Finding Max/Min of Functions of Two Variables with TD

Given a function \( f(x, y) \) with total derivative the problem of determining its maximum and minimum values means to locate those points, \((a, b)\) for which the difference \( f(x, y) - f(a, b) \) does not change sign for all points \((x, y)\) in some disc around \((a, b)\). If \( f(x, y) - f(a, b) \geq 0 \) for all points \((x, y)\) in some disc around \((a, b)\), \( f(x, y) \) has a maximum at \((a, b)\), while if \( f(x, y) - f(a, b) \leq 0 \) on such a disc, \( f(x, y) \) has a minimum at \((a, b)\). Suppose \( f(x, y) \) has an extremum (maximum or minimum) at \((a, b)\) and consider any direction \( \hat{u} = (u_x, u_y) \). The directional derivative

\[
D_\hat{u} f = \nabla f \cdot \hat{u} = \frac{\partial f}{\partial x} u_x + \frac{\partial f}{\partial y} u_y
\]

must be zero at \((a, b)\) since

Both expressions equal \( D_\hat{u} f \) in the limit that \( s \to 0 \). Since at an extremum \( D_\hat{u} f = \nabla f \cdot \hat{u} \) must vanish for every direction \( \hat{u} \), a necessary condition for a point \((x, y)\) to be the input coordinates of an extremum is that gradient of \( f \) must vanish at \((x, y)\). This condition is not sufficient. For example, \( f(x, y) = x^2 - y^2 \) has \( \nabla f = (2x, -2y) \) which vanishes at \((0, 0)\). It is obvious that the origin is not an extremum of this function. Moving along the \( x \) axis \( f(x, 0) = x^2 \) so on this path the origin is a minimum, but along the \( y \) axis \( f(0, y) = -y^2 \), so along this path the origin is a maximum. It just isn’t true that along all paths out of the origin that \( f(x, y) - f(0,0) \) has the same sign. When the gradient vanishes (meaning a horizontal tangent plane parallel to the \( x - y \) plane) but the function does not have an extremum, the point is called a “saddle” point. Points where the gradient is zero are called critical points. They can be maximums, minimums or saddles. In fact, they must be one of the three. Having found the critical points by setting the gradient of the function equal to zero, it is sometimes easy to determine whether the point is a maximum, minimum or saddle by directly analyzing the difference \( f(x, y) - f(a, b) \). For example, consider \( f(x, y) = 3x^2 + 18xy + y^2 \), \( \nabla f = (6x + 18y, 18x + 2y) \), which is zero when \( x = -3y \) and \( y = -9x \), so that the only critical point is the origin. Now,

\[
f(x, y) - f(0,0) = 3(x^2 + 6xy + 9y^2) - 26y^2 = 3(x + 3y)^2 - 26y^2.
\]

Now, along the \( x \) axis,

\[
f(x, 0) - f(0,0) = 3x^2 \geq 0,
\]

but along the line \( y = -\frac{1}{3}x \),

\[
f(x, -\frac{1}{3}x) - f(0,0) = -26y^2 \leq 0.
\]

So, this function has a saddle point at the origin.

In analogy with calculus of a single variable there is a second derivative test for functions of two variables having continuous second derivatives at its critical points. As a precursor we will first develop the Taylor series for a function of two variables. Suppose that \( f \) has continuous partial derivatives to all orders at \((a, b)\). Then the function \( g(t) \) defined as \( g(t) = f(a + th, b + tk) \) has derivatives to all order.

From the chain rule,

\[
\frac{dg}{dt} = \frac{df}{dx} \frac{dx}{dt} + \frac{df}{dy} \frac{dy}{dt} = f_x(a + th, b + tk)h + f_y(a + th, b + tk)k.
\]

Similarly,

\[
\frac{d^2g}{dt^2} = h \frac{d^2f}{dx^2} + h \frac{d^2f}{dy^2} + k \frac{d^2f}{dx^2} + k \frac{d^2f}{dy^2}.
\]

From Clairout’s Theorem,
\[ g''(t) = h^2 f_{xx}(a + th, b + tk) + 2hkf_{xy}(a + th, b + tk) + k^2 f_{yy}(a + th, b + tk). \]

Using mathematical induction one can establish the following result that is similar to the Binomial Theorem:
\[
\frac{d^n g}{dt^n} = \sum_{m=0}^{n} \frac{n!}{(m-n)!n!} h^{n-m} k^m \frac{\partial^n f}{\partial x^{n-m} \partial y^m}, \text{ with all partial derivatives evaluated at } (a + th, b + tk).
\]

The Maclaurin series \( g(t) = g(0) + g'(0)t + \frac{g''(0)}{2!}t^2 + \cdots \). Letting \( t = 1 \) gives the Taylor series for \( f(a+h, b+k) \),
\[
f(a+h, b+k) = f(a,b) + f_x(a,b)h + f_y(a,b)k + \frac{1}{2} [f_{xx}(a,b)h^2 + 2f_{xy}(a,b)hk + f_{yy}(a,b)k^2] + \cdots
\]

At an extremum, \( \nabla f = \mathbf{0} \), so \( \frac{\partial f}{\partial x} = \frac{\partial f}{\partial y} = 0 \), and for small \( h \) and \( k \), the difference
\[ \Delta f = f(a+h, b+k) - f(a,b) \]
will be well approximated by the quadratic expression
\[ \Delta f \approx \frac{1}{2} \left[ f_{xx}(a,b)h^2 + 2f_{xy}(a,b)hk + f_{yy}(a,b)k^2 \right]. \]

Now if \( f_{xx}(a,b) \) and \( f_{yy}(a,b) \) differ in sign \( f_{xy}(a,b) \) (excluding the case where either is zero), \( f(x,y) \) has a saddle at \( (a,b) \). If both \( f_{xx}(a,b) \) and \( f_{yy}(a,b) \) are zero and \( f_{xy}(a,b) \neq 0 \) then \( f(x,y) \) has a saddle at \( (a,b) \) since \( hk \) can change sign. Assume therefore that \( f_{xx}(a,b) \neq 0 \), then
\[ \Delta f \approx \frac{1}{2} \left[ \frac{(f_{xx}(a,b)h)^2 + 2f_{xx}(a,b)f_{xy}(a,b)hk}{f_{xx}(a,b)} + f_{yy}(a,b)k^2 \right] \]
\[ = \frac{1}{2} \left[ f_{xx}(a,b)h + f_{xy}(a,b)k \right]^2 + k^2 \left( f_{yy}(a,b)f_{xx}(a,b) - [f_{xy}(a,b)]^2 \right) \]
\[ \frac{f_{xx}(a,b)}{f_{xx}(a,b)} \]

Now, if \( f_{yy}(a,b)f_{xx}(a,b) - [f_{xy}(a,b)]^2 > 0 \), \( \Delta f \) will have the same sign as \( f_{xx}(a,b) \) and \( f(x,y) \) must have an extremum at \( (a,b) \). If \( f_{yy}(a,b)f_{xx}(a,b) - [f_{xy}(a,b)]^2 < 0 \), then on the path \( f_{xx}(a,b)h + f_{xy}(a,b)k = 0 \) \( \Delta f \) will have the opposite sign to \( \Delta f \) on the path \( k = 0 \) so that \( f(x,y) \) must have a saddle at \( (a,b) \).

**Summary of the Second Derivative Test:**

If \( \nabla f = \mathbf{0} \) at \( (a,b) \) and
1. \( f_{yy}(a,b)f_{xx}(a,b) - [f_{xy}(a,b)]^2 > 0 \) and \( f_{xx}(a,b) > 0 \), then \( f(x,y) \) has a minimum at \( (a,b) \)
2. \( f_{yy}(a,b)f_{xx}(a,b) - [f_{xy}(a,b)]^2 > 0 \) and \( f_{xx}(a,b) < 0 \), then \( f(x,y) \) has a maximum at \( (a,b) \)
3. \( f_{yy}(a,b)f_{xx}(a,b) - [f_{xy}(a,b)]^2 < 0 \), then \( f(x,y) \) has a saddle at \( (a,b) \)
4. \( f_{yy}(a,b)f_{xx}(a,b) - [f_{xy}(a,b)]^2 = 0 \), the test fails and one needs to consider higher order terms in the Taylor expansion of \( f(x,y) \) about \( (a,b) \)