EXAM
TMA4180 OPTIMIZATION THEORY
May 27, 2008

Solution with additional comments

Preliminary

1 Problem

Estimate the speed of convergence for the Steepest Descent method near the solution of the unconstrained problem

$$\min_{(x,y)\in\mathbb{R}^2} \left\{ (1-x)^2 + \mu\,(y-x)^2 \right\} \tag{1}$$

when μ much larger than 1.

Solution:

This may be viewed as a quadratic penalty formulation of the trivial problem

$$\begin{gathered} \min_{(x,y)\in\mathbb{R}^2} (1-x)^2\,, \\ y = x. \end{gathered} \tag{2}$$

with the solution $(1,1)$. It is easy to compute the Hessian:

$$\nabla^2 f = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix} = \begin{bmatrix} 2+2\mu & -2\mu \\ -2\mu & 2\mu \end{bmatrix}. \tag{3}$$

Eigenvalues follow from the characteristic equation, $\lambda^2 - (4\mu+2)\,\lambda + 4\mu = 0$:

$$\begin{aligned} \lambda_1 &= 2\mu + 1 + \sqrt{4\mu^2+1}, \\ \lambda_2 &= 2\mu + 1 - \sqrt{4\mu^2+1}. \end{aligned} \tag{4}$$

With a condition number $\kappa = \lambda_{\max}/\lambda_{\min}$, the standard convergence estimate is

$$\left\| x_j - x^* \right\|_A \le \frac{\kappa-1}{\kappa+1} \left\| x_{j-1} - x^* \right\|_A. \tag{5}$$

Here,

$$\frac{\kappa-1}{\kappa+1} = \frac{\lambda_1-\lambda_2}{\lambda_1+\lambda_2} = \frac{2\sqrt{4\mu^2+1}}{4\mu+2} \approx 1 - \frac{2}{\mu} + O\left(\frac{1}{\mu^2}\right). \tag{6}$$

2 Problem

(a) What is the content of the Duality Theorem in linear programming?

(b) Show that the following two problems are dual problems (A has full row rank):

$$\begin{array}{cc} (\mathcal{P}) & (\mathcal{D}) \\ \min_x c'x & \max_\lambda b'\lambda \\ Ax \ge b,\ x \ge 0. & A'\lambda \le c,\ \lambda \ge 0. \end{array} \tag{7}$$

(Hint: Consider the KKT-equations and start by using (λ, s) as Lagrange multipliers for $\mathcal{P}$ and (x, u) for $\mathcal{D}$).

(c) Find the minimum value of

$$2x_1 + 2x_2 + 3x_3 + 2x_4 \tag{8}$$

when

$$\begin{aligned} 2x_1 + x_2 + x_3 + 0x_4 &\geq 3 \\ x_1 + 2x_2 + 0x_3 + x_4 &\geq 1 \\ x_i \geq 0,\ i = 1, \cdots, 4. \end{aligned} \tag{9}$$

Solution:

(a) The Dual and Primal Problems have equivalent KKT-equations, and variables and Lagrange multipliers switch place. The Duality Theorem states that if any of the problems are unbounded, the other is infeasible. Moreover, the optimal objective values are equal and, since one is a minimum and the other a maximum problem, objectives values for the two are separated by the optimal objective value on the real line.

(b) In order to establish the statement, we need to look at the KKT-equations. For the primal problem ($\mathcal{P}$), and using that $\mathcal{L}_{\mathcal{P}}(x, \lambda, s) = c'x - \lambda'(Ax - b) - s'x$, we obtain

$$\begin{aligned} \nabla_x \mathcal{L}'_{\mathcal{P}} &= c - A'\lambda - s = 0, \\ Ax - b &\geq 0, \\ \lambda'(Ax - b) &= 0, \\ s'x &= 0, \\ \lambda, s, x &\geq 0. \end{aligned} \tag{10}$$

Solving for s,

$$\begin{aligned} \lambda'(Ax - b) &= 0, \\ (c - A'\lambda)' x &= 0,, \\ Ax - b &\geq 0, \\ c - A'\lambda &\geq 0, \\ \lambda, x &\geq 0. \end{aligned} \tag{11}$$

For the dual problem we use $\mathcal{L}_{\mathcal{D}}(\lambda, x, u) = -b'\lambda - x'(c - A'\lambda) - u'\lambda$:

$$\begin{aligned} \nabla_\lambda \mathcal{L}'_{\mathcal{D}} &= -b + Ax - u = 0, \\ x'(c - A'\lambda) &= 0, \\ c - A'\lambda &\geq 0 \qquad (12) \\ u'\lambda &= 0, \qquad (13) \\ \lambda, u, x &\geq 0. \end{aligned}$$

Eliminating u leads directly to Eqn. 11. This establishes the correspondence.

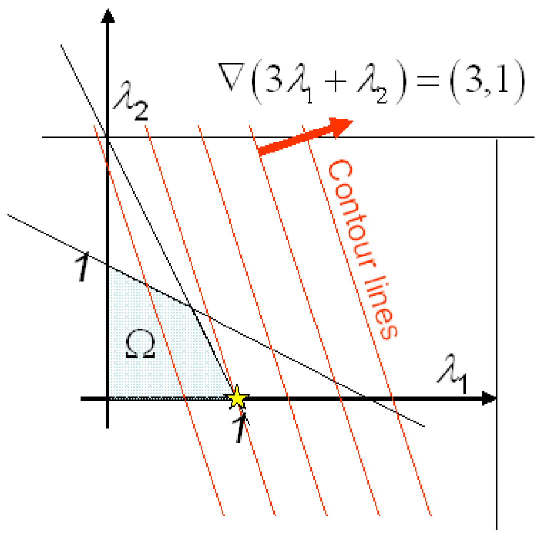

Figure 1: Solution of the dual problem is found at $\lambda_1 = 1$, $\lambda_2 = 0$.

(c) We consider the dual problem, which may be read directly from (b):

$$\max_{\lambda} (3\lambda_1 + \lambda_2),$$
$$\begin{bmatrix} 2 & 1 \\ 1 & 2 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \lambda \leq \begin{bmatrix} 2 \\ 2 \\ 3 \\ 2 \end{bmatrix}, \ \lambda \geq 0. \tag{14}$$

This is easy to solve graphically, see Fig. 1. Since $\min c'x = \max \lambda' b = 3\times 1+0 = 3$, the minimum we seek is 3.

Even if it is not asked for, finding x is easy: Since λ is known, we apply $(c - A'\lambda)' x = 0$ to show that only $x_1$may be different from 0, and $x_2 = x_3 = x_4 = 0$. Looking at the original problem it follows that

$$x^* = \left\{ \frac{3}{2}, 0, 0, 0 \right\}.$$

3 Problem

In a typical Trust Region iteration we have, at iteration n, a sphere D_n with diameter Δ_n around the current iterate x_n. We then form a quadratic approximation

$$m(p) = f(x_n) + \nabla f(x_n) p + \frac{1}{2} p' B_n p$$

and solve for

$$p_n = \arg\min_{p \in D_n} m(p).$$

(a) Explain how we move to x_{n+1} and adjust the diameter of the sphere.

(b)The standard form of the Trust Region sub-problem is

$$\min_{\|x\| \leq 1} q(x), \tag{15}$$

where $q(x) = \frac{1}{2}x'Bx - b'x$ *and the constraint is written* $c(x) = 1 - x'x \geq 0$*. Below we only consider a situation where* B *is positive definite* ($B > 0$)*. Will the problem always have a unique solution? State the KKT-equations and discuss their solution when the global solution* x_g *of the unconstrained problem* ($x_g = \arg\min_{x\in\mathbb{R}^n} q(x)$) *is inside or outside the domain defined by* $c(x) \geq 0$*.*

Solution:

(a) After solving (perhaps approximately) the sub-problem

$$p_n = \arg\min_{p\in D} m(p), \tag{16}$$

we form

$$\rho_n = \frac{f(x_n) - f(x_n + p_n)}{m(0) - m(p_n)} = \frac{\text{Actual decrease}}{\text{Estimated decrease}} \tag{17}$$

($x_n + p_n$ is tentatively x_{n+1}). If $\rho \approx 1$, the approximation $f(x_n + p) \approx m(p)$ works very well, and we move to $x_{n+1} = x_n + p_n$ and increase Δ, $\Delta := \alpha\Delta$, $\alpha > 1$. If $\rho \ll 1$, the approximation is bad; we stay at x_n and try a smaller Δ, $\Delta := \beta\Delta$, $\beta < 1$. Otherwise, we move to $x_{n+1} = x_n + p_n$ and keep the size of Δ.

(b) The problem is

$$\begin{aligned} &\min \left\{ \frac{1}{2}x'Bx - b'x \right\} \\ &\text{when } 1 - x'x \geq 0. \end{aligned} \tag{18}$$

(Here it was assumed that symmetry is part of the definition of *positive definite*. If you don't like that, replace B by $(B + B')/2$ below). We observe that since $B > 0$, the objective is strictly convex and the domain we are considering, $D = \{x; \|x\| \leq 1\}$ is convex as well. Since D is bounded, we always have a unique solution (which is found by solving the KKT-equations). With the Lagrangian $\mathcal{L}(x, \lambda) = \frac{1}{2}x'Bx - b'x - \lambda(1 - x'x)$, we obtain the KKT-equations

$$\begin{aligned} &\nabla_x \mathcal{L}(x, \lambda)' = Bx - b + \lambda x = 0, \\ &\lambda\left(1 - x'x\right) = 0, \\ &\qquad \lambda \geq 0, 1 - x'x \geq 0, \end{aligned} \tag{19}$$

or

$$\begin{aligned} &(B + I\lambda) x = b, \\ &\lambda\left(1 - x'x\right) = 0, \\ &\quad 1 - x'x \geq 0,\ \lambda \geq 0. \end{aligned}$$

The following situations occur:

1. If $\|x_g\| \leq 1$, then $x_g \in D$, $x^* = x_g = B^{-1}b$, and $\lambda = 0$ satisfies all KKT-equations (this also includes the special case when $\|x_g\| = 1$).

2. If $\|x_g\| > 1$, then the overall minimum is outside D and x^* has to be on the boundary of D (since $\nabla\left(\frac{1}{2}x'Bx - b'x\right) \neq 0$ for all $x \in D$). Note that with $x_\lambda \triangleq (B + I\lambda)^{-1} b$, we have $x_0 = x_g$ and $\lim_{\lambda\to\infty} x_\lambda = 0$. We thus need to find (the unique!) $\lambda^* > 0$ so that $x^* = x_{\lambda^*}$ and $\|x_{\lambda^*}\| = 1$.

4 Problem

Consider a functional J defined for all continuous functions $y(x)$ on $[0,1]$ ($y \in C[0,1]$) as

$$J(y) = \int_0^1 f(x, y(x))\, dx, \tag{20}$$

where $f(x,y)$ is smooth and strictly convex in y for all fixed x in $[0,1]$.

(a) Write down the Gâteaux derivative (directional derivative) of J and show that J is strictly convex (Use that $f(x, y+v) - f(x,y) \geq \frac{\partial f}{\partial y}(x,y)\, v$, with strict inequality unless $v = 0$).

(b) Show that an acceptable solution y^* (that is, $y^* \in C[0,1]$) of

$$\frac{\partial f}{\partial y}(x, y^*) = 0,\ \ x \in [0,1], \tag{21}$$

is a unique minimizer for $J(y)$.

A gardener delivers flowers to the wholesale market (=grossistmarkedet). The wholesale basis price $p_0(x)$ per unit varies a lot with the time of the year ($0 \leq x \leq 1$). In addition, if the gardener provides an amount $y(x)$ to the market mafia, the actual price she/he obtains may be modeled as

$$p(x) = p_0(x)\,(1 - \alpha y(x)),\ \ \alpha > 0. \tag{22}$$

(Since flowers have to be sold immediately after they are produced, the price is highly negotiable (=diskutabel)!). The gardener's income for one year is

$$P(y) = \int_0^1 p(x)\, y(x)\, dx. \tag{23}$$

In addition, the overall yearly production is limited by regulations, so that

$$G(y) = \int_0^1 y(x)\, dx = 1. \tag{24}$$

(c) Based on this simplified model, determine the optimal production strategy $y(x)$ for maximizing $P(y)$ when $\alpha \geq 1/2$ and $p_0(x) \geq 1 > 0$. What happens when $\alpha \gg 1$, and what is the obvious (but unrealistic) strategy when $\alpha \to 0$? (Don't forget we seek the minimum of $-P(y)$!)

Solution:

(a) The derivative is computed in the standard way,

$$\delta J(y, v) = \frac{d}{d\varepsilon} \int_0^1 f(x, y + \varepsilon v)\, dx \bigg|_{\varepsilon = 0} = \int_0^1 \frac{\partial f}{\partial y}(x, y(x))\, v(x)\, dx. \tag{25}$$

To prove that J is convex, we check the definition of a convex functional:

$$\begin{aligned} J(y+v) - J(y) &= \int_0^1 \left[f(x, y+v) - f(x,y) \right] dx \\ &\geq \int_0^1 \frac{\partial f}{\partial y}(x, y(x))\, v(x)\, dx = \delta J(y, v). \end{aligned} \tag{26}$$

Since the inequalities for the integral kernel are sharp whenever $v(x) \neq 0$, it follows that J is strictly convex.

It is also possible to say, referring to the theory in Troutman, that since the assumptions imply that f is partly strongly convex, J is strictly convex.

(b) If $y^* \in C[0,1]$ solves

$$\frac{\partial f}{\partial y}(x, y^*(x)) = 0,\ x \in [0,1], \tag{27}$$

we have for all $y \in C[0,1]$ that

$$J(y) - J(y^*) = J(y^* + (y - y^*)) - J(y^*) \geq \int_0^1 \frac{\partial f}{\partial y}(x, y^*(x))(y(x) - y^*(x))\,dx = 0. \tag{28}$$

Hence, y^* is a global minimum, and unique since J is strictly convex.

(c) We formulate the problem as a constrained variational minimization by means of the Lagrangian,

$$\mathcal{L}(y) = -P(y) + \lambda G(y) = \int_0^1 \left[p_0(x)\left(\alpha y^2 - y\right) + \lambda y\right] dx. \tag{29}$$

Since $p_0(x)$ is positive for all x and $\alpha > 0$, $f(x,y) = p_0(x)\left(\alpha y^2 - y\right) + \lambda y$ will always be strictly convex in y. It is therefore sufficient to solve

$$\frac{\partial}{\partial y}\left[p_0(x)\left(\alpha y^2 - y\right) + \lambda y\right] = 2\alpha p_0(x)\, y - p_0(x) + \lambda = 0, \tag{30}$$

and adjust λ so that $G(y) = 1$. Now,

$$y(x, \lambda) = \frac{1}{2\alpha}\left[1 - \frac{\lambda}{p_0(x)}\right], \tag{31}$$

and $G(y) = 1$ leads to

$$1 = \frac{1}{2\alpha}\left[1 - \lambda \int_0^1 \frac{dx}{p_0(x)}\right], \tag{32}$$

Finally,

$$y^*(x) = \frac{1}{2\alpha}\left[1 + (2\alpha - 1) \times \frac{\frac{1}{p_0(x)}}{\int_0^1 \frac{dx}{p_0(x)}},\right]. \tag{33}$$

Since $\alpha \geq 1/2$, we see that $y^*(x) > 0$ over the whole year. When $\alpha = 1/2$, the production may even be kept constant. As α gets bigger, the optimal production strategy approaches

$$y^*_\infty(x) = \frac{1}{\int_0^1 \frac{dx}{p_0(x)}} \frac{1}{p_0(x)}. \tag{34}$$

This means that the production is low when the price is high, and vice versa! When $\alpha < 1/2$, is possible that $y^*(x)$ becomes 0, and it pays to stop the production for some periods with low prices (when α decreases, $y^*(x) = 0$ is first reached around the minimum of $p_0(x)$. If $\alpha = 0$, the optimal solution is of course to sell the complete yearly production around the time of maximum price!

5 Problem

Consider the following functionals and their definition domains:

$$J_1(y) = \int_0^1 \left[y'(x)^2 + 12xy(x) \right] dx, \; y \in C^2[0,1], \; y(0) = 0, \; y(1) = 1, \tag{35}$$

$$J_2(y) = \int_0^{\pi/2} \left[y'(x)^2 - y(x)^2 \right] dx, \; y \in C^2[0,\pi/2], \;\; y(0) = 0, \; y(\pi/2) = 1, \tag{36}$$

$$J_3(y) = \int_0^{2\pi} \left[y'(x)^2 - y(x)^2 \right] dx, \; y \in C^2[0,2\pi], \; y(0) = 0, \; y(2\pi) = 0. \tag{37}$$

Find the stationary points (functions) by solving the Euler equations. Try to determine whether they are minima (For J_2 and J_3, a reasonable discussion without conclusion is sufficient for a full score!).

Solution:

All functionals have the standard form, and we recall that for $F(y) = \int_a^b f(x, y, y')\, dx$, the derivative is given by

$$\delta F(y; v) = \left[\frac{\partial f}{\partial y'} v \right]_a^b - \int_a^b \left[\frac{d}{dx} \frac{\partial f}{\partial y'} - \frac{\partial f}{\partial y} \right] v dx. \tag{38}$$

In the present cases, the allowed functions have fixed end-points, in which case it is sufficient that y^* solves the Euler equation

$$\frac{d}{dx} \frac{\partial f}{\partial y'} - \frac{\partial f}{\partial y} = 0, \tag{39}$$

and satisfies the boundary conditions. The Euler equation for J_1 is

$$2y''(x) + 12x = 0, \tag{40}$$

with the general solution

$$y(x) = x^3 + c_1 + c_2 x. \tag{41}$$

The boundary conditions imply that $y^*(x) = x^3$ is the unique solution. This solution is also a unique minimum of J_1 since the function

$$f_1(x, y, z) = z^2 + 12xy \tag{42}$$

is partially strongly convex, and hence J_1 is *strictly convex.*

For J_2 and J_3 the Euler equation is

$$y'' + y = 0. \tag{43}$$

The boundary conditions for J_2 implies the unique solution

$$y^*(x) = \sin x. \tag{44}$$

Since the function $f(x, y, z) = z^2 - y^2$ is *not* partially strongly convex, J_2 may not be convex, and hence the solution may not be a minimum. The same conclusion applies to J_3, but here the problem

$$\begin{aligned} y'' + y &= 0, \\ y(0) = y(2\pi) &= 0, \end{aligned} \tag{45}$$

has many solutions, the most obvious being

$$y^*(x) = a\sin x,\ \ a \in \mathbb{R}. \tag{46}$$

The functional has the constant value 0 for all of these

$$J_3(y) = \int_0^{2\pi} \left[a^2\cos^2 x - a^2\sin^2 x\right] dx = 0. \tag{47}$$

Without further investigations, we can not say whether these functions are minima, and *the solution ends here.*

Addendum: The functional J_3 has *no* global minimum: The simplest is to consider the function

$$y_0(x) = b\sin\frac{x}{2}. \tag{48}$$

The function is smooth and $y_0(0) = y_0(2\pi) = 0$, but it is not a stationary point. However,

$$J_3(y_0) = \int_0^{2\pi} \left[b^2\frac{1}{4}\cos^2\frac{x}{2} - b^2\sin^2\frac{x}{2}\right] dx = -b^2\frac{3}{4}\pi, \tag{49}$$

which can be made as small as we want by increasing b.

However, the functional J_2 is even strictly convex. Consider

$$J_2(y+v) - J_2(y) = \int_0^{\pi/2} 2\left(y'v' - yv\right) dx + \int_0^{\pi/2} \left(v'^2 - v^2\right) dx. \tag{50}$$

The first term on the RHS is just $\delta J_2(y,v)$, so it is sufficient to show that

$$\int_0^{\pi/2} \left(v'^2 - v^2\right) dx \geq 0. \tag{51}$$

for all allowed variations, that is, $v \in C^2[0,\pi/2]$ and $v(0) = v(\pi/2) = 0$.

I see no other way to decide this than applying a bit of Fourier analysis: The functions $\left\{\sin k\frac{\pi}{L}x\right\}_{k=1}^{\infty}$ and (separately) the set $\left\{\cos k\frac{\pi}{L}x\right\}_{k=1}^{\infty}$ make up complete, orthogonal systems for $L^2[0,L]$. Then, for all absolutely continuous functions u on $[0,L]$ with $u' \in L^2[0,L]$, we have the L^2-identities

$$\begin{aligned} u(x) &= \sum_{k=1}^{\infty} b_k \sin kx, \\ u'(x) &= \sum_{k=1}^{\infty} b_k \frac{k\pi}{L}\cos kx. \end{aligned} \tag{52}$$

This collection of functions contains all allowed variations v for J_2, and if we now apply Parseval's Formula,

$$\|u\|^2 = \int_0^L u(x)^2\, dx = \frac{L}{2}\sum_{k=1}^{\infty} b_k^2 \tag{53}$$

(and similarly for u'), we obtain

$$\int_0^L \left(u'(x)^2 - u(x)^2\right) dx = \frac{L}{2}\sum_{k=1}^{\infty}\left(\left(\frac{k\pi}{L}\right)^2 - 1\right) b_k^2. \tag{54}$$

As long as $L \leq \pi$, all terms in the series are non-negative. For J_2, where $L = \pi/2$, the integral in 51 can never be negative and J_2 is indeed strictly convex, since the sum is positive unless $b_k = 0$, $k = 1, \cdots$. However, if $L = 2\pi$, as is the case for J_3, this is no longer true. The transition occurs at $L = \pi$.