EXTREMA FOR SEVERAL VARIABLE FUNCTIONS

Maxima & Minima
For the next few lectures we will see how (partial) derivatives of functions of several variables \( f \) can be used to determine properties of \( f \).

Let \( f : X \subset \mathbb{R}^n \to \mathbb{R} \) be a scalar-valued function.

- We say that \( a \in X \) is a local minimum if there exists \( r > 0 \) such that
  \[ f(a) \leq f(x) \quad \text{whenever } |x - a| < r. \]

- We say that \( a \) is a local maximum if there exists \( r > 0 \) such that
  \[ f(x) \leq f(a) \quad \text{whenever } |x - a| < r. \]

Example: Let \( f(x, y) = x^2 + y^2 \). Then, \((0, 0)\) is a local minimum: take \( r = 1 \) (for example) then, for any \( x = (x, y) \) such that \( |x| = \sqrt{x^2 + y^2} < 1 \) we have \( f(x, y) = x^2 + y^2 \geq 0 = f(0, 0) \).

Remark: As we will soon see, this example is indicative of the behaviour of a function near a local minimum.

Suppose that \( a = (a, b) \in \mathbb{R}^2 \) is a local maximum of the differentiable function \( f(x, y) \). Then, for \( h \in \mathbb{R} \) sufficiently small, we have

\[ f(a + h, b) - f(a, b) \leq 0. \]

In particular, if \( h > 0 \) is sufficiently small then

\[ \frac{f(a + h, b) - f(a, b)}{h} \leq 0 \quad \Rightarrow \quad \frac{\partial f}{\partial x}(a, b) \leq 0 \]

Meanwhile, if \( h < 0 \) is sufficiently small then

\[ \frac{f(a + h, b) - f(a, b)}{h} \geq 0 \quad \Rightarrow \quad \frac{\partial f}{\partial x}(a, b) \geq 0 \]

Hence, \( \frac{\partial f}{\partial x}(a, b) = 0 \). A similar argument shows that \( \frac{\partial f}{\partial y}(a, b) = 0 \).

Remark:
1. Proceeding as above, we can show that if \(a\) is a local minimum of \(f\) then 
\[
\frac{\partial f}{\partial x}(a) = \frac{\partial f}{\partial y}(a) = 0.
\]

2. An analogous result holds more generally:

**Multivariable Fermat's Theorem**

Suppose that \(f : X \subset \mathbb{R}^n \to \mathbb{R}\) is differentiable. If \(a\) is a local maximum/minimum of \(f\) then \(\nabla f(a) = 0\).

A point \(a \in X\) such that \(\nabla f(a) = 0\), or \(f\) is not differentiable at \(a\) is called a critical point of \(f\).

We have an approach to determining local maxima/minima of a function \(f\):

- Determine the critical points of \(f\);
- Check whether the critical points just found are local maxima/minima.

Consider the following level curve diagram of

\[
f(x, y) = \cos(x) \cos(y), \quad -4 \leq x, y \leq 4
\]

![Level Curve Diagram](image)

**Exercise:** Using the level curve diagram

- indicate all the local maxima/minima (there are nine points altogether);
- find a critical point that is not a local maximum/minimum (there are four critical points that are not local max/min).
How would we determine these points without the level curve diagram? First we determine the critical points of \( f(x, y) = \cos(x) \cos(y) \) - these are those points \((a, b)\) in the domain of \( f \) satisfying

\[
\begin{bmatrix} 0 & 0 \end{bmatrix} = \nabla f(a, b) = [-\sin(a) \cos(b) \quad -\cos(a) \sin(b)]
\]

\[\Rightarrow \quad \sin(a) \cos(b) = 0 = \cos(a) \sin(b)\]

\[\Rightarrow (a, b) = (r\pi/2, s\pi/2), \quad r, s \in \{-2, -1, 0, 1, 2\}\]

*Recall that the domain of \( f \) is \( \{(x, y) \mid -4 \leq x, y \leq 4\} \)*

For each of these critical points we check if they are local maximum/minimaums. For example, if \((a, b) = (0, 0)\) let's check how \( f(x, y) \) changes as we move away from \((0, 0)\): for \( h, k \) sufficiently small and nonzero

\[f(h, k) - f(0, 0) = \cos(h) \cos(k) - \cos(0) \cos(0) = \cos(h) \cos(k) - 1 < 0\]

Hence, \((0, 0)\) is a local maximum. Using the double angle formula \( \cos(A \pm B) = \cos(A) \cos(B) \mp \sin(A) \sin(B) \), a similar computation shows that

local maximum: \((0, 0), (-\pi, -\pi), (-\pi, \pi), (\pi, -\pi), (\pi, \pi)\)

local minimum: \((\pm\pi, 0), (0, \pm\pi)\)

What about the critical point \((\pi/2, \pi/2)\)? This point is neither a local maximum nor a local minimum. The level curve diagram indicates this point is not a local maximum or local minimum. Let’s check how \( f(x, y) \) changes as we move away from \((\pi/2, \pi/2)\): for \( h, k \) sufficiently small and nonzero

\[f(\pi/2 + h, \pi/2 + k) - f(\pi/2, \pi/2) = \cos(\pi/2 + h) \cos(\pi/2 + k) - \cos(\pi/2) \cos(\pi/2)\]

\[= \sin(h) \sin(k)\]

In particular:

\[f(\pi/2 + h, \pi/2 + k) - f(\pi/2, \pi/2) \begin{cases} < 0, & \text{if } h > 0, k < 0 \text{ or } h < 0, k > 0 \\ > 0, & \text{if } h, k > 0 \text{ or } h, k < 0 \end{cases}\]

Near to the critical point \((\pi/2, \pi/2)\), \( f(x, y) \) is both strictly increasing and strictly decreasing.

**Question:**

Is there a methodical way to determine whether a critical point is a local maximum/minimum/neither of a differentiable function \( f \)?

**Taylor’s Theorem: Second Order Formula**

Let \( f(x, y) \) be a function with continuous second partial (mixed) derivatives having domain \( X \subset \mathbb{R}^2 \). We've already seen a first order approximation to \( f(x, y) \) near to \((a, b) \in X\): this is the linearisation \( L(x, y) \) of \( f(x, y) \) near to \((a, b) \)

\[L(x, y) = f(a, b) + f_x(a, b)(x - a) + f_y(a, b)(y - b)\]

If \((a, b)\) is a critical point of \( f(x, y) \), so that \( \nabla f(a, b) = 0 \), then \( L(x, y) = f(a, b) \) is a constant function. A natural question is the following:
**Question:**

Is there a degree two polynomial function

\[ p_2(x, y) = \alpha + \beta(x - a) + \gamma(y - b) + \delta(x - a)^2 + \epsilon(x - a)(y - b) + \eta(y - b)^2 \]

that ‘closely approximates’ \( f(x, y) \) near to \((a, b)\)?

Suppose that such a polynomial function existed. Then, it would seem reasonable to expect

\[ p_2(a, b) = \alpha \]

\[ \frac{\partial p_2}{\partial x}(a, b) = \beta, \quad \frac{\partial f}{\partial y}(a, b) = \gamma \]

\[ \frac{\partial^2 f}{\partial x^2}(a, b) = 2\delta, \quad \frac{\partial^2 f}{\partial x\partial y}(a, b) = 2\epsilon, \quad \frac{\partial^2 f}{\partial y^2}(a, b) = 2\eta \]

In fact, it’s possible to show that such a polynomial does exist, in general:

**Taylor’s Theorem: Second Order Formula**

Let \( f : X \subset \mathbb{R}^n \rightarrow \mathbb{R} \) have continuous second order partial derivatives, \( a = (a_1, \ldots, a_n) \in X \). Then, there exists a degree two polynomial \( p_2(x) \), called the second order Taylor polynomial of \( f(x, y) \) near \( a \), such that

\[ \lim_{x \to a} \frac{|f(x) - p_2(x)|}{|x - a|^2} = 0 \]

Moreover,

\[ p_2(x) = f(a) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a)(x_i - a_i) + \frac{1}{2} \sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(a)(x_i - a_i)(x_j - a_j) \quad (*) \]

**Remark:**

1. For a function \( f(x, y) \) of two variables

\[ p_2(x, y) = f(a, b) + f_x(a, b)(x - a) + f_y(a, b)(y - b) + \frac{1}{2} f_{xx}(a, b)(x - a)^2 \]

\[ + f_{xy}(a, b)(x - a)(y - b) + \frac{1}{2} f_{yy}(a, b)(y - b)^2 \]

2. \((*)\) can be written in the compact form

\[ p_2(x) = f(a) + \nabla f(a)(x - a) + \frac{1}{2} (x - a)^t H f(a)(x - a) \]

where \( H f(a) \) is an \( n \times n \) symmetric matrix called the Hessian of \( f \). We will introduce the Hessian in the next lecture. For \( n = 2 \), the Hessian is

\[
H f(a, b) = \begin{bmatrix}
\frac{\partial^2 f}{\partial x^2}(a, b) & \frac{\partial^2 f}{\partial x \partial y}(a, b) \\
\frac{\partial^2 f}{\partial y \partial x}(a, b) & \frac{\partial^2 f}{\partial y^2}(a, b)
\end{bmatrix}
\]
Example:

1. Consider \( f(x, y) = \cos(x) \cos(y) \), \((a, b) = (0, 0)\). Then,

\[
f(0, 0) = 1, \quad f_x(0, 0) = f_y(0, 0) = 0, \quad f_{xx}(0, 0) = f_{yy}(0, 0) = -1, \quad f_{xy}(0, 0) = 0
\]

Hence, the second order Taylor polynomial of \( f(x, y) \) near \((0, 0)\)

\[
p_2(x) = 1 - \frac{1}{2} x^2 - \frac{1}{2} y^2 \quad \leftarrow \quad p_2(x, y) = f(0, 0)
\]

2. Consider the function \( f(x, y) = x^3 + 3xy + y^3 \). Then,

\[
f(1, 1) = 5, \quad \frac{\partial f}{\partial x}(1, 1) = 6 = \frac{\partial f}{\partial y}(1, 1),
\]

\[
\frac{\partial^2 f}{\partial x^2}(1, 1) = 6 = \frac{\partial^2 f}{\partial y^2}(1, 1), \quad \frac{\partial^2 f}{\partial x \partial y}(1, 1) = \frac{\partial^2 f}{\partial y \partial x}(1, 1) = 3
\]

Hence, the second order Taylor polynomial near to \((1, 1)\) is

\[
p_2(x, y) = 5 + 6(x - 1) + 6(y - 1) + 3(x - 1)^2 + 3(x - 1)(y - 1) + 3(y - 1)^2
\]

Next time: We will see how to use the second order Taylor polynomial to determine the nature of a critical point.