Maxima, Minima and Saddle Points

Remember how we go about finding the max/min points of a 1D curve $f(x)$ - we find what values of $x$ set the derivative $f'(x)$ equal to zero. Whether the turning points are maxima or minima depends on whether $f''(x)$ is negative or positive, respectively. We can use similar concepts to investigate maxima and minima in 2D.

Let’s say that we have a function of two variables $f(x,y)$. What happens if we are at a point $(a,b)$ where

$$f_x = f_y = 0 \quad ?$$

(1)

Here we are writing $f_x$ as shorthand for $\frac{\partial f}{\partial x}$. Equation (1) means that the curve in the x-direction $f(x,b)$ and the curve in the y-direction $f(a,y)$ both have extreme values (maxima or minima) (see Lecture 1). It also means that the plane which is tangent to the surface at the point (the tangent plane) is also horizontal.

If the tangent plane is locally horizontal at $(a,b)$, then this means that the surface $f(x,y)$ is at a critical point. Critical points may be maxima, minima, inflection points or saddle points. The first three are familiar, the last one is new.

A saddle point at $(a,b)$ means that the surface is at a minimum when sliced along one direction, but at a maximum when sliced along the orthogonal direction. The name comes from the surface’s resemblance to . . . (you guessed it).

To locate maxima, minima and saddle points we first identify points $(a,b)$ where

$$f_x = f_y = 0$$

We then use information from the second partial derivatives to determine what type of points we have. The test we use is based on a variable we will call $D$:

$$D = D(a,b) = f_{xx}(a,b)f_{yy}(a,b) - [f_{xy}(a,b)]^2$$

There are four possible outcomes of this test:

1. $D > 0$ and $f_{xx}$ (or $f_{yy}$) $> 0$: $f(a,b)$ is a minimum
2. $D > 0$ and $f_{xx}$ (or $f_{yy}$) $< 0$: $f(a, b)$ is a maximum
3. $D < 0$: $f(a, b)$ is not a local extremum (it’s a saddle point)
4. $D = 0$: indeterminate

**Example** What do the following curves look like in the $x - z$ and $y - z$ planes, where are their critical points, and what sort of critical points are they?

a) $z = x^2 + y^2$

b) $z = -(x^2 + y^2)$

c) $z = x^2 - y^2$

**Example** Identify the location and nature of the critical points of the following curve:

$$z = x^3 - x + 2y^3 - 6y$$

Being able to find maxima/minima in several dimensions is obviously useful for optimization problems. You can also find optimum solutions subject to an imposed constraint, using the method of Lagrange Multipliers. This goes beyond the scope of the current course, but is pretty clever: you can prove e.g. that the shortest distance between two points is a straight line or (slightly more useful) determine the shape that a chain will hang in if supported at both ends (it’s called a catenary).

**Fitting a curve to data**

A very useful application of the 2D max/min approach is in fitting a curve to data, known as the least squares approach.

Say we have a series of $n$ observations $(x_i, y_i)$ and we also have a model which gives a predicted value of $y_i$, call it $y'_i$, for each $x_i$. We want to minimize the misfit between the observed and predicted $y$ values. To do this, we set up a measure of the misfit $E$:

$$E = \sum_{i=1}^{n} (y_i - y'_i)^2$$

where $\sum_{i=1}^{n}$ denotes taking the sum over $n$ points. We want to find the minimum value for this quantity as we vary the model parameters.

If we think the relationship between $x$ and $y$ should be a straight line, then our model should be

$$y'_i = mx_i + b$$
where $m$ and $b$ are the model parameters to be determined by minimizing the misfit $E$.

So we can rewrite $E$ as

$$E = \sum_{i=1}^{n} (y_i - mx_i - b)^2$$

So now $E$ is a function of $m$ and $b$, the two variables we are interested in. In order to find the minimum of $E$ as we vary $m$ and $b$, we can use the approach outlined above and deduce that

$$\frac{\partial E}{\partial m} = \frac{\partial E}{\partial b} = 0$$

at the minimum-misfit point. So now all we have to do is to find these partial differentials.

$$\frac{\partial E}{\partial m} = -2 \sum_{i=1}^{n} (y_ix_i - mx_i^2 - bx_i)$$

and

$$\frac{\partial E}{\partial b} = -2 \sum_{i=1}^{n} (y_i - mx_i - b)$$

At the minimum point, both of these quantities are zero, so that

$$m \sum x_i^2 + b \sum x_i = \sum x_i y_i$$

and

$$m \sum x_i + b \sum 1 = \sum y_i$$

Note that summation from $i = 1$ to $n$ is assumed. If we write

$$s = \sum x_i, \quad t = \sum y_i, \quad u = \sum x_i^2, \quad v = \sum x_i y_i$$

then we end up with the final result

$$m = \frac{ts - nv}{s^2 - un}, \quad b = \frac{vs - ut}{s^2 - un}$$

This is a very important result - given the observations, we can find the straight line which minimizes the sum of the squared errors. (The approach was developed by Gauss). Also note that we are not restricted to straight lines - you can apply this approach to almost any model you care to think of.

**Example** Say we have a data table with only three points:
Table 1:

<table>
<thead>
<tr>
<th>$x_i$</th>
<th>$y_i$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0</td>
<td>3.0</td>
</tr>
<tr>
<td>2.0</td>
<td>4.9</td>
</tr>
<tr>
<td>3.0</td>
<td>7.1</td>
</tr>
</tbody>
</table>

You can see by inspection that the answer is going to be close to $y = 2x + 1$.

From this table we have $n = 3$,

$$s = \sum x_i = 6.0 \quad , \quad t = \sum y_i = 15.0 \quad , \quad u = \sum x_i^2 = 14.0 \quad , \quad v = \sum x_i y_i = 34.1$$

which gives us in turn

$$m = \frac{(15.0 \cdot 6.0) - (3 \cdot 34.1)}{(6.0^2) - (3 \cdot 14.0)} = 2.05 \quad , \quad b = \frac{(34.1 \cdot 6.0) - (14.0 \cdot 15.0)}{(6.0^2) - (3 \cdot 14.0)} = 0.9$$

**Reminder** Mid-term is next Tuesday (25th Oct).