

Contact during exam:
Anne Kværnø: 92663824

Exam in TMA4180 Optimization Theory

Wednesday May 29, 2013
Tid: 09.00 – 13.00

Auxiliary materials: Simple calculator (Hewlett Packard HP30S or Citizen SR-270X)
Rottmann: *Matematisk formelsamling*

With preliminary solutions (not proofread).

Problem 1 Let

$$f(\mathbf{x}) = \frac{x^2}{2} + x \cos y.$$

- a) Find the gradient and the hessian of f .
- b) Find all minima of f .

For the remaining part of this problem we discuss one step of a line search method, starting from $\mathbf{x}_0 = (1, \pi/4)$, with search direction $\mathbf{p} = (-1, 0)$.

- c) Confirm that \mathbf{p} is a descent direction from \mathbf{x}_0 .
- d) State the Wolfe-conditions. What are the purpose of these conditions?
If $c_1 = 0.1$ and $c_2 = 0.8$, what are the admissible values for the steplength α .
- e) Do one step of the line search method, using the exact value for the steplength.

Solution:

a)

$$\text{Gradient: } \nabla f(\mathbf{x}) = (x + \cos y, -x \sin y)^T,$$

$$\text{Hessian: } \nabla^2 f(\mathbf{x}) = \begin{pmatrix} 1 & -\sin y \\ -\sin y & -x \cos y \end{pmatrix}$$

b) There are two sets of solutions for the first order condition: $\nabla f(\mathbf{x}) = 0$:

1. The first solution is $x = 0$, $y = (n + 1/2)\pi$ for $n = 0, \pm 1, \pm 2, \dots$. In this case the Hessian is

$$\nabla^2 f(\mathbf{x}) = \begin{pmatrix} 1 & (-1)^{n+1} \\ (-1)^{n+1} & 0 \end{pmatrix}$$

with eigenvalues $\lambda = \frac{1}{2}(1 \pm \sqrt{5})$, so these are not minima (they are saddle points).

2. The second solution is $x = (-1)^{n+1}$, $y = n\pi$, for $n = 0, \pm 1, \pm 2, \dots$. The Hessian in this point is simply the identity matrix, which obviously is SPD, so these are strict local minima. Further

$$f((-1)^{n+1}, n\pi) = \frac{1}{2} + (-1)^{n+1} \cos n\pi = -\frac{1}{2}$$

so we can conclude that these are also global minima.

c) The direction \mathbf{p}_0 is descent if $\nabla f(\mathbf{x}_0)^T \mathbf{p}_0 < 0$. In our case this is

$$(1 + \sqrt{2}/2, -\sqrt{2}/2) \cdot \begin{pmatrix} -1 \\ 0 \end{pmatrix} = -1 - \sqrt{2}/2 < 0$$

so \mathbf{p}_0 is a descent direction.

d) The Wolfe conditions are (general)

$$\begin{aligned} f(\mathbf{x}_k + \alpha_k \mathbf{p}_k) &\leq f(\mathbf{x}_k) + c_1 \alpha_k \nabla f(\mathbf{x}_k)^T \mathbf{p}_k \\ \nabla f(\mathbf{x}_k + \alpha_k \mathbf{p}_k)^T \mathbf{p}_k &\geq c_2 \nabla f(\mathbf{x}_k)^T \mathbf{p}_k. \end{aligned}$$

for two parameters c_1 and c_2 satisfying $0 < c_1 < c_2 < 1$.

Choosing the steplengths such that the Wolfe conditions are satisfied ensures sufficiently decrease of the objective function from one iteration to the next one, to ensure convergence the line search method, see Theorem 3.2 in N&W for details. This is critical when the one-dimensional minimization problem of a line search step is solved approximately.

With the given f , \mathbf{x}_0 and \mathbf{p}_0 , this becomes

$$\begin{aligned} \alpha &\leq (1 - c_1)(2 + \sqrt{2}) \\ \alpha &\geq (1 - c_2)(1 + \sqrt{2})/2 \end{aligned}$$

or, with $c_1 = 0.1$ and $c_2 = 0.8$

$$0.3414 \leq \alpha \leq 3.073.$$

e) One step with a line search method means to solve the following one-dimensional problem

$$\min_{\alpha > 0} f(\mathbf{x}_0 + \alpha \mathbf{p}_0) = \min_{\alpha > 0} \left\{ \frac{1}{2}(1 - \alpha)^2 + \frac{\sqrt{2}}{2}(1 - \alpha) \right\}$$

with solution $\alpha_0 = 1 + \sqrt{2}/2$. The new iterate is then

$$\mathbf{x}_1 = \mathbf{x}_0 + \alpha_0 \mathbf{p}_0 = \left(-\frac{\sqrt{2}}{2}, \frac{\pi}{4} \right)^T.$$

The value of the objective function is then $f(\mathbf{x}_1) = 0$.

Problem 2 Given the linear problem

$$\min x_1 - 3x_2$$

subject to

$$-x_1 + 2x_2 \leq 6$$

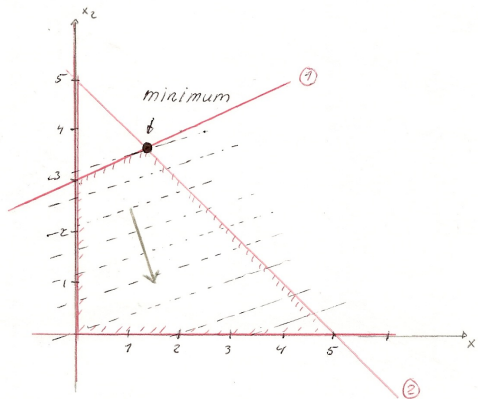
$$x_1 + x_2 \leq 5$$

$$x_1, x_2 \geq 0$$

- Make a sketch of the feasible domain, and solve the problem graphically.
- Write the problem in standard form.
- Write up the dual of this problem.
State the relations between the primal and the dual problem (the duality theorem).
- Explain the idea of the simplex method for linear problems.
Perform one step of the method on the given problem, starting from $(x_1, x_2) = (0, 0)$.

Solution:

a)



The minimum is at the solution of

$$-x_1 + 2x_2 = 6, \quad x_1 + x_2 = 5,$$

that $x_1 = 4/3$ and $x_2 = 11/3$.

b) The standard form of an LP problem is

$$\min c^T x \quad \text{subject to} \quad Ax = b, \quad x > 0.$$

In our case, using slack variables x_3 and x_4 , this is

$$\begin{aligned} & \min x_1 - 3x_2 \\ & \text{subject to} \\ & -x_1 + x_2 + x_3 = 6 \\ & x_1 + x_2 + x_4 = 5 \\ & x_1, x_2, x_3, x_4 \geq 0 \end{aligned}$$

so

$$A = \begin{pmatrix} -1 & 2 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{pmatrix}, \quad b = (6, 5)^T, \quad c^T = (1, -3, 0, 0).$$

c) The dual problem is

$$\max b^T \lambda \quad \text{subject to} \quad A^T \lambda \leq c$$

where λ is the vector of Lagrange multipliers for the equality constraints. In our case, this is:

$$\begin{aligned} & \max 6\lambda_1 + 5\lambda_2 \\ & \text{subject to} \\ & -\lambda_1 + \lambda_2 \leq 1 \\ & 2\lambda_1 + \lambda_2 \leq -3 \\ & \lambda_1 \leq 0 \\ & \lambda_2 \leq 0 \end{aligned}$$

For the relation between the primal and the dual problem, see Theorem 13.1 in the N&W.

d) The rough idea of the simplex method is the following: The feasible domain Ω of a LP problem is a polytope. Start from a vertex (or a basic feasible point) of this polytope. Move from vertex to vertex along edges, reducing the objective function $c^T x$. If a minimum exist, it is on a vertex, and will eventually be found.

How this can be done mathematically is better explained with the example. In our case we start with $x_1 = x_2 = 0$. Since the point has to be feasible, $Ax = b$ has to be satisfied, so $x_3 = 6$ and $x_4 = 5$. This fulfill the definition of a basic feasible point (p. 363 in N&W), with $\mathcal{B} = \{3, 4\}$. Do the corresponding splitting of the system, that is:

$$x = [0, 0, 6, 5]^T = [0^T, x_B^T]^T, \quad A = [N, B], \quad c^T = [c_1^T, c_2^T].$$

From the definition of a feasible point, B has to be invertible. In this case, B is simply the 2×2 identity matrix The idea is now to search for an $x(t) \in$

Ω s.t. $x(0) = [0^T, x_B^T]^T$ and $c^T x(t)$ is reduced as t increases. Let $x(t) = [v(t)^T, y(t)^T]^T$. We get

$$Ax(t) = Nv(t) + By(t) = b \quad y(t) = B^{-1}(b - Nv(t) = x_B - B^{-1}Nv(t).$$

$$c^T x(t) = c_2^T v(t) + c_1^T y(t) = c_1^T x_B + (c_2^T - c_1^T B^{-1}N)v(t)$$

In our case, the vector $c_2^T - c_1^T B^{-1}N = [1, -3]^T$. So, by choosing $v(t) = [0, t]^T$ $c^T x(t)$ will decrease with increasing value of t . But the last two components are $y(t) = [6, 5]^T - [2, 1]^T t$. They should both be non-negative, so we can only increase t up to $t = 3$, in which case $y(3) = [0, 2]^T$, $v(3) = [0, 3]^T$ and the new point is

$$x = [0, 3, 0, 2]^T.$$

Which is what was expected from the figure in point **a**).

Problem 3

- a) Define a (strict) convex function and a convex set.
- b) Let the set Ω be defined by

$$\Omega = \{\mathbf{x} \in \mathbb{R}^n : c_i(\mathbf{x}) \leq 0, i = 1, \dots, m\}$$

Show that Ω is convex if all the functions $c_i, i = 1, 2, \dots, m$ are convex.

Consider the following constrained optimization problem:

$$\min x_1$$

subject to

$$\begin{aligned} x_1 - x_2^2 - x_3^2 &\geq 0 \\ x_2 + x_3 - 1 &\geq 0 \end{aligned}$$

- c) Set up the KKT conditions for this problem.
- d) Find the KKT point(s). What can you say about the optimality of these points (if they exist)?

Solution:

- a) A set Ω is convex if, for all $\mathbf{x}, \mathbf{y} \in \Omega$

$$\theta\mathbf{x} + (1 - \theta)\mathbf{y} \in \Omega, \quad \text{for all } \theta \in (0, 1).$$

A function f defined on a convex set Ω is convex iff for all $\mathbf{x}, \mathbf{y} \in \Omega$

$$f(\theta\mathbf{x} + (1 - \theta)\mathbf{y}) \leq \theta f(\mathbf{x}) + (1 - \theta)f(\mathbf{y}), \quad \text{for all } \theta \in (0, 1).$$

It is strictly convex if the inequality above is strict.

- b) We have:

$$\begin{aligned} c_i(\theta\mathbf{x} + (1 - \theta)\mathbf{y}) &\leq \theta c_i(\mathbf{x}) + (1 - \theta)c_i(\mathbf{y}) && c_i \text{ is convex} \\ &\leq 0 && c_i(\mathbf{x}), c_i(\mathbf{y}) \leq 0, \\ &&& \theta, 1 - \theta > 0. \end{aligned}$$

Which is true for all $i = 1, \dots, n$. So $\theta\mathbf{x} + (1 - \theta)\mathbf{y} \in \Omega$ and we can conclude that Ω is convex.

c) The Lagrangian is

$$\mathcal{L}(\mathbf{x}, \lambda) = x_1 - \lambda_1(x_1 - x_2^2 - x_3^2) - \lambda_2(x_2 + x_3 - 1)$$

and the KKT-conditions are

$$\nabla_x \mathcal{L}(\mathbf{x}, \lambda) = 0 \quad (\text{a})$$

$$\lambda_i c_i(\mathbf{x}) = 0, \quad i = 1, 2 \quad (\text{b})$$

$$\lambda_i \geq 0, \quad i = 1, 2 \quad (\text{c})$$

$$c_i(\mathbf{x}) \geq 0, \quad i = 1, 2 \quad (\text{d})$$

which becomes

$$1 - \lambda_1 = 0, \quad (\text{a}')$$

$$2\lambda_1 x_2 - \lambda_2 = 0$$

$$2\lambda_1 x_3 - \lambda_2 = 0$$

$$\lambda_1(x_1 - x_2^2 - x_3^2) = 0 \quad (\text{b}')$$

$$\lambda_2(x_2 + x_3 - 1) = 0$$

$$\lambda_1, \lambda_2 \geq 0, \quad (\text{c}')$$

$$x_1 - x_2^2 - x_3^2 \geq 0, \quad (\text{d}')$$

$$x_2 + x_3 - 1 \geq 0$$

d) Clearly, $\lambda_1 = 1$, so (c') is satisfied for this multiplier. Further, from (b'), $x_1 = x_2^2 + x_3^2$. If $\lambda_2 = 0$ then from (a') $x_2 = x_3 = 0$ and then $x_1 = 0$. But this solution do not satisfy the last inequality of (d'), so we conclude that $\lambda_2 \neq 0$. The last two equations of (a') together with the last of (b') is

$$2x_2 - \lambda_2 = 0, \quad 2x_3 - \lambda_2 = 0, \quad x_2 + x_3 - 1 = 0,$$

with the solution $\lambda_2 = 1$ and $x_2 = x_3 = 1/2$. We then have the following solution

$$x_1 = x_2 = x_3 = \frac{1}{2}, \quad \lambda_1 = \lambda_2 = 1. \quad (*)$$

Finally, we have to prove that the LICQ-condition holds, that is $\nabla c_i(x)$ are linear independent for all active constraints in this point. In our case, both constraints are active, and

$$\nabla c_1(x) = (1, -2x_2, -2x_3), \quad \nabla c_2(x) = (0, 1, 1)$$

which obviously are linear independent. So (*) is a KKT-point.

Since the objective function x_1 is linear and thus convex, the feasible domain Ω is convex (since c_i , $i = 1, 2$ are both concave), and we have found an extreme point, we know that this is a global minimum. And since there was only one solution of the KKT-conditions, we know that the minimum is unique.

Problem 4 Let

$$F(y) = \int_1^2 \frac{(y'(x))^2}{x} dx$$

be defined on

$$\mathcal{D} = \{y \in C^1[0, 1] ; y(1) = 0, y(2) = 3\}.$$

a) Solve the problem

$$\min_{y \in \mathcal{D}} F(y).$$

b) Solve the problem from a), but now with the additional constraint

$$\int_1^2 y(x) dx = 1.$$

Solution:

a) First of all, notice that the integrand only depends on y' . On \mathcal{D} the function $f(\underline{x}, z) = z^2/\underline{x}$ is strongly convex, since

$$f(\underline{x}, z+w) - f(\underline{x}, z) - f_z(\underline{x}, z)w = \frac{1}{\underline{x}} ((z+w)^2 - z^2 - 2zw) = \frac{w^2}{\underline{x}} \geq 0$$

with equality if and only if $w = 0$.

The Euler-Lagrange equation

$$\frac{d}{dx} f_z[y(x)] = f_y[y(x)]$$

becomes

$$\frac{d}{dx} \frac{2y'(x)}{x} = 0 \quad \Rightarrow \quad \frac{2y'(x)}{x} = C$$

for some constant C . So

$$y'(x) = \frac{Cx}{2} \quad \Rightarrow \quad y(x) = \frac{1}{4}Cx^2 + D$$

Insert the boundary conditions from \mathcal{D} , $y(1) = 0$ and $y(2) = 3$, and the solution becomes

$$y(x) = x^2 - 1$$

which is a minimizer, since F is strictly convex on \mathcal{D}

b) In this case, we want to minimize

$$\tilde{F}(y) = \int_1^2 \frac{y'(x)^2}{x} dx + \lambda \int_1^2 y(x) dx$$

over \mathcal{D} , The Euler-Lagrange equations becomes

$$\frac{d}{dx} \frac{2y'(x)}{x} = \lambda$$

with solution

$$y(x) = \frac{1}{6}\lambda x^2 + \frac{1}{2}Cx^2 + D$$

Insert the boundary conditions gives $C = 2 - 7\lambda/9$ and $D = -1 + 2\lambda/9$, so

$$y(x) = \frac{1}{6}\lambda x^3 + \frac{1}{2}\left(2 - \frac{7}{9}\lambda\right)x^2 - 1 + \frac{2}{9}\lambda$$

Finally, λ is determined from the constraint:

$$\int_1^2 y(x)dx = \frac{4}{3} - \frac{13}{216}\lambda = 1 \quad \Rightarrow \quad \lambda = 72/13$$

and the solution becomes

$$y(x) = \frac{1}{13}(12x^3 - 15x^2 + 3).$$

Since the constraint is linear and thus convex, \tilde{F} is strictly convex, and $y(x)$ is the minimizer of the constrained problem.

Problem 5 Consider the integral functional

$$J(y) = \int_0^1 \left(y(x) + \frac{(y'(x))^2}{2} + y(x) \cdot y''(x) \right) dx$$

for $y \in C^2[0, 1]$.

- a) Find the Gâteaux derivative $\delta J(y; v)$ of J .
- b) Find $y \in C^2[0, 1]$ such that $\delta J(y; v) = 0$ for all $v \in C^2[0, 1]$.
Is this a minimizer of J ? Justify your answer.

Solution:

- a) Use

$$\delta J(y; v) = \frac{\partial}{\partial \varepsilon} J(u + \varepsilon v)|_{\varepsilon=0}.$$

which becomes

$$\begin{aligned} \delta J(y; v) &= \int_0^1 \frac{\partial}{\partial \varepsilon} \left((y + \varepsilon v) + \frac{(y' + \varepsilon v')^2}{2} + (y + \varepsilon v) \cdot (y'' + \varepsilon v'') \right) |_{\varepsilon=0} dx \\ &= \int_0^1 (v + y'v' + y''v + yv'') dx \end{aligned}$$

By using partial integration the Gâteaux derivative can be rewritten:

$$\begin{aligned} \delta J(y; v) &= \int_0^1 v dx - \int_0^1 y''v dx + y'v|_0^1 + \int_0^1 y''v dx + \int_0^1 y''v dx + yv'|_0^1 - y'v|_0^1 \\ &= \int_0^1 (1 + y'')v dx + yv'|_0^1 \end{aligned} \quad (1)$$

- b) From (1) we have that this is true if

$$1 + y'' = 0, \quad \text{and} \quad y(0) = y(1) = 0.$$

which is

$$y(x) = \frac{1}{2}x(1 - x).$$

This is not a minimizer, nor a maximizer. For inst. $y(x) = K$ (some constant) is a function in $C^2[0, 1]$. For this choice of $y(x)$, $J(y) = K$ which can be as small or large we want. So J do not have ha minimizer.