Math 438 **ACTIVITY 19:** Optimization of Functions of Several Variables

Why

Many optimization applications require working with non-linear functions. This activity is designed to review the theory and techniques for finding stationary points for functions of several variables and generalize techniques for deciding when the stationary points are actually optimal points. This involves introduction of the concepts of Hessian matrix, Quadratic Form, and definiteness which will be important during the rest of the course.

LEARNING OBJECTIVES

1. Discover how to compute the Hessian matrix.
2. Discover how to find stationary points for functions of several variables.
3. Learn some criteria for determining if a stationary point gives a local optimal value.
4. Learn to recognize positive/negative definite and postive/negative semidefinite matrices (and why we care)

CRITERIA

1. Success in completing the exercises.
2. Success in involving all members of the team in the solution
3. Understanding of the use of the gradient and Hessian in seeking local optima of multivariable functions.

RESOURCES

1. Sections 4.1 - 4.4 (especially 4.4) of Strategic Mathematics
2. Your recall of max/min situations and derivative concepts in Calculus I, II, III
3. 50 minutes

PLAN

1. Before class: read the text and examples
2. Before class: read the model
3. In class: Complete the exercise.

VOCABULARY

Stationary Point

The point $\mathbf{X}$ is a *stationary point* for the function $f : \mathbf{R_n} \to \mathbf{R}$ if all partial derivatives of f are zero at $\mathbf{X}$.

Hessian Matrix

The *Hessian matrix* $\mathbf{H}$ associated with a function $f : \mathbf{R_n} \to \mathbf{R}$ is the $n \times n$ matrix of second-order partial derivatives of f. If f is twice continuously differentiable, then $\mathbf{H}$ is symmetric.

Quadratic form

A function $f : \mathbf{R_n} \to \mathbf{R}$ is a *quadratic form* if there is a symmetric $n \times n$ matrix $\mathbf{a}$ for which $f(\mathbf{X}) = \mathbf{XAX^T}$.

Definiteness

If $\mathbf{A}$ is a symmetric $n \times n$ matrix, then $\mathbf{A}$ is $\left\{\begin{array}{c}\text{positive definite}\\ \text{positive semidefinite}\\ \text{negative definite}\\ \text{negative semidefinite}\end{array}\right\}$ if $\mathbf{XAX^T} \left\{\begin{array}{c} > 0\\ \geq 0\\ < 0\\ \leq 0\end{array}\right\}$ for every vector$\mathbf{X} \in \mathbf{R_n}$ (that is, in every direction). Otherwise $\mathbf{A}$ is *indefinite*.

Optimal point

A function $f : \mathbf{R_n} \to \mathbf{R}$ has a *local (relative) minimum*(respectively: *maximium*) at a point $\mathbf{X_0}$ if there is a neighborhood of $\mathbf{X_0}$ for which $f(\mathbf{X}) \leq f(\mathbf{X_0})$ (respespectively: $f(\mathbf{X}) \geq f(\mathbf{X_0})$ for every point $\mathbf{X}$ in the neighborhood. The terms *relative optimum* and *relative extremum* are generic for "relative maximum or relative minimum". The point $\mathbf{X_0}$ is then an *optimal point* and $f(\mathbf{X_0})$ is an *optimal value* (or maximum,or minimum, as appropriate).

Second derivative test for relative optimum (Theorem - several variables version)

A function $f : \mathbf{R_n} \to \mathbf{R}$ has a relative minimum (respectively: maximum) at a stationary point $\mathbf{X_0}$ if $\mathbf{H}(\mathbf{X_0})$ is positive definite (Respectively: negative definite).
This is the natural extension of the one-variable second-derivative test, and shares the "not always conclusive" (sufficient but not necessary condition) weakness of that test.

MODEL

A manufacturing company operates its two available machines to polish metal products. The two machines are equally efficient , but their maintenance costs are different. The daily maintenance and operating costs are given (in dollars) by $f(x_1, x_2) = 60 - x_1 - 1.2x_2 + 0.05x_1^2 + 0.1x_1x_2 + 0.07x_2^2$, where x_i is the number off hours of operation of machine i. We would like to find the number of hours to operate each machine in order to minimize daily cost. (At this stage of our development, we put no constraints on the number of hours or the combination).

First, we find the gradient vector and set it equal to 0 to find the stationary points.

$\nabla f(x_1, x_2) = (-1 + 0.1x_1 + 0.1x_2, -1.2 + 0.1x_1 + 0.14x_2)$
Solving the system $\begin{matrix} 0.1x_1 + 0.1x_2 = 1 \\ 0.1x_1 + .014x_2 = 1.2 \end{matrix}$ we find the only stationary point is at $x_1 = 5, x_2 = 5$

Second, we check that $(5, 5)$ gives a relative minimum; we calculate the Hessian and verify that it is positive definite at $(5, 5)$ (in fact,it's positive definite everywhere—but we only care at $(5, 5)$).

$\mathbf{H} = \begin{bmatrix} 0.1 & 0.1 \\ 0.1 & 0.14 \end{bmatrix}$
To verify the definiteness, we use theorem 4.4.5 on p.162. To reduce $\mathbf{H}$ to row-echelon form we first divide h_{11} by 0.1 (so $d_{11} = 0.1$) and subtract—we get $\begin{bmatrix} 1 & 1 \\ 0 & 0.04 \end{bmatrix}$. We now divide h_{22} by 0.04 (so $d_{22} = 0.04$); the diagonal matrix is $\mathbf{D} = \begin{bmatrix} 0.1 & 0 \\ 0 & 0.04 \end{bmatrix}$. A diagonal matrix with only positive entries on the diagonal is positive definite, since for such a matrix $\mathbf{M}$ and non-zero vector $\mathbf{h}$, $\mathbf{hMh^T} = m_{11}h_1^2 + m_{22}h_2^2 + \cdots + m_{nn}h_n^2$, which is a sum of nonnegative numbers, some of which must be positive. (For our example, $\begin{bmatrix} h_1 & h_2 \end{bmatrix} \begin{bmatrix} 0.1 & 0 \\ 0 & 0.04 \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \end{bmatrix} = 0.1h_1^2 + 0.04h_2^2$ with at least one of h_1, h_2 not equal to 0.)

EXERCISE

1. In the model, what is the minimum daily cost that the company can realize?
2. We used a diagonal matrix to determine whether the Hessian matrix was positive definite. What are the conditions for positive or negative definiteness for a diagonal matrix?
3. If the Hessian matrix is indefinite at a stationary point, does that mean that the point gives neither a maximum nor a minimum? [Recall the Calc III second-derivative test]
4. If the Hessian matrix is singular (cannot be reduced to an identity) at a stationary point, how can we determine its definiteness?
5. Verify that a 2×2 symmetric matrix $\begin{bmatrix} a & b \\ b & c \end{bmatrix}$ is negative definite if $a < 0$ and $ac - b^2 > 0$ [Text includes–p.156–a proof that such a matrix is positive definite if $a > 0$ and $ac - b^2 > 0$] (For the Hessian matrix, this corresponds to one case (local max) of the Calc III second derivative test: $f_{xx}f_{yy} - f_{xy}^2 > 0$ and $f_{xx} < 0$).
6. Find and classify (max, min, saddle point, don't know) the stationary points for $f(x_1, x_2) = x_1^3 + x_1^2 - 6x_2^2 + x_2 - 1$