An argument for the second-derivative test

We turn our attention now to developing a test to distinguish among local minima, local maxima, and saddle points for critical inputs. The test involves the second partial derivatives of the function $f : \mathbb{R}^2 \to \mathbb{R}$. We assume that f has continuous second partial derivatives, so that the mixed partials are equal.

Let (x_0, y_0) be a critical input for f. By definition, the first partial derivatives of f are zero for input (x_0, y_0). Using this fact, the second-order Taylor polynomial approximation of f based at (x_0, y_0) gives

$$f(x,y) - f(x_0,y_0) \approx \frac{1}{2}\left[f_{xx}(x_0,y_0)(x-x_0)^2 + 2f_{xy}(x_0,y_0)(x-x_0)(y-y_0) + f_{yy}(x_0,y_0)(y-y_0)^2\right]. \quad (1)$$

We want to determine the conditions under which the expression on the right side is

- positive for all (x, y) near (x_0, y_0), in which case (x_0, y_0) is a local maximizer;
- negative for all (x, y) near (x_0, y_0), in which case (x_0, y_0) is a local minimizer; or
- positive for some (x, y) near (x_0, y_0) and negative for other (x, y) near (x_0, y_0), in which case (x_0, y_0) is a local "saddlizer."

To simplify the expression on the right side of Equation (1), we introduce the notation

$$A = f_{xx}(x_0, y_0), \qquad B = f_{xy}(x_0, y_0)\ \text{, and} \qquad C = f_{yy}(x_0, y_0)$$

These are constants for a given function and a given critical input. We also let

$$u = x - x_0 \qquad \text{and} \qquad v = y - y_0.$$

We focus on $(u, v) = (0, 0)$, because this input in u and v corresponds to the critical input $(x, y) = (x_0, y_0)$. We drop the factor of $1/2$ because this will not affect the sign of the expression. With this notation, the expression on the right side in Equation (1) gives the outputs of a function

$$g(u, v) = Au^2 + 2Buv + Cv^2.$$

This is a quadratic function in the input variables u and v.

To understand the shape of the graph of g near $(0, 0)$, we examine slices with vertical planes containing the $z = g(u, v)$ axis. These planes are defined by lines throught the origin in the uv-plane. Along the line given by $v = mu$, the outputs of g are

$$g(u, mu) = Au^2 + 2Bu(mu) + C(mu)^2 = (A + 2Bm + Cm^2)u^2.$$

From this last expression, we deduce that the concavity of the curve in the section defined by $v = mu$ is determined by the sign of

$$A + 2Bm + Cm^2.$$

The sign clearly depends on the value of m, that is, on which section we have. If the concavity is positive for every value of m, then the critical input at $(u, v) = (0, 0)$ must be a local minimizer. If the concavity is negative for every value of m, then the critical input must be a local maximizer. If the concavity is negative in some sections and positive in others, then the critical input corresponds to a saddle.

Note that the line $u = 0$ is not included among the lines $v = mu$. The line $u = 0$ is the v cross section. The concavity in this section is given by C. Note this is consistent with the fact that the $v = 0$ section has concavity given by A.

To distinguish among the possible cases, we look at the condition under which the concavity is zero. Let m_0 be a value for which

$$A + 2Bm_0 + Cm_0^2 = 0.$$

The case $C = 0$ is a special situation, which you should analyze on your own. Assuming that C is not equal to zero, we use the quadratic formula to solve for m_0 giving

$$m_0 = \frac{-2B \pm \sqrt{4B^2 - 4AC}}{2C} = \frac{-B \pm \sqrt{B^2 - AC}}{C}. \tag{2}$$

There are two possibilities, depending on the sign of the quantity $B^2 - AC$. If $B^2 - AC > 0$, then there are two solutions for m_0, corresponding to two sections in which the concavity is zero. These lines divide the uv-plane into four "wedges," and the concavity alternates from positive to negative in these wedges (see Figure 1). The critical input $(0, 0)$ thus corresponds to a saddle. If $B^2 - AC < 0$, then there are no (real-valued) solutions. Hence, the concavity is positive in all sections or negative in all sections. Note that in order for this case to occur, A and C must have the same sign. If the signs are both positive, then all sections are concave up, meaning $(0, 0)$ is a local minimizer. If the signs are both negative, then all sections are concave down, meaning $(0, 0)$ is a local maximizer.

We summarize the conclusions of the preceding argument in the following theorem. The argument above constitutes a proof of the theorem. For historical reasons, the convention is to define the quantity $AC - B^2$ as the discriminant. We denote this as D.

Theorem 1. *Suppose* (x_0, y_0) *is a critical input for a function* $f : \mathbb{R}^2 \to \mathbb{R}$. *Also, suppose* f *has continuous second partial derivatives in some open disk centered at* (x_0, y_0). *Let*

$$D = AC - B^2 = f_{xx}(x_0, y_0) f_{yy}(x_0, y_0) - [f_{xy}(x_0, y_0)]^2 .$$

If $D < 0$, *then* (x_0, y_0) *corresponds to a saddle point. If* $D > 0$, *then* (x_0, y_0) *is a local extremizer. In this case, if* $A > 0$ *and* $C > 0$ *then* (x_0, y_0) *is a local minimizer. If* $A < 0$ *and* $C < 0$, *then* (x_0, y_0) *is a local maximizer.*

Note that the theorem is silent in the case $D = AC - B^2 = 0$.

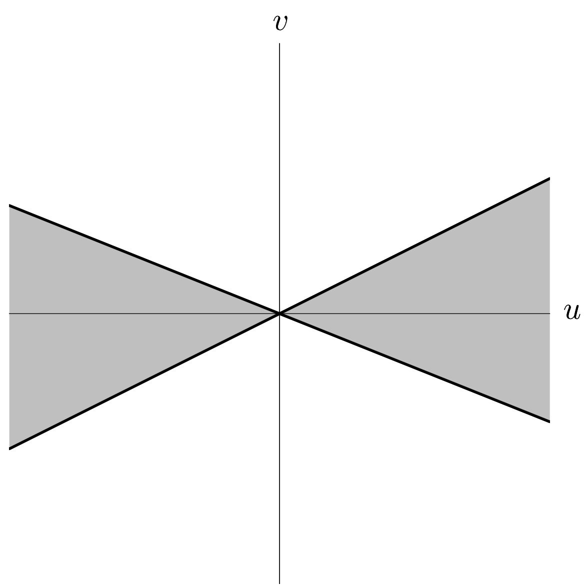

Figure 1: For a quadratic saddle there are two sections in which the concavity is zero. These correspond to the two lines through the origin in the uv-plane. The concavity is the same sign for all sections in the gray "wedges" and of opposite sign for all sections in the white "wedges."