MATH 4420/5420 Optimization Test 2 Professor Carlson

SHOW ALL YOUR WORK

1. a) Define what it means for a set $\Omega \subset R^N$ *to be convex.*

A set Ω is convex if the points

$$tX + (1-t)Y, \quad 0 \le t \le 1$$

are in Ω whenever X and Y are in Ω.

b) If Ω *is a convex set, define what it means for a function* $f : \Omega \to R$ *to be convex. Define strictly convex.*

The function $f : \Omega \to R$ is convex if

$$f(tX + (1-t)Y) \le tf(X) + (1-t)f(Y),$$

for all $X, Y \in \Omega$. The function f is strictly convex if

$$f(tX + (1-t)Y) < tf(X) + (1-t)f(Y)$$

for $0 < t < 1$ and $X \ne Y$.

c) Show that the sum of two convex functions is convex.

Suppose f and g are convex. Then for $X, Y \in \Omega$ and $0 \le t \le 1$,

$$(f+g)(tX + (1-t)Y) = f(tX + (1-t)Y) + g(tX + (1-t)Y)$$

$$\le tf(X) + (1-t)f(Y) + tg(X) + (1-t)g(Y) = t(f+g)(X) + (1-t)(f+g)(Y).$$

2. a) Find the vector V *in* R^3 *of the form*

$$\alpha(1,1,1) + \beta(-2,0,2)$$

which is closest to $(2,2,3)$*. Measure distance in the usual way*

$$d(X,Y)^2 = \|X - Y\|^2 = (X - Y) \bullet (X - Y).$$

Notice that the problem amounts to finding the vector in a plane which is closest to a given point. There will be a minimizer, and the minimizer will be a critical point of the square of the distance.

The square of the distance from V of the given form to $(2,2,3)$ is

$$f(\alpha, \beta) = (\alpha - 2\beta - 2, \alpha - 2, \alpha + 2\beta - 3) \bullet (\alpha - 2\beta - 2, \alpha - 2, \alpha + 2\beta - 3)$$

$$= (\alpha - 2\beta - 2)^2 + (\alpha - 2)^2 + (\alpha + 2\beta - 3)^2.$$

Look for critical points

$$\frac{\partial f}{\partial \alpha} = 2(\alpha - 2\beta - 2) + 2(\alpha - 2) + 2(\alpha + 2\beta - 3) = 0.$$

$$\frac{\partial f}{\partial \beta} = -4(\alpha - 2\beta - 2) + 4(\alpha + 2\beta - 3) = 0.$$

The equations can be solved easily after simplification,

$$\frac{\partial f}{\partial \alpha} = \alpha = 7/3.$$

$$\frac{\partial f}{\partial \beta} = \beta = 1/4.$$

Thus the vector we want is

$$V = \frac{7}{3}(1,1,1) + \frac{1}{4}(-2,0,2).$$

3. Recall the following theorem.

If Ω is a convex open set and $f : \Omega \to R$ has first partial derivatives, then f is convex if and only if

$$f(X) + \nabla f(X) \bullet (Y - X) \leq f(Y), \quad X, Y \in \Omega.$$

a) Use this result to show that any critical point of such a function f is a global minimizer.

A critical point is a point where $\nabla f = 0$. If $\nabla f(X) = 0$ then $f(X) \leq f(Y)$ for all $Y \in \Omega$, so X is a global minimizer.

b) Suppose that $f : \Omega \to R$ has continuous second partial derivatives. What is the Hessian H_f of f? What condition on the Hessian implies that f is convex?

The Hessian is the matrix of mixed second partial derivatives. The function f is convex if the Hessian is a positive semidefinite matrix at each point of Ω.

c) Compute the Hessian of the function

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

Is the Hessian positive definite?

$$Hf(x_1, x_2) = \begin{pmatrix} \partial^2 f/\partial x_1^2 & \partial^2 f/\partial x_1 \partial x_2 \\ \partial^2 f/\partial x_1 \partial x_2 & \partial^2 f/\partial x_2^2 \end{pmatrix} = \begin{pmatrix} 4 & -2 \\ -2 & 2 \end{pmatrix}.$$

The principle minors are 4 and $8 - 4 = 4$, which are both positive, so Hf is everywhere positive definite.

4. Suppose we want to find the local minimum of a real-valued function $f(x)$ of a real variable with as many derivatives as you need. Assume the function is expected to have a local minimum near x_0.

(a) Give a brief description of the bisection method for finding the local minimum.

At a minimum x_1 a differentiable function must have derivative $f'(x_1) = 0$. Let $g(x) = f'(x)$. The bisection method requires g to be continuous, and to change sign on an interval $[a, b]$ containing x_1. The method generates a sequence of approximations to the root x_1 of g as follows. Let $a_0 = a$, $b_0 = b$. For $n = 0, 1, 2, \ldots$ define $c_n = (a_n + b_n)/2$. If $g(c_n) = 0$, we're done. If not, and $f(a_n)f(c_n) < 0$, then define $a_{n+1} = a_n$, $b_{n+1} = c_n$. If $f(a_n)f(c_n) > 0$, then define $a_{n+1} = c_n$, $b_{n+1} = b_n$. At each step the length of the interval is halved, and each interval $[a_n, b_n]$ contains a root of g by the Intermediate Value Theorem. The desired root is $x_1 = \lim_{n\to\infty} a_n$.

(b) Give a brief description of Newton's method for finding the local minimum.

Again, Newton's method uses the fact that the desired minimum will satisfy $f'(x_1) = 0$. Let $g(x) = f'(x)$ and assume that g does not vanish at the root x_1. For purposes of analysis we usually assume that $g'(x)$ is continuous.

Starting at $t_0 = x_0$, define a sequence $\{t_n\}$ by

$$t_{n+1} = t_n - \frac{g(t_n)}{g'(t_n)}.$$

If t_0 is sufficiently close to x_1, the sequence $\{t_n\}$ will converge to the root x_1.

(c) Suppose we want to find the minimum of a real-valued function $f(X)$ of a several real variables, $X = (x_1, \ldots, x_N)$, with as many derivatives as you need. Briefly describe the steepest descent method for finding the minimum.

The steepest descent method also begins with the selection of an initial guess X_0. Since the function f decreases most rapidly near X_0 in the direction of the gradient, the following functions of one variable are introduced,

$$g(t) = f(X + t\nabla f(X)).$$

Beginning with $X = X_0$ one tries to minimize $g(t)$, for instance by looking a critical points. The selected minimizer of g becomes X_1, and the process repeats, thus defining a sequence X_n of vectors which reduce the value of f.