TMA 4180 Optimeringsteori

Midterm Test - Solution

March 4, 2010, 10:15 –12:00.

Problem 1

(a) *Define a convex set and a convex function defined on a convex set.*

(b) *Show that if the function f is convex, then the set*

$$\{x\ ;\ f(x) \leq c\} \tag{1}$$

is convex.

Solution:

(a) A set Ω is convex if for all $x_1, x_2 \in \Omega$, $x_\theta = \theta x_1 + (1-\theta)x_2 \in \Omega$, $0 < \theta < 1$. The function f is convex on Ω if for all $x_1, x_2 \in \Omega$, $f(x_\theta) \leq \theta f(x_1) + (1-\theta) f(x_2)$.

(b) It follows from the definition of a convex function that if $f(x_1),\ f(x_2) \leq c$, then

$$f(x_\theta) \leq \theta f(x_1) + (1-\theta) f(x_2) \leq \theta c + (1-\theta)c = c. \tag{2}$$

Thus, $x_\theta \in \{x\ ;\ f(x) \leq c\}$, and the set is convex.

Problem 2

(a) *State the necessary first and second order conditions, and the second order sufficient condition on a smooth function f at x^*, for x^* to be a local minimum of the unconstrained problem*

$$\min_{x \in \mathbb{R}^n} f(x). \tag{3}$$

(Hint: First order conditions deal with ∇f and second order conditions with $\nabla^2 f$)

Let

$$f(x) = (1-x_1)^2 + \mu(x_2 - x_1)^2,\ x = (x_1, x_2)' \in \mathbb{R}^2, \tag{4}$$

where μ is a small, positive constant, $0 < \mu \ll 1$.

(b) *Show that the function f is strictly convex by inspecting the Hessian.*

(c) *What is A, x^*,κ, and the meaning of $\|\cdot\|_A$ in the error estimate*

$$\|x_j - x^*\|_A \leq \frac{\kappa - 1}{\kappa + 1} \|x_{j-1} - x^*\|_A, \tag{5}$$

for the Steepest Descent method applied the unconstrained problem in (3), where f is the function in Eqn. (4)? Estimate $\frac{\kappa-1}{\kappa+1}$.

(Hints for (b) and (c): The eigenvalues of the matrix $A = \nabla^2 f$ are

$$\lambda_1 = 2\mu + 1 + \sqrt{4\mu^2 + 1}, \tag{6}$$

$$\lambda_2 = 2\mu + 1 - \sqrt{4\mu^2 + 1}. \tag{7}$$

Solution:

(a) The first order necessary condition, identifying *potential minima*, is

$$\nabla f(x^*) = 0. \tag{8}$$

Furthermore, if $\nabla f(x^*) = 0$, it is *necessary* for a local minimum that

$$\nabla^2 f(x^*) \geq 0, \tag{9}$$

and *sufficient* for a strict local minimum if

$$\nabla^2 f(x^*) > 0. \tag{10}$$

(b) A smooth function is strictly convex if the Hessian is positive definite in every point. Here the Hessian is

$$\nabla^2 f = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f}{\partial x_1 \partial x_2} & \frac{\partial^2 f}{\partial x_2^2} \end{bmatrix} = \begin{bmatrix} 2+2\mu & -2\mu \\ -2\mu & 2\mu \end{bmatrix}. \tag{11}$$

This constant matrix is diagonal dominant and therefore positive definite. Alternatively, observe that both eigenvalues given in the hint are positive.

(c) The matrix A is the Hessian $\nabla^2 f$ at the solution Eqn. (11) and $\kappa = \lambda_{\max}/\lambda_{\min}$ is the condition number of A. The unique solution is clearly $x^* = (1,1)$, where $f(x^*) = 0$. The so-called A-norm is defined as

$$\|y\|_A = \sqrt{y' A y}. \tag{12}$$

Here,

$$\frac{\kappa - 1}{\kappa + 1} = \frac{\lambda_1 - \lambda_2}{\lambda_1 + \lambda_2} = \frac{2\sqrt{4\mu^2+1}}{4\mu+2} \approx \frac{2\left(1+2\mu^2\right)}{4\mu+2} = \frac{1+2\mu^2}{1+2\mu} \approx \frac{1}{1+2\mu}, \tag{13}$$

thus defining a slow geometric convergence when $0 < \mu \ll 1$.

Problem 3:

Explain, without proofs, how the search directions are found in the Conjugate Gradient method applied to the quadratic model problem

$$\min_{x\in\mathbb{R}^n} \left\{ \frac{1}{2} x' A x - b' x \right\}, \ A > 0. \tag{14}$$

Solution:

For the CG method the new search direction, p_k, out from the current point x_k is a linear combination of the old search direction, p_{k-1}, and the current gradient direction, $g_k = Ax_k - b$:

$$p_k = -g_k + \beta_k p_{k-1}. \tag{15}$$

(g_k is orthogonal to span $\{p_0, \cdots, p_{k-1}\}$). The parameter β_k is chosen so that $p_k' A p_{k-1} = 0$, which leads to

$$\beta_k = \frac{g_k' A p_{k-1}}{p_{k-1}' A p_{k-1}}. \tag{16}$$

Problem 4:

Solve the Least Square optimization problem

$$\min_{x\in\mathbb{R}^2} \|Ax - b\|^2, \tag{17}$$

where

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ -1 & 0 \end{bmatrix}, \ b = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}. \tag{18}$$

Solution:

The problem is solved by writing

$$q\left(x\right)=\left\|Ax-b\right\|^{2}=\left(Ax-b\right)'\left(Ax-b\right)=x'A'Ax-2\left(A'b\right)'x+\left\|b^{2}\right\|, \tag{19}$$

which is a *quadratic model problem*. Hence, the solution is found from a vanishing gradient, $\nabla q'=2A'Ax-2\left(A'b\right)=0$, or

$$A'Ax-A'b=0, \tag{20}$$

the *Normal Equations*. We have

$$A'A=\begin{bmatrix}2&1\\1&2\end{bmatrix},\ A'b=\begin{bmatrix}2\\3\end{bmatrix}, \tag{21}$$

and then,

$$2x_1+x_2=2, \tag{22}$$
$$x_1+2x_2=3, \tag{23}$$

with the solution $x_1=\frac{1}{3},\ x_2=\frac{4}{3}$.

Problem 5:

Consider the following constrained optimization problem for $(x_1,x_2)\in R^2$,

$$\min_{x\in\Omega}\left\{-4x_1-x_2\right\}, \tag{24}$$

where Ω *is defined in terms of the constraints*

$$0\le x_1\le 2, \tag{25}$$
$$0\le x_2, \tag{26}$$
$$x_2\le 3-x_1. \tag{27}$$

(a) *Reformulate the constraints into four constraints of the form*

$$c_i\left(x\right)\ge 0,\ i=1,\cdots,4, \tag{28}$$

and write down all KKT-equations and inequalities.

(b) *Solve the problem graphically by making a sketch of* Ω.

(c) *Identify the active and inactive constraints and the corresponding Lagrange multipliers at the solution.*

Solution:

(a) The constraints may be written

$$c_1\left(x\right)=x_1\ge 0, \tag{29}$$
$$c_2\left(x\right)=2-x_1\ge 0, \tag{30}$$
$$c_3\left(x\right)=x_2\ge 0, \tag{31}$$
$$c_4\left(x\right)=3-x_1-x_2\ge 0. \tag{32}$$

Hence, the Lagrangian is

$$\mathcal{L}\left(x,\lambda\right)=-4x_1-x_2-\lambda_1x_1-\lambda_2\left(2-x_1\right)-\lambda_3x_2-\lambda_4\left(3-x_1-x_2\right), \tag{33}$$

and $\nabla_x\mathcal{L}\left(x,\lambda\right)=0$ gives the equations

$$-\lambda_1+\lambda_2+\lambda_4=4, \tag{34}$$
$$-\lambda_3+\lambda_4=1, \tag{35}$$

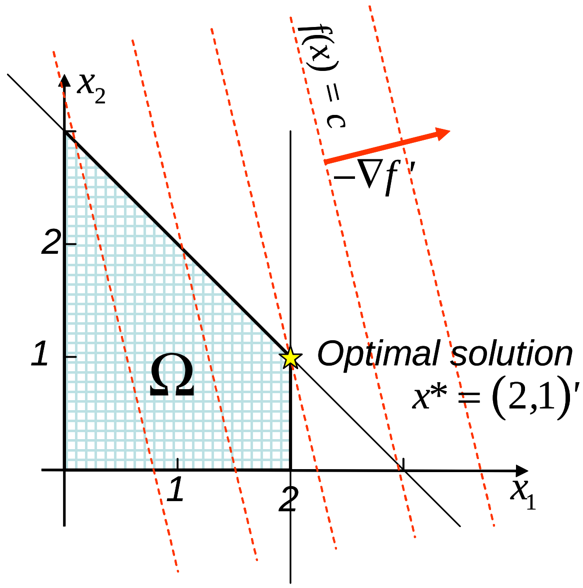

Figure 1: Graph of the level curves of f, the constant negative gradient vector $-\nabla f'$, and Ω.

along with the rest of the KKT equations:

$$\lambda_1 x_1 = 0, \tag{36}$$
$$\lambda_2 (2 - x_1) = 0, \tag{37}$$
$$\lambda_3 x_2 = 0, \tag{38}$$
$$\lambda_4 (3 - x_1 - x_2) = 0, \tag{39}$$

plus all 4 inequalities in Eqn. (29)–(32), and the requirements $\lambda_1, \cdots, \lambda_4 \geq 0$.

(b) The function $f(x)$ has level curves defined by

$$-4x_1 - x_2 = \text{const.}, \tag{40}$$

and the negative gradient direction is therefore constant,

$$-\nabla f' = 4\mathbf{i} + \mathbf{j}. \tag{41}$$

This, along with the constraints in Eqn. (29)–(32) that defines Ω is shown in Fig. 1. The solution is clearly $x^* = (2, 1)'$ with $f(x^*) = -4 \times 2 - 1 = -9$.

(c) Eqns. (36)–(39) give that $\lambda_1 = \lambda_3 = 0$ (c_1 and c_3 are not active), whereas c_2 and c_4 are active, so that λ_2 and λ_4 may be different from 0. It then follows from Eqns. (34) and (35) that

$$\lambda_4 = 1 \text{ and } \lambda_2 = 3.$$

As a final check,

$$\begin{aligned} \nabla f(x^*)' &= \lambda_2 \nabla c_2 (x^*)' + \lambda_4 \nabla c_4 (x^*)' \\ &= 3(-\mathbf{i}) + 1 \times (-\mathbf{i} - \mathbf{j}) = -4\mathbf{i} - \mathbf{j}. \end{aligned} \tag{42}$$