

# Homework 3

CS 730, Semester II, 2007–08

February 21, 2008

1. An elementary but helpful introduction to questions of optimality will be gained by investigating a small problem in standard format:

$$\begin{aligned} &\text{minimize } f_0(x_1, x_2) := 3x_1^2 + 3x_1x_2 + \frac{1}{2}x_2^2 \\ &\text{over all } x \in X := \{(x_1, x_2) : -1 \leq x_1 \leq 1, -1 \leq x_2 \leq 1\} \\ &\text{satisfying } f_1(x_1, x_2) := 2x_1 + x_2 - 1 \leq 0. \end{aligned}$$

- (a) Draw a picture in  $\mathbb{R}^2$  of the feasible set  $C$ .
- (b) Verify that this problem is “well posed” - in the precise sense defined in class.
- (c) Show that at the point  $(0, 0)$ , which lies in the interior of  $C$ , the partial derivatives of  $f_0$  with respect to both  $x_1$  and  $x_2$  vanish, and yet  $f_0$  has neither a local minimum nor a local maximum at  $(0, 0)$ . (Consider the behavior of  $f_0$  along various lines through  $(0, 0)$ , not only the two axes.)
- (d) Determine the optimal value and the optimal solution(s). Also identify the points, if any, that are just locally optimal. (In this two-dimensional situation you can arrive at the answers by applying what you know about one-dimensional minimization. The idea is to study what happens to  $f_0$  along families of parallel lines as they intersect  $C$ , these lines being chosen for instance to line up with one of the edges of  $C$ .)
- (e) Implement and “solve” this model in GAMS.

2. The issues behind the basic theorems on well-posedness will be explored through the properties of the following problem in standard format:

minimize  $f_0(x_1, x_2) := x_1 + x_2^2$  over all  $x \in X := \mathbb{R}^2$  satisfying

$$f_1(x_1, x_2) := -\frac{2x_1}{1+x_1^2} + x_2 \leq 0$$

$$f_2(x_1, x_2) := -\frac{2x_1}{1+x_1^2} - x_2 \leq 0$$

- (a) Draw a picture of the feasible solution set  $C$ . Determine what the  $\epsilon$ -feasible set  $C_\epsilon := \{x \in \mathbb{R}^2: f_1(x) \leq \epsilon, f_2(x) \leq \epsilon\}$  will look like for small  $\epsilon > 0$ . Caution: there is a “phantom” portion of  $C_\epsilon$  which has no counterpart in  $C$ . (Hint: write the  $f_1$  constraint in the form  $x_2 \leq g(x_1)$ ; use calculus to graph  $g$  and understand its asymptotic properties.)
- (b) Verify that the optimal value in this problem is 0 and the unique optimal solution is  $\bar{x} = (0, 0)$ . Establish moreover that every optimizing sequence  $\{x^\nu\}$  must converge to  $\bar{x}$ . (Be careful to keep to the precise *definitions* of optimal value, optimal solution, and optimizing sequence.)
- (c) Show that this problem *isn't* well posed. Produce in fact an *asymptotically feasible* sequence of points  $x^\nu$  with the property that  $f_0(x^\nu) \rightarrow -\infty$ ; this will then be an *asymptotically optimal* sequence that doesn't converge to the unique optimal solution  $\bar{x}$  to the problem. (Verify that your sequence does fit these terms as *technically defined*.)
- (d) Show that the addition of the constraint  $f_3(x_1, x_2) := -x_1 - 1 \leq 0$  would cause the problem to be well posed after all. What could be said then about the convergence of asymptotically optimizing sequences?
- (e) Implement and “solve” this model in GAMS.
3. This exercise concerns various properties of partial derivatives and convexity in unconstrained optimization.
- (a) For the function  $f(x) = f(x_1, x_2) := x_1^3 + x_2^3 - (x_1 + x_2)^2$  on  $\mathbb{R}^2$ , obtain formulas for  $\nabla f(x)$  and  $\nabla^2 f(x)$  and say what you can about where  $f$  might have a local or global minimum or maximum (unconstrained).
- (b) Figure out what you can about the convexity or strict convexity of the function  $g(x) = g(x_1, x_2) := 3 + x_2 + \frac{1}{2}(x_1 - x_2)^2 + \frac{1}{4}x_1^4$  on  $\mathbb{R}^2$ . Does this function attain a local or global minimum on  $\mathbb{R}^2$ ?
- (c) Show that if  $h(x) = \max\{h_1(x), h_2(x)\}$  for convex functions  $h_i$  on  $\mathbb{R}^n$ , then  $h$  is convex as well. Moreover, if both  $h_1$  and  $h_2$  are strictly convex, then  $h$  is strictly convex. (Argue right from the definitions.)
4. The purpose of this exercise is to get you started in thinking about the systematic expression of optimality conditions by focusing on the problem of minimizing a differentiable function  $f_0$  over the set  $C = \mathbb{R}_+^n$ , which is the “orthant” consisting of all  $x = (x_1, \dots, x_n)$  such that  $x_j \geq 0$  for  $j = 1, \dots, n$ .

- (a) Show that if  $\bar{x} = (\bar{x}_1, \dots, \bar{x}_n)$  is a locally optimal solution to this problem, then the following must be satisfied at  $\bar{x}$  (and thus is a *necessary* condition for local optimality):

$$\begin{aligned} (\partial f_0 / \partial x_j)(\bar{x}) &= 0 \text{ for each index } j \text{ with } \bar{x}_j > 0, \text{ whereas} \\ (\partial f_0 / \partial x_j)(\bar{x}) &\geq 0 \text{ for each index } j \text{ with } \bar{x}_j = 0. \end{aligned} \quad (1)$$

(Get this by reduction to elementary calculus; fix all but one of the  $x_j$ 's and let the remaining one vary. It might help if you first think about the cases of the given minimization problem in which  $n = 1$ , or  $n = 2$ .) Verify that (1) can be expressed by

$$\bar{x}_j \geq 0, (\partial f_0 / \partial x_j)(\bar{x}) \geq 0, \bar{x}_j \cdot (\partial f_0 / \partial x_j)(\bar{x}) = 0 \text{ for } j = 1, \dots, n. \quad (1')$$

The literature calls this a problem a *complementarity* problem. It can be generalized further as we shall see later in the course.

- (b) In terms of the gradient vector  $\nabla f_0(x) = ((\partial f_0 / \partial x_1)(x), \dots, (\partial f_0 / \partial x_n)(x))$  and the notation  $y \cdot z = y_1 z_1 + \dots + y_n z_n$  for the scalar product of two vectors  $y = (y_1, \dots, y_n)$  and  $z = (z_1, \dots, z_n)$ , show that condition (1) is *equivalent* to having, for  $C = \mathbb{R}_+^n$ ,

$$\nabla f_0(\bar{x}) \cdot [x - \bar{x}] \geq 0 \text{ for all } x \in C. \quad (2)$$

(Don't forget that "equivalence" refers to two implications, one in each direction.)

- (c) Show that when  $f_0$  is an *convex* function, condition (2) is in fact also *sufficient* for the *global* optimality of a point  $\bar{x} \in C$ : it guarantees that  $f_0(x) \geq f_0(\bar{x})$  for all  $x \in C$ .