TMA 4180: Optimeringsteori

Eksam May 24, 2007

Problems and solutions

(Preliminary version)

This solution is more detailed than what is required for a full score.

Problem 1:

The function f is defined for all $(x, y) \in \mathbb{R}^2$:

$$f(x, y) = x^2 + 2y^2 - 2xy - 2y^3 + y^4. \tag{1}$$

Determine the global minima (if they exist).

Solution: We start by computing ∇f and $\nabla^2 f$:

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

Candidates for solutions will be where $\nabla f(x, y) = 0$, or

$$\begin{aligned} y &= x, \\ 4y - 6y^2 + 4y^3 &= 2x. \end{aligned} \tag{3}$$

The solutions of Eqn. 3 are easily seen to be

$$\begin{aligned} x_{(1)} &= (0, 0)', \\ x_{(2)} &= \left(\frac{1}{2}, \frac{1}{2}\right)', \\ x_{(3)} &= (1, 1)'. \end{aligned} \tag{4}$$

We check the Hessians:

$$\begin{aligned} \nabla^2 f(0, 0) &= \begin{bmatrix} 2 & -2 \\ -2 & 4 \end{bmatrix} > 0, \\ \nabla^2 f\left(\frac{1}{2}, \frac{1}{2}\right) &= \begin{bmatrix} 2 & -2 \\ -2 & 1 \end{bmatrix}, \text{ indefinite}, \\ \nabla^2 f(1, 1) &= \begin{bmatrix} 2 & -2 \\ -2 & 4 \end{bmatrix} > 0. \end{aligned} \tag{5}$$

Thus, $(0, 0)'$ and $(1, 1)'$ are minima, and both are strict since the Hessian is positive definite. The function values in both minima are equal to 0. The point in the middle is actually a saddle point, and the function value is also larger (equal to 1/16). Both $(0, 0)'$ and $(1, 1)'$ are *global* because $f(x, y) \to \infty$ when $\|(x, y)\| \to \infty$.

The problem may also be solved by the following *trick*:

$$f(x, y) = x^2 + 2y^2 - 2xy - 2y^3 + y^4 = (y - x)^2 + \left(y - y^2\right)^2. \tag{6}$$

The global minimum is 0, and this is obtained for $y = y^2$, $y = x$, or $(0, 0)'$ and $(1, 1)'$.

Problem 2

(a) Show that $x_0 = (2, 2, -1)'/3$ *is the only KKT point for the problem*

$$\begin{aligned} &\min_{x \in \mathbb{R}^3} (x_1 + x_2) \\ c_1(x) &= x_1 + x_2 + x_3 - 1 = 0 \\ c_2(x) &= x_1^2 + x_2^2 + x_3^2 - 1 \geq 0 \end{aligned} \tag{7}$$

(b) Is x_0 *really a solution? (Hint: Illustrate the situation as seen in the plane* $c_1(x) = 0$*, unless you apply the second order conditions)*

Solution:

(a) We observe that Ω is unbounded, and that $f(x) = x_1 + x_2$ is unbounded below on Ω. We can only hope for *local* minima, and $\nabla \mathcal{L} = 0$ gives

$$\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \lambda_1 + \begin{pmatrix} 2x_1 \\ 2x_2 \\ 2x_3 \end{pmatrix} \lambda_2. \tag{8}$$

There are no solutions for $\lambda_2 = 0$. Thus, $\lambda_2 \neq 0$, and the inequality constraint is active. Subtracting the first two equations shows that $x_1 = x_2$, and we are left with four equations ($\nabla \mathcal{L} = 0$ and the constraints):

$$\begin{aligned} \lambda_1 + 2x_1\lambda_2 &= 1, \\ \lambda_1 + 2x_3\lambda_2 &= 0, \\ 2x_1 + x_3 &= 1, \\ 2x_1^2 + x_3^2 &= 1. \end{aligned} \tag{9}$$

The two last equations have solutions

$$x_a = (0, 0, 1)', \; x_b = (2, 2, -1)'/3. \tag{10}$$

The multipliers are found from the first pair:

$$\lambda_a = (1, -1/2)', \; \lambda_b = (1/3, 1/2)'. \tag{11}$$

Since λ_2 needs to be larger than 0, the only KKT-point is $x_b = x_0$.

(b) The intersection between the surface of the sphere ($c_2(x) = 0$) and the plane ($c_1(x) = 0$) is a circle, and Ω is the domain in the plane outside the circle. The contours (level curves) of f are still straight lines, and the situation is illustrated in fig. 1. The points x_a and x_b are marked, and we observe that x_a is actually a solution if c_2 has changed sign.

The first order conditions are not sufficient to decide whether x_b is a minimum, but moving along the the rim of Ω (the circle), it is clear that this is not the case.

It is also possible to check the Hessian of the Lagrange function

$$\begin{aligned} &\nabla_x^2 \left[x_2 + x_3 - \lambda_1^* (x_1 + x_2 + x_3 - 1) - \lambda_2^* \left(x_1^2 + x_2^2 + x_3^2 - 1 \right) \right] \\ &= -2\lambda_2^* I_{3\times 3} = -I_{3\times 3}. \end{aligned} \tag{12}$$

The second order test considers the *projected Hessian*, $Z' \nabla_x^2 L(x_0, \lambda^*) Z$, where $Z = [z_1, \cdots, z_j]$ is a basis for $\mathcal{N}(A)$,

$$A = \begin{bmatrix} \nabla c_1(x_a) \\ \nabla c_1(x_b) \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ \frac{4}{3} & \frac{4}{3} & -\frac{2}{3} \end{bmatrix}. \tag{13}$$

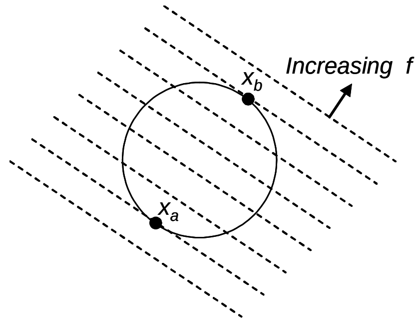

Figure 1: The problem in the plane $x_1 + x_2 + x_3 = 1$.

Since A has rank 2, the null space $\mathcal{N}(A)$ is spanned by $\begin{pmatrix} -1 & 1 & 0 \end{pmatrix}'$. But no (non-zero) projection of $-I_{3\times 3}$ will ever be positive semi-definite, so we conclude that *no local minimum exists.*

Problem 3

(a) Express the following LP-problem in Standard Form

$$\begin{aligned} &\min_x x'c, \\ &Ax \geq b, \\ &x,\ c \in \mathbb{R}^n, b \in \mathbb{R}^m,\ A \in \mathbb{R}^{m\times n}. \end{aligned} \tag{14}$$

(b) The dual problem to the LP-problem

$$\begin{aligned} &\min_x x'c, \\ &Ax = b,\ x \geq 0, \end{aligned} \tag{15}$$

is

$$\begin{aligned} &\max b'\lambda, \\ &A'\lambda \leq c. \end{aligned} \tag{16}$$

How is the dual problem defined, and what is the most important result about the dual problem?

(c) Apply duality to determine the minimum value of the objective function

$$f(x) = 2x_1 + 6x_2 + 3x_3 + 5x_4 + 2x_5 \tag{17}$$

when

$$\begin{aligned} x_1 + x_2 + 0 + x_4 - 2x_5 = 1, \\ -x_1 + x_2 + x_3 + 0 - x_5 = 1, \\ x_i \geq 0,\ i = 1, \cdots, 5. \end{aligned} \tag{18}$$

Solution:

(a) Since we need non-negative variables and an equality constraint instead of an inequality, we introduce $x = y - z,\ y, z \geq 0$, and a slack variable $s \geq 0$ so that

$$A(y - z) + s = b. \tag{19}$$

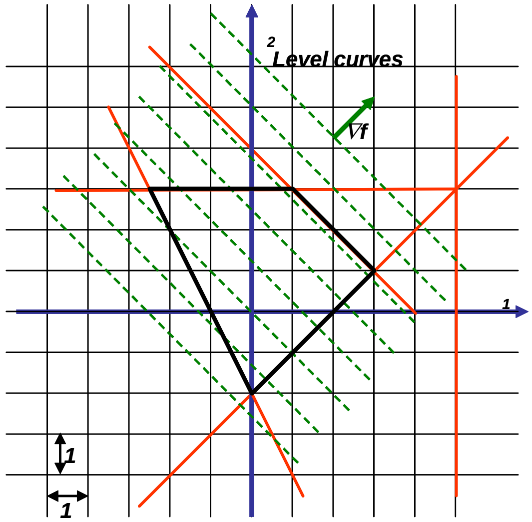

Figure 2: Feasible domain defined from the five inequalities (In order to check which side of the line belongs to Ω, check the origin!). The level curves are parallel to one of the sides in the polygon, so that the solution is not unique. However, $(2, 2)$ is a solution, and the objective value is $2 + 2 = 4$.

New variables and matrices then become

$$\tilde{x} = \begin{bmatrix} y \\ z \\ s \end{bmatrix}, \ \tilde{c} = \begin{bmatrix} c \\ -c \\ 0 \end{bmatrix}, \ \tilde{A} = \begin{bmatrix} A & (-A) & I \end{bmatrix}, \tag{20}$$

and the standard problem is

$$\min_{\tilde{x}} \tilde{c}'\tilde{x}$$
$$\tilde{A}\tilde{x} = b, \ \ \tilde{x} \geq 0. \tag{21}$$

(b) The primal and dual problems have equivalent KKT-equations. The central result is the Duality Theorem which states that if either the dual or the primal has a feasible solution, then both problems have a solution, and the optimal objective values are equal.

(c) Since only the optimal value of the objective is asked for, we turn to the dual problem stated in (b):

$$\max_{\lambda \in \tilde{\Omega}} (\lambda_1 + \lambda_2) \tag{22}$$

when

$$\begin{bmatrix} 1 & -1 \\ 1 & 1 \\ 0 & 1 \\ 1 & 0 \\ -2 & -1 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix} \leq \begin{bmatrix} 2 \\ 4 \\ 3 \\ 5 \\ 2 \end{bmatrix} \tag{23}$$

The easiest is to make a graph similar to the one in Fig. 2. The solution of the dual problem is found along the line $\lambda_2 = 4 - \lambda_1$, between $\lambda_1 = 1$ and $\lambda_1 = 3$, and $\max_\lambda (\lambda_1 + \lambda_2) = 4 = \min_{x \in \Omega} f(x)$.

Problem 4

(a) *What is the idea behind Tikhonov regularization of inverse problems?*

(b) *Prove that the solution of*

$$\min_x \left\{ \|Ax-b\|^2 + \mu \|x\|^2 \right\} \tag{24}$$

is unique and may be expressed in terms of the Singular Value Decomposition of A as

$$x_\mu = \sum_{k=1}^{r=\text{rank}(A)} \frac{\sigma_k}{\sigma_k^2+\mu} \left(u_k' b\right) v_k \tag{25}$$

What happens when $\mu \to 0$ and $\mu \to \infty$?

(Recall that a matrix $A \in \mathbb{R}^{m\times n}$ has a Singular Value Decomposition $A = U\Sigma V'$ where $V = [v_1, \cdots, v_n]$, $U = [u_1, \cdots, u_m]$ and $(\Sigma)_{kk} = \sigma_k \neq 0$ for $k = 1, \cdots, r = \text{rank}\,(A)$. Moreover, $V'V = VV' = I_{n\times n}$, $U'U = UU' = I_{m\times m}$. The Moore-Penrose generalized inverse is defined $A^+c = \sum_{k=1}^r \frac{1}{\sigma_k}\left(u_k' c\right) v_k$)

Solution:

(a) Tikhonov regularization consists of stabilizing the solution of an inverse problem (formulated as an optimization problem) by adding a penalty term to the objective. Typically, it takes the form

$$\min_x \{f(x) + \mu N(x)\}, \tag{26}$$

where μ is a weighting factor and N a regularization function, e.g.,

- $N(x) = \|x\|^2$ for punishing the size of the x.
- $N(x) = \|x - x_0\|^2$ for punishing x when going too far away from what we think is a reasonable solution x_0.
- $N(x) = \|L(x)\|^2$, where L gets large when x becomes irregular, and thus punishing irregularities.

(b) We observe that

$$\begin{aligned} \|Ax-b\|^2 + \mu\|x\|^2 &= (Ax-b)'(Ax-b) + \mu x' I x \\ &= x'\left(A'A + \mu I\right)x - 2\left(A'b\right)'x + \|b\|^2 \end{aligned} \tag{27}$$

Since $A'A \geq 0$, we have $A'A + \mu I > 0$, and the (unique) solution is obtained for $(A'A + \mu I)\,x^* = A'b$. In order to get the form given in the problem, we need the SVD of $A'A + \mu I$:

$$A'A + \mu I = V\Sigma' U' U \Sigma V' + \mu V V' = V\left(\Sigma'\Sigma + I\mu\right)V' = VCV', \tag{28}$$

where $C = \text{diag}\{c_k\}$,

$$c_k = \begin{cases} \sigma_k^2 + \mu, & 1 \leq k \leq r, \\ \mu, & r < k \leq n. \end{cases} \tag{29}$$

The solution may now be written

$$x_\mu = \left(VC^+V'\right)\left(V\Sigma U'\right)b = V\left(C^{-1}\Sigma\right)\left(U'b\right) = \sum_{k=1}^r \frac{\sigma_k}{c_k}\left(u_k' b\right) v_k, \tag{30}$$

which is identical to the formula given in the problem.

Finally,

$$\lim_{\mu \to \infty} x_\mu = 0,$$
$$\lim_{\mu \to 0} x_\mu = A^+ b. \tag{31}$$

Typically, the limit when $\mu \to 0$ is numerically unstable and the reason for the regularization.

Problem 5

When landing, an old-fashioned aircraft brakes with a combination of air-brakes (flaps) and mechanical brakes. Using dimensionless variables throughout, the total braking force may be written $F = -\alpha v^2 - u$, where v is the velocity of the aircraft and u the mechanical braking force. The brakes reduce the velocity from $v = 1$ at the start of the landing field, $x = 0$, to $v = 0$ at $x = 1$. A simple argument involving kinetic energy ($E(x) = v(x)^2$), leads to $\frac{dE(x)}{dx} = -\alpha v^2(x) - u(x)$. Locally, heat builds up in the mechanical brakes, and we shall assume that generated heat per length unit, $\frac{dQ}{ds} \propto u^2$.

The objective is now to find the optimal braking strategy which stops the aircraft at $x = 1$, and at the same time, minimizes the generated heat.

(a) Show that problem may be formulated as

$$\begin{aligned} &\min_u \int_0^1 u^2(x)\,dx, \\ &\frac{d}{dx} v^2(x) = -\alpha v^2(x) - u(x), \\ &v(0) = 1,\ v(1) = 0. \end{aligned} \tag{32}$$

Prove that the dynamics and boundary conditions lead to a linear functional $G(u)$ that forms a constraint for the minimization.

(Hint: The general solution of the differential equation $y'(x) + \alpha y(x) + u(x) = 0,\ y(1) = 0$, is

$$y(x) = \int_x^1 u(s)\, e^{\alpha(s-x)} ds). \tag{33}$$

(b) Determine the optimal breaking force $u^*(x)$ and show that the solution is unique.

Solution:

(a) The total amount of generated heat (assuming nothing is conducted or radiating away) is

$$Q = \int_0^1 u^2(x)\,dx. \tag{34}$$

The dynamics follows from the energy dissipation equation stated in the text,

$$\frac{d}{dx} v^2(x) = -\alpha v^2(x) - u(x), \tag{35}$$

and the boundary conditions are $v(0) = 1$, $v(1) = 0$.

If we now introduce $y(s) = v^2(s)$ as a new variable, the differential equation becomes

$$\begin{aligned} &y'(s) + \alpha y(x) + u(x) = 0, \\ &y(0) = v(0)^2 = 1,\ y(1) = v(1)^2 = 0. \end{aligned} \tag{36}$$

Applying the hint, all solutions vanishing at $x=1$ may be written as

$$y(x)=\int_x^1 u(s)\,e^{\alpha(s-x)}ds, \tag{37}$$

and since we need to fulfill $y(0)=1$, we obtain the constraint

$$1=y(0)=G(u)=\int_0^1 u(s)\,e^{\alpha(s-0)}ds=\int_0^1 u(x)\,e^{\alpha x}dx. \tag{38}$$

(b) The problem then consists of

$$\begin{aligned}&\min_u \int_0^1 u^2(x)\,dx,\\ &\int_0^1 u(x)\,e^{\alpha x}dx=1.\end{aligned} \tag{39}$$

The first functional is strictly convex since the kernel is strongly convex, and since the constraint involves a linear functional, the Lagrangian,

$$\mathcal{L}(u,\lambda)=\int_0^1 u^2(x)\,dx+\lambda\int_0^1 u(x)\,e^{\alpha x}dx, \tag{40}$$

will be strictly convex. Thus, if there is a solution, *it is necessarily unique.* The kernel of $\mathcal{L}$ is

$$u^2(x)+\lambda u(x)\,e^{\alpha x}, \tag{41}$$

and the Euler equation is therefore

$$\frac{d}{dx}\frac{\partial}{\partial u'}\left(u^2+\lambda u e^{\alpha x}\right)-\frac{\partial}{\partial u}\left(u^2+\lambda u e^{\alpha x}\right)=0, \tag{42}$$

or

$$2u+\lambda e^{\alpha t}=0, \tag{43}$$

Thus,

$$u(x)=-\frac{\lambda}{2}e^{\alpha x}. \tag{44}$$

The value of λ is determined from the constraint,

$$1=\int_0^1\left(-\frac{\lambda}{2}e^{\alpha x}\right)e^{\alpha x}dx=\left(-\frac{\lambda}{2}\right)\frac{1}{2\alpha}\left(e^{2\alpha}-1\right), \tag{45}$$

and we obtain

$$u^*(x)=\frac{2\alpha}{e^{2\alpha}-1}e^{\alpha x}. \tag{46}$$

It is also easy to find the velocity as a function of x,

$$\begin{aligned}v^2(x)=y(x)&=\int_x^1 u(s)\,e^{\alpha(s-x)}ds\\ &=\frac{2\alpha}{e^{2\alpha}-1}\int_x^1 e^{\alpha s}ds=\frac{2}{e^{2\alpha}-1}\left(e^{\alpha}-e^{-\alpha x}\right).\end{aligned} \tag{47}$$