Norwegian University of Science and Technology
Institutt for matematiske fag

Exam supervisor:
Name: Simone Cifani (98830093)

EXAM IN OPTIMIZATION THEORY (TMA4180)

Friday May 22, 2009
Time: 09:00–13:00

Examination aids: C Simple calculator, Formula handbook (Rottmann)

Problem 1 Consider the unconstrained minimization problem

$$\min_{(x,y)\in\mathbb{R}^2} f(x,y),$$

$$f(x,y) = 1 + 2y + x^2 + 2xy + 2y^2.$$

a) Compute the gradient and the Hessian of f for arbitrary $\mathbf{x} = [x,y]^T \in \mathbb{R}^2$ and show that $\mathbf{x}^* = [1,-1]^T$ is the unique global minimum.

Answer:

$$\nabla f(x,y) = [2x+2y, 2+2x+4y], \qquad \nabla^2 f(x,y) = \begin{bmatrix} 2 & 2 \\ 2 & 4 \end{bmatrix}$$

Clearly, the gradient vanishes at $\mathbf{x}^* = [1,-1]^T$ and the Hessian is positive definite everywhere.

b) Suppose that a line search method has been given $\mathbf{x}_0 = [0,-1]^T$ as initial point and that a search direction $\mathbf{p} = [1,1]^T$ has been selected. Verify that $\mathbf{p}$ is a descent direction and determine the next approximation $\mathbf{x}_1 = \mathbf{x}_0 + \alpha\mathbf{p}$ such that

$$\alpha = \arg\min_{\bar{\alpha}>0} f(\mathbf{x}_0 + \bar{\alpha}\mathbf{p}).$$

Answer: We find that $\nabla f(x_0) = [-2,-2]$ so that $\nabla f(x_0)\cdot\mathbf{p} = -4 < 0$ thus $\mathbf{p}$ is a descent direction. We compute

$$f(x_0 + \bar{\alpha}p) = f(\bar{\alpha}, \bar{\alpha}-1) = 5\bar{\alpha}^2 - 4\bar{\alpha} + 1$$

so that the unique minimum is at $\alpha = \frac{2}{5}$ corresponding to $x_1 = [\frac{2}{5}, -\frac{3}{5}]^T$.

c) Suppose now that we want to solve the line search problem approximately. Given constants $0 < c_1 < c_2 < 1$, show that the Wolfe conditions are satisfied for $\mathbf{x}_0$ and $\mathbf{p}$ as in **b)** whenever

$$\frac{2}{5}(1 - c_2) \leq \alpha \leq \frac{4}{5}(1 - c_1).$$

Answer: This is just a matter of checking each of the conditions

1. $f(\mathbf{x}_0 + \alpha\mathbf{p}) \leq f(\mathbf{x}_0) + c_1\alpha\nabla f(\mathbf{x}_0)\mathbf{p}$
2. $\nabla f(\mathbf{x}_0 + \alpha\mathbf{p})\,\mathbf{p} \geq c_2\nabla f(\mathbf{x}_0)\,\mathbf{p}$

substituting the values for $\mathbf{x}_0, \mathbf{p}$.

Problem 2 Consider the constrained minimization problem

$$\min_{x\in\mathbb{R}^2} 4x_1 + x_2, \tag{2.1}$$

subject to

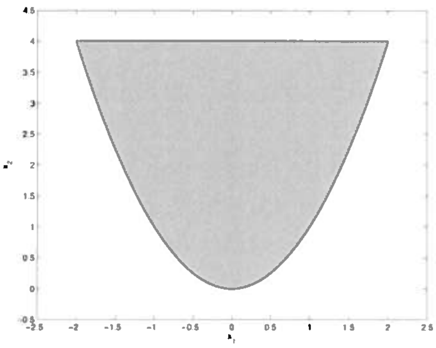
$$c_1(x) = x_2 - x_1^2 \geq 0, \tag{2.2}$$

$$c_2(x) = A - x_2 \geq 0, \quad A > 0. \tag{2.3}$$

a) Sketch the domain Ω defined by the constraints c_1 and c_2 and show that Ω is convex. Is Ω strictly convex? Does the LICQ hold for all points in Ω?

Answer:

In the figure, we have used $A = 4$. It is also clear from the sketch that Ω is convex, however, since it is defined by constraints $x_1^2 - x_2 \leq 0$, $x_2 \leq A$, whose left hand sides are both convex (check Hessian), Corollary 1 of the Basic Tools Note asserts the convexity. Ω is not strictly convex since for instance every point on the straight line segment between $(-2, 4)$ and $(2, 4)$ belongs to Ω. As for the LICQ, we compute the matrix $A(x)$ with rows ∇c_1 and ∇c_2



$$A(x) = \begin{bmatrix} -2x_1 & 1 \\ 0 & -1 \end{bmatrix}$$

So $\text{rank}(A(x)) = 2$ on $\{(x_1, x_2) \in \Omega : x_1 \neq 0$. In other words, the LICQ holds everywhere except on the x_2-axis.

b) Write down the KKT conditions for this problem. Show that if a point x^* is a KKT point with $A \geq 4$, then the corresponding Lagrange multiplier $\lambda_2^* = 0$.
Hint. Assume to the contrary $\lambda_2^* > 0$ and then analyze the KKT conditions to obtain a contradiction.

Answer: KKT conditions

$$\begin{aligned} 4 + 2\lambda_1 x_1 &= 0 \\ 1 - \lambda_1 + \lambda_2 &= 0 \\ \lambda_1(x_2 - x_1^2) &= 0 \\ \lambda_2(A - x_2) &= 0 \\ \lambda_1, \lambda_2 &\geq 0 \end{aligned}$$

Assume $\lambda_2 > 0$. The 2nd equation gives $\lambda_1 > 1$, the last two equations imply $x_1^2 = x_2 = A$. Solving the 1st equation for λ_1 and squaring, leads to $\lambda_1^2 = 4/A > 1$ so that $A < 4$, a contradiction.

c) Suppose now that $A \geq 4$ and determine all KKT points. Have you found a global minimum?

Answer: We can assume $\lambda_2^* = 0$, so that $\lambda_1 = 1$, $x_1 = -2$ and $x_2 = 4$. There is only this one KKT point, and since both Ω and $f(x_1, x_2)$ are convex, we conclude that $(-2, 4)$ is the unique global minimum.

d) Consider now the minimization problem, but only subject to the constraint (2.2). Formulate the logarithmic barrier problem, and determine the solution to the resulting unconstrained minimization problem, x_μ in terms for the barrier parameter μ. We have seen that a function $\lambda_1(\mu)$ can be defined such that $\lambda_1(\mu) \to \lambda_1$ as $\mu \to 0$ where λ_1 is the associated Lagrange multiplier for the constrained problem. Determine $\lambda_1(\mu)$.

Answer: We write down the barrier function

$$Q(x, \mu) = 4x_1 + x_2 - \mu \log(x_2 - x_1^2)$$

Compute $\nabla_x Q(x, \mu) = [4 + \frac{2\mu x_1}{x_2 - x_1^2}, 1 - \frac{\mu}{x_2 - x_1^2}]$ First order conditions give $x_\mu = (-2, 4 + \mu)$, so as $\mu \to 0$ we get $x^* = (-2, 4)$ as before. The function $\lambda_1(\mu)$ is defined as

$$\lambda_1(\mu) = \frac{c_1(x_\mu)}{\mu} = \frac{4 + \mu - (-2)^2}{\mu} \equiv 1.$$

Problem 3 Explain what is meant by the *standard form* of a linear programming problem, and transform the following one to such a form.

$$\min_{x_1,x_2} 3\,x_1 - x_2,$$

$$x_2 \le x_1 + 4,$$
$$x_2 \le 4 - x_1,$$
$$x_2 \ge 0.$$

Answer: By *standard form* we mean a formulation

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

For our specific problem, we introduce non-negative slack variables z_1, z_2 and turn the given inequalities into equalities

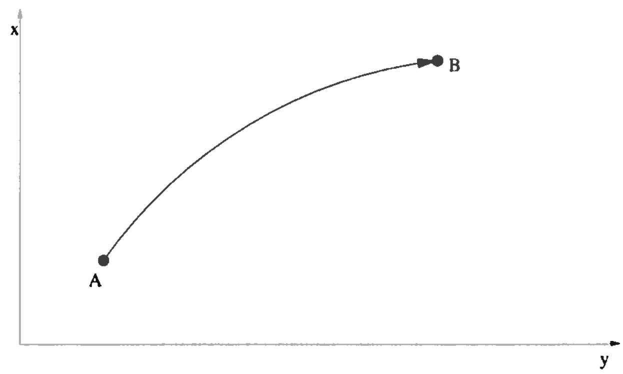
$$-x_1 + x_2 + z_1 = 4, \qquad z_1 \ge 0, x_2 \ge 0$$
$$x_1 + x_2 + z_2 = 4, \qquad z_2 \ge 0$$

Finally, to get only non-negative variables, we need to split x_1 into $x_1 = x_1^+ - x_1^-$ where $x_1^+ = \max(x_1, 0)$, $x_1^- = \max(-x_1, 0)$. Substitute this splitting for every occurence of x_1. We then define the 5-vector $x = (x_1^+, x_1,^-, x_2, z_1, z_2)^T$ and we obtain the standard form, where

$$c = (3, -3, -1, 0, 0)^T,\ A = \begin{pmatrix} -1 & 1 & 1 & 1 & 0 \\ 1 & -1 & 1 & 0 & 1 \end{pmatrix},\ b = (4, 4)^T$$

Problem 4

A cross-country runner wants to move from location A to B (see figure) in a marsh with variable wetness, and is faced with the problem of choosing the fastest path. The speed of which she can run is assumed to depend on the x-coordinate such that her speed at a location (x, y) is $p(x)$ for a positive continuous function $p(x)$. One can therefore work out that along a path $y(x)$ from A to B, the time she spends will be



$$F(y) = \int_{x_A}^{x_B} \frac{\sqrt{y'(x)^2 + 1}}{p(x)}\, dx, \tag{4.1}$$

where x_A and x_B are the x-coordinates of the points A and B respectively. Our main focus in this problem will be

$$\text{Find } y \text{ such that } F(y) = \min_{u \in \mathcal{D}} F(u), \quad \mathcal{D} = \{y \in C^1[x_A, x_B] : y(x_A) = y_A, y(x_B) = y_B\} \quad (4.2)$$

Note that the x-axis is vertical in the figure.

a) Show that the function $f(\underline{x}, y, z) = \dfrac{\sqrt{1+z^2}}{p(\underline{x})}$ is strongly convex on $[x_A, x_B] \times \mathbb{R}^2$. What are the consequences for the functional $F(y)$, and what does it tell us about the existence and uniqueness of the solution to the problem (4.2)?

Answer: We can use Proposition 3.10 in the book of Troutman (p 62). We therefore compute

$$f_{zz} = \frac{(1+z^2)^{-3/2}}{p(\underline{x})} > 0$$

and thus strongly convex on $[x_A, x_B] \times \mathbb{R}^2$. The consequence is that $F(y)$ is strictly convex, and there exists a unique solution to (4.2).

b) Set $p(x) = x$ in the rest of this problem, and show that any function $y \in \mathcal{D}$ satisfying

$$\frac{y'(x)}{\sqrt{1+y'(x)^2}} = \frac{x}{r}, \qquad (4.3)$$

for a constant r, will be a solution to (4.2).

Answer: We can use Theorem 3.7 of Troutman, and conclude that a $y \in \mathcal{D}$ satisfying

$$f_z(x, y'(x)) = \text{const}$$

would be the (unique) solution to our problem. By substituting $p(x) = x$, differentiating f once, and defining const $= 1/r$, the result follows. Note that we have excluded the case that const $= 0$, but this would correspond to an empty x-interval.

c) Verify that functions $y = y(x)$ satisfying

$$x^2 + (y - y_C)^2 = r^2$$

are solutions to (4.3), and determine the constants y_C og r in terms of x_A, y_A, x_B, y_B. Draw a sketch where all these constants are shown.

Answer: To verify the solution, use implicit differentiation to deduce

$$y'(x) = -\frac{x}{y - y_C}, \qquad 1 + y'(x)^2 = \frac{x^2 + (y-y_C)^2}{(y-y_C)^2} = \frac{r^2}{(y-y_C)^2}$$

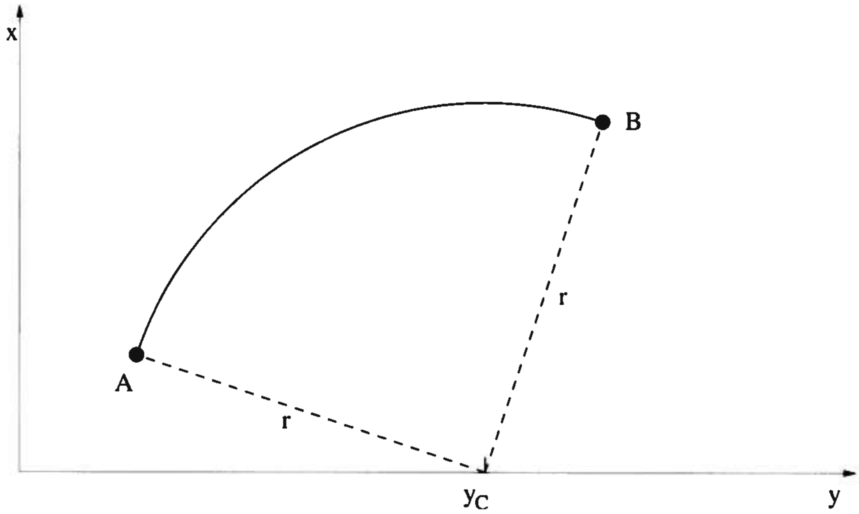
and the answer to the first question follows easily. In order to determine y_C and r, we substitute the boundary points (x_A, y_A) and $(x_B, y_B$ into the solution ansatz and subtract the two, this removes r^2 from the equation and we solve for y_C as

$$y_C = \frac{y_A + y_B}{2} + \frac{x_B - x_A}{y_B - y_A}\frac{x_A + x_B}{2}$$

The expression for r then follows by using one of the two conditions, e.g.

$$r = \sqrt{x_A^2 + (y_A - y_C)^2}$$

Finally, we include a figure, showing that the solution is a circular arc, centered on the y-axis at $(0, y_c)$ with radius r.



Problem 5 In this final problem you shall answer just *yes* or *no* to each question **a)** and **b)**, and in question **c)** you shall just select one of the three alternatives without any further explanation or discussion.

a) In the linear Conjugate Gradient method, the search vectors generated are orthogonal with respect to the standard inner product, i.e. $\langle p_i, p_j \rangle = 0$ whenever $i \neq j$, whereas the residual vectors $r_i = b - Ax_i$ are A-orthogonal, $\langle r_i, r_j \rangle_A = \langle Ar_i, r_j \rangle = 0$ for $i \neq j$, where A is the (SPD) matrix used to define the quadratic form? *Yes* or *No*.

Answer: No

b) In the Trust Region method one needs, in every iteration, to consider a local problem of the form

$$\min_p m(p), \quad m(p) = f + g^T p + \frac{1}{2} p^T B p. \quad \text{subject to } \|p\| \leq \Delta.$$

The question you are to answer is: Can one always obtain a unique global minimum for this problem as a vector p^* which satisfies, for some $\lambda \geq 0$, the conditions

$$\begin{aligned}(B + \lambda I)p^* &= -g \\ \lambda(\Delta - \|p^*\|) &= 0 \\ (B + \lambda I) &\quad \text{is positive semidefinite}\end{aligned}$$

Yes or *No*.

Answer: Yes

c) In an active set method for solving a quadratic programming problem with linear constraints, suppose that one has a working set $\mathcal{W}_0$ for the point x_0. In calculating the next iterate , $x_1 = x_0 + p$ with the reduced problem, one finds that $p = 0$, but concludes that x_0 is not a KKT point for the total problem because some of the Lagrange multipliers are negative. What is the next course of action

A. Find the most negative $\lambda_j \in \mathcal{W}_0$ for x_0 and remove the index j from $\mathcal{W}_0$, i.e. set $\mathcal{W}_1 = \mathcal{W}_0 \backslash \{j\}$ and consider the reduced problem with this new smaller working set.

B. Include a new index among the currently inactive constraints, setting $\mathcal{W}_1 = \mathcal{W}_0 \cup \{j'\}$ where j' is the smallest index not belonging to $\mathcal{W}_0$. Consider the reduced problem with this larger working set.

C Discard all constraints, i.e. set $\mathcal{W}_1 = \emptyset$ and restart the algorithm.

Answer: A