Local Maxima and Minima

References are to Salas/Hille/Etgen’s Calculus, 8th Edition

We study the behavior of the scalar-valued function \( f(\mathbf{r}) \) of the 2-dimensional vector variable \( \mathbf{r} \) near a stationary point \( \mathbf{r}_0 \) (one where \( \nabla f(\mathbf{r}_0) = 0 \)). We wish to determine whether \( f \) has a local maximum or minimum at \( \mathbf{r}_0 \). We assume throughout that \( f \) has class \( C^2 \) (all second-order partial derivatives of \( f \) exist and are continuous).

There is one situation that occurs naturally in \( n \) dimensions, for \( n \geq 2 \).

**Definition 1** The point \( \mathbf{r}_0 \) is a saddle point of \( f \) if it is stationary \( (\nabla f(\mathbf{r}_0) = 0) \), but is neither a local maximum nor a local minimum of \( f \). (For a good picture, see p. 916 or p. 834; you can fill in the horse.)

**The second derivative** We reduce to a one-dimensional problem by restricting attention to the values of \( f \) along a line through \( \mathbf{r}_0 \). We choose any vector \( \mathbf{h} = h_1 \mathbf{i} + h_2 \mathbf{j} \) and set \( g(t) = f(\mathbf{r}_0 + t\mathbf{h}) \). Then by the chain rule,

\[
g'(t) = \nabla f(\mathbf{r}_0 + t\mathbf{h}) \cdot \mathbf{h} = f_x(\mathbf{r}_0 + t\mathbf{h}) h_1 + f_y(\mathbf{r}_0 + t\mathbf{h}) h_2.
\]

In particular, \( g'(0) = 0 \), and Taylor’s Theorem simplifies to

\[
g(t) = f(\mathbf{r}_0 + t\mathbf{h}) = g(0) + \frac{t^2}{2} g''(\theta t) = f(\mathbf{r}_0) + \frac{t^2}{2} g''(\theta t),
\]

for some \( \theta \) (depending on \( t \)), where \( 0 < \theta < 1 \). It is now clear by continuity that:

(a) If \( g''(0) > 0 \), then \( g \) has a strict local minimum at \( t = 0 \). In other words, \( g(t) > g(0) \) for all small \( t \neq 0 \). So \( g(t) \) does not have a local maximum at \( t = 0 \), and \( f(\mathbf{r}) \) does not have a local maximum at \( \mathbf{r}_0 \).

(b) If \( g''(0) < 0 \), then \( g \) has a strict local maximum at \( t = 0 \). In other words, \( g(t) < g(0) \) for all small \( t \neq 0 \). So \( g(t) \) does not have a local minimum at \( t = 0 \), and \( f(\mathbf{r}) \) does not have a local minimum at \( \mathbf{r}_0 \).

(c) If \( g''(0) = 0 \), we have no information about \( g''(\theta t) \) or \( g(t) \).

We compute that

\[
g''(0) = Q(\mathbf{h}) = Ah_1^2 + 2Bh_1h_2 + Cb_2^2,
\]

a quadratic form in \( \mathbf{h} \), which we abbreviate to \( Q(\mathbf{h}) \), with coefficients

\[
A = f_{xx}(\mathbf{r}_0), \quad B = f_{xy}(\mathbf{r}_0), \quad C = f_{yy}(\mathbf{r}_0).
\]

**Classification** There are six types of quadratic form \( Q(\mathbf{h}) \):

(i) **positive-definite**: \( Q(\mathbf{h}) > 0 \) for all \( \mathbf{h} \neq 0 \);

(ii) **positive-semidefinite**: \( Q(\mathbf{h}) \) takes both positive and zero values for \( \mathbf{h} \neq 0 \);

(iii) **negative-definite**: \( Q(\mathbf{h}) < 0 \) for all \( \mathbf{h} \neq 0 \);

(iv) **negative-semidefinite**: \( Q(\mathbf{h}) \) takes both negative and zero values for \( \mathbf{h} \neq 0 \);

(v) **indefinite**: \( Q(\mathbf{h}) \) takes both positive and negative values (and hence also zero values) for \( \mathbf{h} \neq 0 \);

(vi) **zero**: \( Q(\mathbf{h}) = 0 \) for all \( \mathbf{h} \).
Local Maxima and Minima

Except for (vi), the type of a given quadratic form \( Q(h) \) is not obvious in general.

Case 1: \( A \neq 0 \). We complete the square, by rewriting equation (4) as

\[
Q(h) = A \left( h_1 + \frac{B}{A} h_2 \right)^2 + \left( C - \frac{B^2}{A} \right) h_2^2 = A \left( h_1 + \frac{B}{A} h_2 \right)^2 - \frac{D}{A} h_2^2, \tag{5}
\]

where we introduce the discriminant \( D = B^2 - AC \) of \( Q(h) \). [WARNING: Many books (with good reason) change the sign and consider \( AC - B^2 \) rather than \( B^2 - AC \).] Since the expressions \( h_2 \) and \( h_1 + (B/A)h_2 \) can be chosen independently, all we have to do is check the signs of the coefficients \( A \) and \( -D/A \).

Case 2: \( A = 0 \). Here, \( D = B^2 \), and we write

\[
Q(h) = (2Bh_1 + Ch_2)h_2. \tag{6}
\]

**Theorem 7** The behavior of \( f \) near \( r_0 \) is given by the following table:

<table>
<thead>
<tr>
<th>Type of ( Q(h) )</th>
<th>Conditions</th>
<th>Local minimum?</th>
<th>Local maximum?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive-definite</td>
<td>( D &lt; 0, A &gt; 0 )</td>
<td>yes, strict</td>
<td>no</td>
</tr>
<tr>
<td>Positive-semidef.</td>
<td>( D = 0, (A &gt; 0 \text{ or } C &gt; 0) )</td>
<td>maybe</td>
<td>no</td>
</tr>
<tr>
<td>Negative-definite</td>
<td>( D &lt; 0, A &lt; 0 )</td>
<td>no</td>
<td>yes, strict</td>
</tr>
<tr>
<td>Negative-semidef.</td>
<td>( D = 0, (A &lt; 0 \text{ or } C &lt; 0) )</td>
<td>no</td>
<td>maybe</td>
</tr>
<tr>
<td>Indefinite</td>
<td>( D &gt; 0 )</td>
<td>no, saddle point</td>
<td>no, saddle point</td>
</tr>
<tr>
<td>Zero</td>
<td>( A = B = C = 0 )</td>
<td>maybe</td>
<td>maybe</td>
</tr>
</tbody>
</table>

This includes everything in Theorem 15.5.3, and a little more. Simple examples show that all statements are best possible. These results remain valid in \( n \) dimensions (with the second column greatly generalized, in ways that are not obvious).

**Proof** We read off the conditions from equations (5) and (6). The “no” statements follow directly from (3) and equation (5) or (6). If there exists \( h \) such that \( Q(h) > 0 \), then \( f \) does not have a local maximum; if there exists \( h \) such that \( Q(h) < 0 \), then \( f \) does not have a local minimum.

For the positive-definite case, we use equation (2) to write

\[
f(r_0 + h) = g(1) = f(r_0) + \frac{1}{2}Q(h),
\]

where \( Q(h) \) is given by equation (5), except that we must now evaluate \( A, B, C \) and \( D \) at \( r_0 + \theta h \) rather than \( r_0 \). Since these functions are all continuous, we still have \( D < 0 \) and \( A > 0 \) for all small \( h \), and it is clear that \( Q(h) > 0 \) for all small \( h \neq 0 \).

The negative-definite case is entirely similar. Alternatively, we can simply consider \(-f\) instead of \( f \). \( \square \)