problem

$$\begin{aligned} &\min_{x \in \mathbb{R}^n} \frac{x'Cx}{2} \\ &\text{when} \quad Ax = b, \end{aligned} \tag{32}$$

$x \in \mathbb{R}^n$, $C \in \mathbb{R}^{n \times n}$, $A \in \mathbb{R}^{r \times n}$, and $b \in \mathbb{R}^r$, $r < n$. Assume that C is positive definite ($C > 0$), and $\operatorname{rank}(A) = r$.

(a) Outline at least one of several methods for solving the problem.

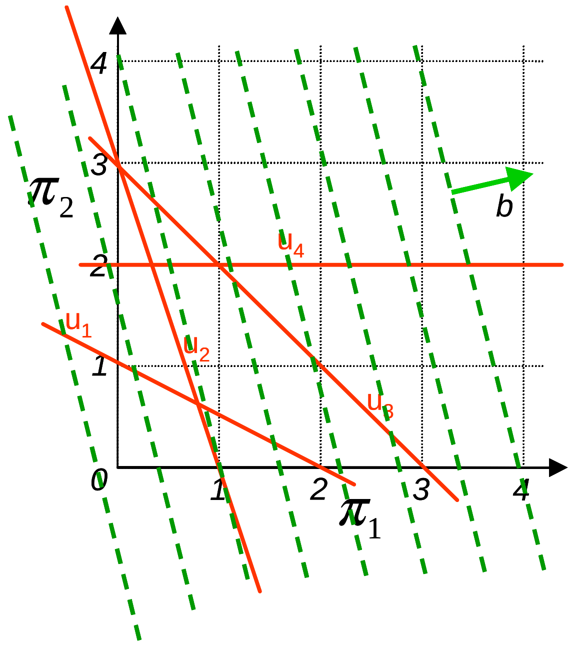

Figure 1: The feasible domain is below the minimum of u_1, u_2 and u_4, and hence unbounded. Moving down along u_2, it follows that the objective may be as large as we wish.

(b) Solve the problem

$$\min_{x\in\mathbb{R}^3}\left\{\frac{x_1^2+2x_2^2+3x_3^2}{2}\right\}$$
$$x_1+x_2+x_3\geq 3, \tag{33}$$

using any suitable method.

Solution:

(a) The problem is a quadratic optimization problem with equality constraints. The matrix C is non-singular and A has full row rank.

(i) Elimination of variables

Bring $Ax=b$ over to the equivalent *reduced Echelon form.* This changes A and b so that

$$\tilde{A}x=[I_{r\times r}\quad N]\left[\begin{array}{c} x_I \\ x_N \end{array}\right]=\tilde{b}. \tag{34}$$

Eliminate x_I from the problem by introducing

$$x=\left[\begin{array}{c} b-Nx_N \\ x_N \end{array}\right] \tag{35}$$

and obtain an unconstrained quadratic problem in x_N (Of course, this is equivalent to solving for x_I, that is, express x_I in terms of x_N and b).

(ii) The null-space method

Since A has full row rank, there are always solutions to $Ax=b$. Assume that one solution is x_0. Then all solutions have the form $x=x_0+Zu$, where Z contains a basis for $N(A)$ and

$u \in \mathbb{R}^{n-r}$. Introduce $x = x_0 + Zu$ into $x'Cx$ and solve the unconstrained quadratic problem for u^*. Note that the quadratic part is $u'Z'CZu$, where $Z'CZ > 0$, and the solution, u^*, is thus unique.

(iii) Solving the KKT equations

The KKT equations are

$$\nabla \mathcal{L}(x, \lambda)' = Cx - A'\lambda = 0, \tag{36}$$

$$Ax = b, \tag{37}$$

which we may write as

$$\begin{bmatrix} C & -A' \\ A & 0 \end{bmatrix} \begin{bmatrix} x \\ \lambda \end{bmatrix} = \begin{bmatrix} 0 \\ b \end{bmatrix} \tag{38}$$

The existence of a unique solution (x^*, λ^*) follows from Thm. 16.1 in N&W. However, in the present case, the first equation gives

$$x = C^{-1}A'\lambda, \tag{39}$$

and the 2nd gives λ:

$$A\left(C^{-1}A'\lambda\right) = b, \tag{40}$$

that is,

$$\begin{aligned} \lambda^* &= \left(AC^{-1}A'\right)^{-1} b, \\ x^* &= C^{-1}A'\lambda^*. \end{aligned} \tag{41}$$

(iv) The Moore-Penrose inverse

A method that has not been discussed in the lectures is to change variables to $y = C^{1/2}x$, since this transforms the problem into

$$\begin{aligned} &\min \|y\|^2, \\ &AC^{-1/2}y = b. \end{aligned} \tag{42}$$

By the "*least norm*"-property of the Moore-Penrose inverse,

$$y^* = \left(AC^{-1/2}\right)^+ b. \tag{43}$$

(b) Since the unique unconstrained minimum ($x = 0$) is outside the feasible domain, the inequality constraint has to be active. The most elementary way is to eliminate one unknown from the problem, say $x_3 = 3 - x_1 - x_2$.We then have to consider the unconstrained minimum of

$$g(x_1, x_2) = x_1^2 + 2x_2^2 + 3(3 - x_1 - x_2)^2. \tag{44}$$

Since the problem is convex, we only go for $\nabla g(x_1, x_2) = 0$, that is,

$$\frac{\partial g}{\partial x_1} = 2x_1 - 6(3 - x_1 - x_2) = 0, \tag{45}$$

$$\frac{\partial g}{\partial x_2} = 4x_2 - 6(3 - x_1 - x_2) = 0, \tag{46}$$

and

$$\begin{aligned}4x_1 + 3x_2 &= 9,\\ 3x_1 + 5x_2 &= 9.\end{aligned} \tag{47}$$

The solution is

$$x_1^* = \frac{18}{11}, \tag{48}$$

$$x_2^* = \frac{9}{11}, \tag{49}$$

$$x_3^* = 3 - x_1^* - x_2^* = \frac{6}{11}. \tag{50}$$

An even simpler alternative is to use the KKT equations for the equality constrained problem,

$$\begin{aligned}x_1 - \lambda &= 0,\\ 2x_2 - \lambda &= 0,\\ 3x_3 - \lambda &= 0,\\ x_1 + x_2 + x_3 &= 3,\end{aligned} \tag{51}$$

with the obvious solution $\lambda^* = \frac{18}{11}$ (yes, it should be non-negative) and the same solution for x^* as above.

4 Problem

A biker moves because she/he is able to produce a thrust (force) U against the ground. However, due to friction in the body and the bike, this force is a nonlinear function of power $P(t)$ produced by the biker. We shall assume that $u \propto P^{1/2}$, where u is produced force measured per mass unit. The biker wants to go up a hill with slope α and length L in a certain time T with minimum energy consumption.

The velocity $y(t)$ follows from *Newton's Law*,

$$\dot{y}(t) = \frac{dy\,(t)}{dt} = u(t) - g\sin\alpha, \tag{52}$$

where g is the acceleration of gravity. The total energy consumption, $\int_0^T P(t)dt$, leads to the functional

$$J(y) = \int_0^T u^2(t)dt = \int_0^T (\dot{y}(t) + g\sin\alpha)^2\,dt. \tag{53}$$

The biker starts with velocity $y\,(0) = 0$, and since the distance to be covered is L,

$$\int_0^T y(t)dt = L. \tag{54}$$

With suitable units, the problem may now be written

$$\begin{aligned} \min_{y\in D} J(y) = \min_{y\in D} \int_0^T (\dot{y}(t)+1)^2\, dt, \\ D = \left\{ y \in C^2 [0,T] ; y(0) = 0 \right\}, \\ \int_0^T y(t)dt = 1. \end{aligned} \tag{55}$$

(a) Show that $J(y)$ as well as the Lagrangian are strictly convex.

(b) Solve the problem for the two situations:

$$\begin{array}{ll} (i) & y(0) = y(T) = 0, \\ (ii) & y(0) = 0,\ y(T) = \text{free}. \end{array} \tag{56}$$

(c) Compute the minimum energy consumption as a function of T for situation (i). The corresponding minimum energy consumption for situation (ii) is

$$J\left(y^*_{(ii)}\right) = \frac{3\left(T^2+2\right)^2}{4T^3}. \tag{57}$$

Discuss the optimal solutions.

Solution:

(a) The integrand $f(x,y,z) = (z+1)^2$ is *strongly* convex since

$$(z+w+1)^2 - (z+1)^2 = 2(z+1)w + w^2 = \frac{\partial (z+1)^2}{\partial z} w + w^2. \tag{58}$$

With one fixed endpoint, this implies that the functional $J(y) = \int_0^T (\dot{y}(t)+1)^2\, dt$ is *strictly convex*. Moreover, $G(y) = \int_0^T y(t)dt$ is linear and hence convex. Thus,

$$L(y,\lambda) = J(y) + \lambda G(y) = \int_0^T \left\{ (\dot{y}(t)+1)^2 + \lambda y(t) \right\} dt, \tag{59}$$

is strictly convex regardless the value of λ.

(b) The (unique) solution is found from $\delta L(y(t,\lambda^*), \lambda^*) = 0$, where λ^* needs to be adjusted so that $y^*(t,\lambda^*)$ satisfies the constraint.

The Euler Equation (EE) for the Lagrangian is

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{y}} - \frac{\partial L}{\partial y} = \frac{d}{dt} 2(\dot{y}(t)+1) - \lambda = 2\ddot{y} - \lambda = 0. \tag{60}$$

and the general solution, with $y(0) = 0$, becomes

$$y(t) = Bt + \frac{\lambda}{4} t^2. \tag{61}$$

For situation (i), $y(T) = 0$, and $B = -\frac{T\lambda}{4}$. The λ to use follows from the constraint

$$\begin{aligned} \int_0^T y(t,\lambda)dt &= \int_0^T \left(-\frac{T\lambda}{4} t + \frac{\lambda}{4} t^2 \right) dt \\ &= \frac{\lambda}{4} \int_0^T \left(-Tt + t^2 \right) dt = -\frac{1}{24} T^3 \lambda = 1. \end{aligned} \tag{62}$$

Hence

$$\lambda^* = -\frac{24}{T^3} \tag{63}$$

and

$$y^*_{(i)} = \frac{6}{T^3} t\left(T - t\right). \tag{64}$$

For a free velocity at the end point we need to apply the natural boundary condition,

$$\frac{\partial f\left(y, \dot{y}\right)}{\partial \dot{y}}\left(T\right) = 2\left(\dot{y}\left(T\right) + 1\right) = 0. \tag{65}$$

Introducing $Bt + \frac{\lambda}{4}t^2$, we obtain

$$B + \frac{\lambda}{2}T + 1 = 0, \tag{66}$$

and

$$y_{(ii)}\left(t, \lambda\right) = -\left(\frac{\lambda}{2}T + 1\right)t + \frac{\lambda}{4}t^2 \tag{67}$$

We finally need to fulfil the constraint:

$$\int_0^T y(t,\lambda)dt = \left(-\frac{\lambda}{2}T - 1\right)\frac{T^2}{2} + \frac{\lambda}{4}\frac{T^3}{3} = 1, \tag{68}$$

giving

$$\lambda = -\frac{3\left(T^2 + 2\right)}{T^3}, \tag{69}$$

and

$$\begin{aligned} y^*_{(ii)}\left(t\right) &= \left(-\frac{\lambda}{2}T - 1\right)t + \frac{\lambda}{4}t^2 \\ &= \left(\frac{3}{T^2} + \frac{1}{2}\right)t - \frac{3}{4T}\left(\frac{2}{T^2} + 1\right)t^2. \end{aligned} \tag{70}$$

(c) The total energy consumption for situation (i) is given by

$$\begin{aligned} J\left(y^*_{(i)}\right) &= \int_0^T \left(\dot{y}(t) + 1\right)^2 dt \\ &= \int_0^T \left(\frac{6}{T^3}\left(T - 2t\right) + 1\right)^2 dt \\ &= \frac{1}{T^3}\left(T^4 + 12\right). \end{aligned} \tag{71}$$

For situation (ii), the similar result is

$$J\left(y^*_{(ii)}\right) = \frac{3}{4T^3}\left(T^2 + 2\right)^2. \tag{72}$$

With the additional freedom in the second situation, we expect that

$$J\left(y^*_{(ii)}\right) \leq J\left(y^*_{(i)}\right), \tag{73}$$

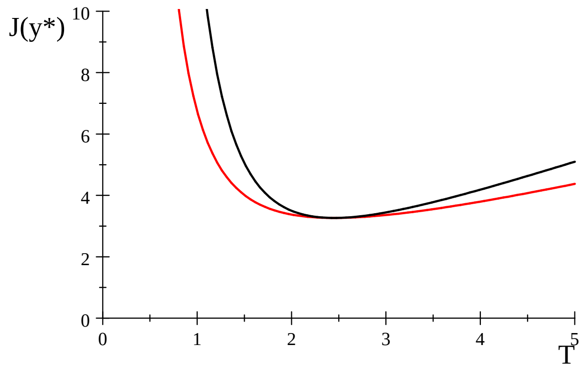

Figure 2: Minimum energy consumption as functions of T for situation (i), black curve, and (ii), red curve. Curves coincide at $T = \sqrt{6}$. See text about a better solution for $T > \sqrt{6}$.

and indeed

$$J\left(y^*_{(i)}\right) - J\left(y^*_{(ii)}\right) = \frac{1}{T^3}\left(T^4 + 12\right) - \frac{3}{4T^3}\left(T^2 + 2\right)^2 = \frac{\left(T^2 - 6\right)^2}{4T^3} \geq 0. \tag{74}$$

Actually, the solutions coincide for $T = \sqrt{6}$. We observe, somewhat surprisingly, that $\lim_{T\to 0} J\left(y^*\right) = \lim_{T\to\infty} J\left(y^*\right) = \infty$ for both solutions. The minimum occurs for $T = \sqrt{6}$. The velocity $y\left(t\right)$ is parabolic, and the terminal velocity for case (ii) comes out negative when $T > \sqrt{6}$. The solution is therefore not particularly interesting since there is no gain of energy into the body from going downhill (!). According to the above, it appears that the best, if $T > \sqrt{6}$, is just to apply the optimal solution for $T = \sqrt{6}$, and then relax for the remaining time. We observe that this otherwise physically acceptable solution is *not* a solution of the Euler equation. The energy consumptions are shown in Fig. 2.