TMA 4180 Optimization Theory
Exam May 21, 2010
Solution with some additional comments
(Revision May 25)

Problem 1

Let

$$f(\mathbf{x}) = 3x^2 - 12x + y^4 - 2y^2 - 5, \quad \mathbf{x} = (x, y) \in \mathbb{R}^2.$$

(a) Compute the gradient and the Hessian of f, and determine the domain $D \in \mathbb{R}^2$ where the function f is strictly convex.

(b) Solve

$$\min_{\mathbf{x}\in\mathbb{R}^2} f(\mathbf{x}).$$

Solution:

(a)

$$\nabla f(\mathbf{x}) = \left(6x - 12, 4y^3 - 4y\right),$$
$$\nabla^2 f = \begin{bmatrix} 6 & 0 \\ 0 & 12y^2 - 4 \end{bmatrix}.$$

The function f is strictly convex when both eigenvalues λ_1 and λ_2 are strictly positive ($\nabla^2 f > 0$). Since

$$\lambda_2 = 6,$$
$$\lambda_1 = 12y^2 - 4,$$

the domain D will be $\left\{(x, y)\,;|y| > 1/\sqrt{3}\right\}$.

(b) Consider the necessary first order condition $\nabla f = 0$, that is,

$$6x - 12 = 0,$$
$$4y^3 - 4y = 0.$$

Moreover, $f(\mathbf{x}) \to \infty$ when $|\mathbf{x}| \to \infty$, so *solutions exist.*

The candidates for solutions are obviously $(2, 0)$, $(2, 1)$, $(2, -1)$. The last two points are in the domain D and are then *strict local minima.* The function values in both points are equal, $f(\mathbf{x}^*) = -18$, and the points are actually global minima since $f(2, 0) = -17$, and $(2, 0)$ is a saddle point (Just checking $f(\mathbf{x})$ for all 3 candidates is also sufficient in the present case).

Problem 2

In Trust Region iterative methods for the unconstrained problem

$$\min_{x \in \mathbb{R}^n} f(x),$$

we consider, for each iteration step, sub-problems of the form

$$\min_{p'p \le \Delta^2} m(p), \tag{1}$$

where

$$m(p) = f(x_k) + b'p + \frac{1}{2}p'Bp. \tag{2}$$

(a) State the Lagrangian and the KKT-equations for the sub-problem in Eq. 1 and 2, and discuss the solution when we assume that B is positive definite, $B > 0$.

(Hint: Since $B > 0$, the equation $(B + \lambda I)p = -b$ has a unique solution for all $\lambda \ge 0$)

(b) How is the size of the trust region adjusted during the iteration?

Solution:

(a) The constraint may be written as $c(p) = (\Delta^2 - p'p)/2 \ge 0$, and the Lagrangian may therefore be written

$$\mathcal{L}(p, \lambda) = f(x_k) + b'p + \frac{1}{2}p'Bp - \lambda\left(\Delta^2 - p'p\right)/2,$$

with the KKT-equations equations

$$\begin{aligned} \nabla_p \mathcal{L}(p, \lambda)' &= Bp + b + \lambda p = 0, \\ \lambda\left(\Delta^2 - p'p\right) &= 0, \\ \Delta^2 - p'p &\ge 0, \\ \lambda &\ge 0. \end{aligned}$$

(Since the hint should actually have been stated for $(B + 2\lambda I)p = -b$, a missing factor $1/2$ has been ignored during the evaluation).

The first equation,

$$(B + \lambda I)p = -b,$$

has a unique solution $p(\lambda)$ for all $\lambda \ge 0$. If $|p(0)| \le \Delta$, then $p^* = p(0)$ is clearly the solution of the subproblem. Otherwise, if $|p(0)| > \Delta$, we increase λ from 0 until $|p(\lambda_0)| = \Delta$. Observe that $p(\lambda) \to 0$ when $\lambda \to \infty$, so such a value $\lambda_0 > 0$ always exists. Since $m(p)$ is strictly convex and domain for p is bounded and convex, the solution to the sub-problem is unique. Verification of the KKT-equations for $p^* = p(\lambda_0)$ is straightforward.

(b) If the current approximate solution is x_k and $x_{k+1} = x_k + p^*$, we consider the ratio

$$\rho = \frac{\text{Actual decrease}}{\text{Estimated decrease}} = \frac{f(x_k) - f(x_{k+1})}{f(x_k) - m(p^*)}.$$

If $\rho \approx 1$, Δ is *increased* for the next subproblem, say $\Delta := 2\Delta$; if $\rho \ll 1$, Δ is *decreased*, say $\Delta := \Delta/2$. Otherwise, Δ is unchanged. Moreover, $x_{k+1} := x_k + p^*$ unless ρ is very small or even negative.

Problem 3

Consider the constrained optimization problem

$$\min_{x\in\Omega} f(x), \tag{3}$$
$$\Omega = \{x \ ; \ c_i(x) \geq 0, \ i \in \mathcal{I}\}, \tag{4}$$

where the objective function $f(x)$ *and* $-c_i(x)$ *are convex for all* $i \in \mathcal{I}$.

(a) Show that Ω *is convex.*

(b) Assume that (x^*, λ^*) *is a KKT-point,*

$$\begin{aligned} \nabla_x \mathcal{L}(x^*, \lambda^*) &= 0, \\ \lambda_i^* \cdot c_i(x^*) &= 0, \ i \in \mathcal{I}, \\ \lambda_i^* &\geq 0, \ i \in \mathcal{I}, \\ x^* &\in \Omega, \end{aligned} \tag{5}$$

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

Show that x^* *is a global minimum for the problem defined in Eqs. 3 and 4.*

(c) Let, for $(x, y) \in \mathbb{R}^2$,

$$f(x,y) = (x-2)^2 + (y+2)^2,$$
$$c_1(x,y) = x - y + 1 \geq 0,$$
$$c_2(x,y) = y \geq 0, \tag{6}$$
$$c_3(x,y) = 4 - (x+1)^2 - y^2 \geq 0, \tag{7}$$

and explain why this is a problem of the form above. Find the solution by making a simple sketch. Show that the solution is a regular KKT-point.

Solution:

(a) First of all, $\Omega = \cap_{i\in\mathcal{I}} \Omega_i$ is convex if each Ω_i is convex. Let $x_1, x_2 \in \Omega_i$ and $x_\theta = \theta x_1 + (1-\theta) x_2$, $\theta \in [0,1]$. Then, since $-c_i$ is convex,

$$-c_i(x_\theta) \leq -\theta c_i(x_1) - (1-\theta) c_i(x_2) \leq 0.$$

Hence, $c_i(x_\theta) \geq 0$ and Ω_i is convex.

(b) We observe that $\mathcal{L}(x, \lambda^*) = f(x) + \sum_{i\in\mathcal{I}} \lambda_i^* (-c_i(x))$ is convex since $\lambda_i^* \geq 0$. Let x be an arbitrary point in Ω:

$$f(x) \geq f(x) + \sum_{i\in\mathcal{I}} \lambda_i^* (-c_i(x)) \tag{8}$$
$$= \mathcal{L}(x, \lambda^*) \tag{9}$$
$$\geq \mathcal{L}(x^*, \lambda^*) + \nabla_x \mathcal{L}(x^*, \lambda^*)(x - x^*) \tag{10}$$
$$= \mathcal{L}(x^*, \lambda^*) = f(x^*), \tag{11}$$

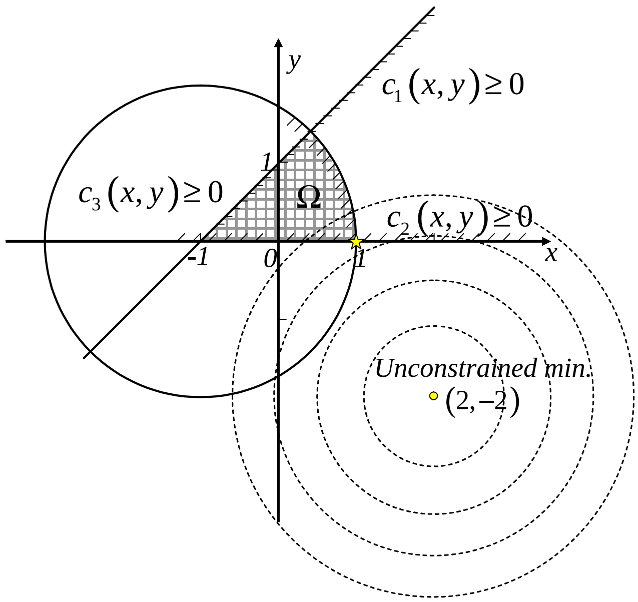

Figure 1: Sketch of $f(x, y)$, and the constraints forming Ω. The obvious solution is indicated by a star, and the global minimum of f is outside Ω.

which is all we need. Alternatively, it is acceptable to say that since $\mathcal{L}$ is convex and $\nabla_x \mathcal{L}(x^*, \lambda^*) = 0$, then x^* is a global minimum for $\mathcal{L}(x, \lambda^*)$. Hence,

$$f(x^*) = \mathcal{L}(x^*, \lambda^*) \leq \mathcal{L}(x, \lambda^*) = f(x) + \sum_{i \in \mathcal{I}} \lambda_i^* (-c_i(x)) \leq f(x). \tag{12}$$

(c) First of all, f is convex since the Hessian is positive definite. The constraints c_1 and c_2 are linear and convex regardless of signs. Moreover, $-c_2$ is strictly convex since it also has a positive definite Hessian.

The domain Ω in the xy-plane and contour lines of f are sketched in Fig. 1. It is obvious that the solution is at $x^* = (1, 0)'$. All constraints are fulfilled, but only $c_2$and c_3 are active. For the gradients, we note that

$$\begin{aligned} &\nabla c_2(\mathbf{x}^*)' = \mathbf{j}, \\ &\nabla c_3(\mathbf{x}^*)' = -2(x^* + 1)\,\mathbf{i} - 2y^*\mathbf{j} = -4\mathbf{i} \end{aligned} \tag{13}$$

Thus,$\nabla c_1(\mathbf{x}^*)$ and $\nabla c_2(\mathbf{x}^*)$ are linearly independent (in fact orthogonal). The solution is at a regular point (LICQ satisfied) and

$$\nabla f(\mathbf{x}^*) = 2(1-2)\,\mathbf{i} + 2(0+2)\,\mathbf{j} = -2\mathbf{i}+4\mathbf{j} \;=4\nabla c_2(x^*) + \frac{1}{2}\nabla c_3(x^*). \tag{14}$$

Also, $\lambda^* = \left(0, 4, \frac{1}{2}\right)' \geq 0$, and all KKT-equations are fulfilled.

Problem 4

Let

$$F(y) = \int_0^1 y'(x)^2 \, dx.$$

Solve the problem

$$\min_{y \in D} F(y),$$
$$D = \left\{ y \in C^1[0,1];\ y(0) = 0,\ y(1) = \text{free} \right\},$$
$$G(y) = \int_0^1 y(x) \, dx = 1.$$

Solution:

Since $f(x, y, z) = z^2$ is strongly convex, and $y(x)$ has one fixed boundary, F is strictly convex. The constraint G is convex since it is linear. The domain D is convex, and we introduce the strictly convex Lagrangian

$$\mathcal{L}(y, \lambda) = \int_0^1 \left(y'(x)^2 + \lambda y(x) \right) dx$$

defined on D. A solution of the Euler equation for $\mathcal{L}$ will be the unique solution to the problem if we are able to find a suitable λ. The Euler equation and the boundary conditions are

$$\frac{d}{dx}(2y'(x)) - \lambda = 2y''(x) - \lambda = 0,$$
$$y(0) = 0,$$
$$\frac{\partial f}{\partial y'}(1) = 2y'(1) = 0.$$

The general solution is easily seen to be

$$y(x, \lambda) = A + Bx + \frac{\lambda}{4}x^2,$$

and the boundary conditions imply that $A = 0$ and $B = -\lambda/2$, so that

$$y(x, \lambda) = \frac{\lambda}{4}x(x-2).$$

It remains to determine λ from the integral constraint:

$$\int_0^1 \frac{\lambda}{4}x(x-2) \, dx = \frac{\lambda}{4}\left(-\frac{2}{3}\right) = 1,$$

which gives $\lambda = -6$. The final solution is therefore

$$y^*(x) = \frac{3}{2}x(2-x).$$

Problem 5

Suppose you are at a gently sloping sand beach, standing in water up to your knees. Running in water is heavier than running on the shore, so if you want to run to a point on the shore in the shortest possible time, you could run straight towards the point, as a dog would do, or alternatively, run the shortest way to the shore, and then on land towards the end point.

Consider the following situation: The shoreline is parallel to the y*-axis, and located at* $x = 2$*, with the sea for* $x < 2$ *and land for* $x > 2$*. Your start-position in the sea is at* $(x, y) = (1, 0)$*, and the end point on the shore is located at* $(2, y_e)$*,* $y_e > 0$*. Your running speed is given by* $v(x) = x$*, so you run twice as fast on the shore, compared to where you stand now (*$v(2) = 2$ *is also your maximal running speed). We assume that the path may be described by the function* $y(x)$*, where* $y(1) = 0$*,* $y(2) = y_e \geq 0$*. Only paths where* $y'(x) \geq 0$ *are of interest since* $y_e > 0$*.*

(a) Show that the variational problem for the total time may be formulated as

$$\min_{y \in \mathcal{D}} J(y),$$

where

$$J(y) = \int_1^2 \frac{\sqrt{1 + y'(x)^2}}{x} dx,$$
$$D = \left\{ y \in C^1[1,2]\,;\ y(1) = 0,\ y(2) = y_e \geq 0, y'(x) \geq 0 \right\}.$$

Prove that J *is a strictly convex functional on the convex domain* D*.*

(b) Write down the Euler equation for the problem in (a), and show that the general solution of the equation is always a part of a circle,

$$(y - a)^2 + x^2 = r^2. \tag{15}$$

Determine the solution when $y(1) = 0$ *and* $y(2) = y_e = 1$*.*

(c) Consider the optimal solution when y_e *increases from 0 towards positive values. What is the (probable) optimal solution when* $y_e \geq \sqrt{3}$*?*

Solution:

(a) The velocity $v = \frac{ds}{dt}$, and

$$ds = \sqrt{1 + y'^2} dx.$$

The total time is therefore

$$\int_{(1,0)}^{(2,y_e)} dt = \int_{(1,0)}^{(2,y_e)} \frac{ds}{v(x)} = \int_1^2 \frac{\sqrt{1 + y'(x)^2}}{x} dx.$$

Here, $y(x) \in C^1[1,2]$ is a sufficient condition for $J(y)$ to exist, whereas the boundary conditions ensure that the start and finish is OK.

The domain D is convex if $y_1, y_2 \in D$ implies that $y_\theta = \theta y_1 + (1 - \theta) y_2 \in D$ for $\theta \in [0, 1]$. All conditions are clearly satisfied for y_θ.

The integrand is strongly convex, since $x > 0$ for $x \in [1,2]$, and $\sqrt{1+z^2}$ is strictly convex:

$$\frac{d^2\sqrt{1+z^2}}{dz^2} = \frac{1}{(z^2+1)^{\frac{3}{2}}} > 0.$$

The functional $J(y)$ is then strictly convex since the end-points are fixed (The direct proof is somewhat cumbersome).

(b) Since there is no y-dependence, the Euler equation becomes

$$\frac{d}{dx}\frac{\partial}{\partial y'}\left(\frac{1}{x}\sqrt{1+y'(x)^2}\right) = \frac{d}{dx}\left(\frac{1}{x}\frac{y'(x)}{\sqrt{1+y'(x)^2}}\right) = 0.$$

Thus,

$$\frac{1}{x}\frac{y'(x)}{\sqrt{1+y'(x)^2}} = C_1$$

where C_1 is an arbitrary constant. Solving for y' we obtain

$$y'(x) = \frac{xC_1}{\sqrt{1-C_1^2x^2}},$$

which leads to the general solution

$$y(x) = \int \frac{xC_1}{\sqrt{1-C_1^2x^2}}dx = -\frac{1}{C_1}\sqrt{1-C_1^2x^2} + C_2.$$

This may be rewritten as

$$(y(x)-C_2)^2 + x^2 = \frac{1}{C_1^2},$$

which defines a circle with center at $(0,a)$, $a = C_2$, and radius $r = 1/C_1$,

$$(y-a)^2 + x^2 = r^2.$$

The boundary conditions lead to

$$a^2 + 1 = r^2,$$
$$(1-a)^2 + 4 = r^2,$$

with the solution $a = 2$ and $r = \sqrt{5}$. The solution may alternatively be written as

$$y(x) = 2 - \sqrt{5-x^2}.$$

(c) Simple geometry shows that, regardless position of $y_e \in [0,\infty)$, it is always possible to find a (center at $(0,a)$) and radius r so as to fit the boundary conditions. However, the solution passes over land $(x > 2)$ where the speed $v(x)$ increases above 2 and $y'(x)$ becomes negative as soon as y_e becomes larger than a. *This violates our assumptions.* By making a sketch, it is easy to see that this occurs if y_e becomes larger than $\sqrt{3}$.

Alternatively, consider the boundary conditions,

$$a^2 + 1 = r^2,$$
$$(y_e - a)^2 + 4 = r^2,$$

giving

$$a = \frac{3 + y_e^2}{2y_e}$$

The condition $y_e < a$ leads to $y_e < \sqrt{3}$.

The optimal solution for $y_e > \sqrt{3}$ thus seems to be to run along the circle $\left(y - \sqrt{3}\right)^2 + x^2 = 4$ to $\left(2, \sqrt{3}\right)$, and then along the shoreline up to y_e.