The second derivative test

The purpose of this note is to help you appreciate the terrain around the second derivate test in two (or more) variables:

Proposition 1. *Consider a quadratic expression of the form*

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

whose coefficients A, B, C are constants that are not all 0.

i) $Q(u,v)$ is positive for all $(u,v) \neq (0,0)$ if and only if $A > 0$ and $AC - B^2 > 0$. If this is the case, then there is a linear change of variables: $u = u(s,t)$, $v = v(s,t)$, such that $Q(u,v) = s^2 + t^2$ (sum of squares, so positive whenever $(s,t) \neq (0,0)$).

ii) $Q(u,v)$ is negative for all $(u,v) \neq (0,0)$ if and only if $A < 0$ and $AC - B^2 > 0$. If this is the case, then there is a linear change of variables: $u = u(s,t)$, $v = v(s,t)$, such that $Q(u,v) = -(s^2 + t^2)$ (negative of sum of squares, so negative whenever $(s,t) \neq (0,0)$).

iii) If $AC - B^2 < 0$, there is a linear change of variables: $u = u(s,t)$, $v = v(s,t)$, such that $Q(u,v) = s^2 - t^2$ (a saddle).

iv) If $AC - B^2 = 0$, there is a substitution as above so that $Q(u,v) = s^2$ (the other variable is absent).

Note that in statements (i) and (ii), A and C necessarily have the same sign, because $AC > B^2 \geq 0$.

We'll take Proposition 1 as given, and use it for deciding the parallel result:

Proposition 2. *Let $f(x,y)$ be a function with continuous 2nd order partial derivatives on some region D, and let $(x_0, y_0) \in D$. Suppose that $\nabla f(x_0, y_0) = \mathbf{0}$. Let $A = f_{xx}(x_0, y_0)$, $B = f_{xy}(x_0, y_0)$, and $C = f_{yy}(x_0, y_0)$. Then*

i) if $A > 0$ and $AC - B^2 > 0$, then f has a local minimum at (x_0, y_0);

ii) if $A < 0$ and $AC - B^2 > 0$, then f has a local maximum at (x_0, y_0);

iii) if $AC - B^2 < 0$, then f has a saddle point at (x_0, y_0).

iv) if $AC - B^2 = 0$, there is insufficient information to say anything.

We will treat only case (i). Write $\mathbf{x}_0$ for (x_0, y_0), etc. By definition, the condition for a local minimum at $\mathbf{x}_0$ is that $f(\mathbf{x}) \geq f(\mathbf{x}_0)$ for all $\mathbf{x}$ in some neighborhood of $\mathbf{x}_0$. Evidently, this holds if and only if the inequality holds along every line through $\mathbf{x}_0$.

The lines through $\mathbf{x}_0$ are all parametrized as $\mathbf{x} = \mathbf{x}_0 + t\mathbf{u}$, where $\mathbf{u}$ can be taken to be a unit vector; the parametrization is such that $t = 0$ gives $\mathbf{x}_0$. In other words, we want to look at the functions

$$g(t) = f(\mathbf{x}_0 + t\mathbf{u}) \tag{1}$$

near $t = 0$, where $\mathbf{u}$ is fixed, but arbitrary. That's a standard way of saying that we will consider all directions simultaneously.

From last week's stuff, we should know that

$$g'(t) = \nabla f(\mathbf{x}_0 + t\mathbf{u}) \bullet \mathbf{u}, \tag{2}$$

so $g'(0) = 0$ for all $\mathbf{u}$ if and only if $\nabla f(\mathbf{x}_0) = \mathbf{0}$.

We want to look at the second derivative, but first, we put $\mathbf{u} = u\mathbf{i} + v\mathbf{j}$, i.e., write $\mathbf{u}$ in terms of its components (no big deal). Then

$$g'(t) = uf_x(\mathbf{x}_0 + t\mathbf{u}) + vf_y(\mathbf{x}_0 + t\mathbf{u}). \tag{3}$$

We see two expressions like what's in (1), except in (3) we have the partial derivatives of f instead of f. Differentiating (3), we get (compare (2))

$$g''(t) = u\nabla f_x(\mathbf{x}_0 + t\mathbf{u}) \bullet \mathbf{u} + v\nabla f_y(\mathbf{x}_0 + t\mathbf{u}) \bullet \mathbf{u}, \tag{4}$$

which we unwind as in (3):

$$g''(t) = u^2 f_{xx}(\mathbf{x}_0 + t\mathbf{u}) + uvf_{xy}(\mathbf{x}_0 + t\mathbf{u}) + vuf_{yx}(\mathbf{x}_0 + t\mathbf{u}) + v^2 f_{yy}(\mathbf{x}_0 + t\mathbf{u}). \tag{5}$$

Putting $t = 0$ (as we were dying to do), we obtain

$$g''(0) = u^2 f_{xx}(\mathbf{x}_0) + uvf_{xy}(\mathbf{x}_0) + vuf_{yx}(\mathbf{x}_0) + v^2 f_{yy}(\mathbf{x}_0). \tag{6}$$

Now, if f is at all reasonable (e.g., if f has continuous second partial derivatives), the mixed partials are equal so we can write that as

$$g''(0) = u^2 f_{xx}(\mathbf{x}_0) + 2uvf_{xy}(\mathbf{x}_0) + v^2 f_{yy}(\mathbf{x}_0) = Au^2 + 2Buv + Cv^2, \tag{7}$$

in the notation of Proposition 2. If $A > 0$ and $AC - B^2 > 0$, we may apply (i) of Proposition 1 to conclude that (7) is always positive when $(u, v) \neq (0, 0)$. This gives (i) of Proposition 2—actually that f has a *strict* local minimum at $\mathbf{x}_0$ (the strict inequality $f(\mathbf{x}) > f(\mathbf{x}_0)$ holds for all $\mathbf{x}$ in some deleted neighborhood of $\mathbf{x}_0$).

In short, critical point theory in two variables is a direct consequence of understanding quadrics in two variables.