Optimization Prof. Carlson

Homework 8

From the text, end of chapter 2, #2, 3, 10, 12, 21

From the text, end of chapter 2

2. Show that each of the following functions is convex or strictly convex on the specified convex set D.

$$(a) \quad f(x_1, x_2) = 5x_1^2 + 2x_1x_2 + x_2^2 - x_1 + 2x_2 + 3, \quad D = R^2.$$

The Hessian is

$$\begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f}{\partial x_1 \partial x_2} & \frac{\partial^2 f}{\partial x_2^2} \end{pmatrix} = \begin{pmatrix} 10 & 2 \\ 2 & 2 \end{pmatrix}.$$

Since the principle minors are positive the Hessian is positive and the function is strictly convex.

$$(b) \quad f(x_1, x_2) = x_1^2/2 + 3x_2^2/2 + \sqrt{3}x_1x_2, \quad D = R^2.$$

The Hessian is

$$\begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f}{\partial x_1 \partial x_2} & \frac{\partial^2 f}{\partial x_2^2} \end{pmatrix} = \begin{pmatrix} 1 & \sqrt{3} \\ \sqrt{3} & 3 \end{pmatrix}.$$

The Hessian has characteristic polynomial

$$\lambda^2 - 4\lambda,$$

so the eigenvalues are 0 and 4. Since the Hessian is positive semidefinite the function is convex.

$$(c) \quad f(x_1, x_2) = (x_1 + 2x_2 + 1)^8 - \log(x_1x_2)^2, \quad D = \{x_1 > 1, x_2 > 1\}.$$

The functions x_1+2x_2+1 is convex (linear) and x^8 is convex and increasing for $x > 0$. The function

$$g = -\log(x_1x_2) = \log\frac{1}{x_1x_2}$$

has Hessian

$$Hg = \begin{pmatrix} 1/x_1^2 & 0 \\ 0 & 1/x_2^2 \end{pmatrix}.$$

The Hessian is positive definite in D, and x^2 is is convex and increasing for $x > 0$. Then the sum of convex functions is convex. Since the second function is strictly convex, so is the sum f.

$$(d) \quad f(x_1, x_2) = 4\exp(3x_1 - x_2) + 5\exp(x_1^2 + x_2^2), \quad D = R^2.$$

The arguments of the exponential functions are convex; both are simple Hessian calculations. The functions $4e^x$ and $5e^x$ are increasing convex functions, so the exponential terms are convex. Finally, the sum of convex functions is convex.

$$(e) \quad f(x_1, x_2) = c_1x_1 + c_2/x_1 + c_3x_2 + c_4/x_2, \quad D = \{x_1 > 0, x_2 > 0\}.$$

Each term is individually convex as a function of one variable, as a computation of the second derivative shows. Then the sum of convex functions is convex.

3. A **quadratic** *function in n variables is any function that can be written*

$$f(X) = a + B \bullet X + X \bullet AX.$$

Here A is real symmetric.

—it a) Show that

$$f(x_1, x_2) = (x_1 - x_2)^2 + (x_1 + 2x_2 + 1)^2 - 8x_1x_2$$

is quadratic.

Expand to get

$$f(x_1, x_2) = 2x_1^2 - 6x_1x_2 + 5x_2^2 + 2(x_1 + 2x_2) + 1.$$

Take

$$a = 1, \quad B = \begin{pmatrix} 2 \\ 4 \end{pmatrix}, \quad A = \begin{pmatrix} 2 & -3 \\ -3 & 5 \end{pmatrix}.$$

b) In general if f is quadratic then

$$\nabla f = B + 2AX, \quad Hf(X) = 2A.$$

c) Show that a quadratic function is (strictly) convex if and only if A is positive (definite) semidefinite.

That the function is (strictly) convex if A is positive (definite) semidefinite follows immediately from Thm 2.3.7. For the other direction, note that $f(Y)$ is linear (resp. concave) on the span of Y if Y is in the null space (resp. negative eigenspace).

d) If f is quadratic and A is positive definite, show that $0 = 2AX + B$ has a unique solution, and this is the strict global minimizer of f.

First f is strictly convex. The solutions of $0 = 2AX + B$ are the critical points, which by 2.3.6 are global minimizers. This exists since A is invertible, and is unique by 2.3.4.

10. Suppose $f(X)$ is a C^1 convex function defined on C. Prove that $X^* \in C$ is a global minimizer of f if and only if

$$\nabla f(X^*) \bullet (X - X^*) \geq 0$$

for all $X \in C$.

Since f is convex we have

$$(A) \quad f(X^*) + \nabla f(X^*) \bullet (X - X^*) \leq f(X)$$

by 2.3.5.

If X^* is a global minimizer then $f(X) - f(X^*) \geq 0$, and so

$$\nabla f(X^*) \bullet (X - X^*) \geq 0$$

for all $X \in C$.

Suppose that X^* is not a global minimizer. Then there is some X such that $f(X) < f(X^*)$. By (A) we get

$$\nabla f(X^*) \bullet (X - X^*) \leq f(X) - f(X^*) < 0,$$

so it is not the case that for all $X \in C$

$$\nabla f(X^*) \bullet (X - X^*) \geq 0.$$

12. Suppose that f has continuous second partials and there is some Z_0 where $Hf(Z_0)$ is not positive semidefinite. Show that f is not convex.

There is an eigenvector Y for $Hf(Z_0)$ such that $Y \bullet Hf(Z_0)Y < 0$. By continuity this inequality holds for all Z in a neighborhood of Z_0. By Taylor's theorem there is some W near Z_0 such that if $Z = Z_0 + \epsilon Y$, for $\epsilon > 0$ and sufficiently small, then

$$f(Z) = f(Z_0) + \nabla f(Z_0) \bullet (\epsilon Y) + \frac{1}{2}(\epsilon Y) \bullet Hf(W)(\epsilon Y),$$

so the condition

$$f(Z) \geq f(Z_0) + \nabla f(Z_0) \bullet (Z - Z_0)$$

fails.