In this week's notes, we are interested in finding the *local maxima* / *local minima* of a multivariate function f, i.e. the points $(a_1, \dots, a_n)$ such that

$$\begin{matrix} f(a_1,\dots,a_n) \geq f(x_1,\dots,x_n) \\ f(a_1,\dots,a_n) \leq f(x_1,\dots,x_n) \end{matrix} \quad \text{for all } (x_1,\dots,x_n) \text{ close to } f(a_1,\dots,a_n).$$

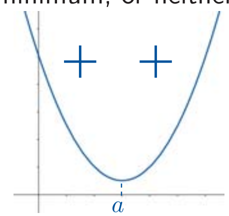
If $\nabla f(a_1, \dots, a_n) \neq \mathbf{0}$, then f is increasing in the direction $\nabla f(a_1, \dots, a_n)$ and decreasing in the direction $-\nabla f(a_1, \dots, a_n)$, so $(a_1, \dots, a_n)$ cannot be a local maximum or minimum. So a local maximum or minimum must be a critical point.

Definition: A point $(a_1, \dots, a_n)$ is a *critical point* of $f : \mathbb{R}^n \to \mathbb{R}$ if $\nabla f(a_1, \dots, a_n) = \mathbf{0}$, i.e. if all its partial derivatives are 0.

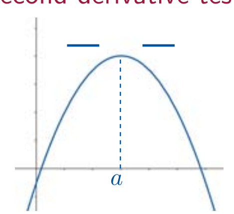
But not every critical point is a local maximum or minimum as we will see.

§13.1: Classifying Critical Points

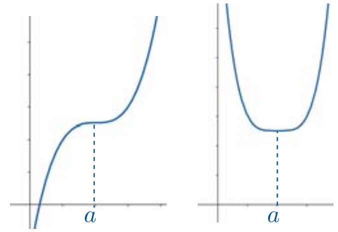
Recall that a critical point of a single-variable function f is where the derivative f' is zero. A standard way to determine whether it is a local maximum, a local minimum, or neither, is the second derivative test:



if $f''(a) > 0$ then a is a local minimum

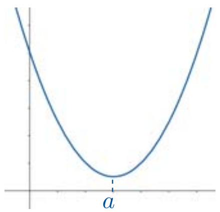


if $f''(a) < 0$ then a is a local maximum

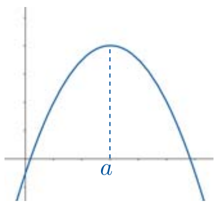


if $f''(a) = 0$ then we need to investigate further

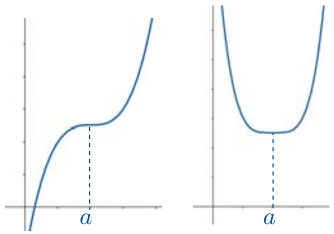
The reason is clear from considering the change in the slope of the graph, but because graphs of multivariate functions are hard to visualise, we give a different justification on the next page.



if $f''(a) > 0$ then a is a local minimum



if $f''(a) < 0$ then a is a local maximum



if $f''(a) = 0$ then we need to investigate further

The second-order Taylor polynomial of f at a is

$$f(a+h) \approx f(a) + \underbrace{f'(a)}_{\text{is 0 if } a \text{ is a critical point}}h + \frac{f''(a)}{2!}\underbrace{h^2}_{\text{is positive if } h \neq 0 \text{, i.e. } x \neq a}$$

$$= f(a) + \frac{f''(a)}{2!}h^2 \qquad \begin{cases} > f(a) & \text{if } f''(a) > 0 \text{ and } h \neq 0 \\ < f(a) & \text{if } f''(a) < 0 \text{ and } h \neq 0 \end{cases}$$

Here is a simplified example of how to use second order Taylor polynomials to classify critical points of multivariate functions.

Example: Find and classify the critical points of $f(x, y) = y^2 - x^3 + x$.

Now we develop a multivariate second derivative test by copying the previous example's argument in general.
The second-order Taylor polynomial of f about (a,b) is

$$f(a+h,b+k) \approx f(a,b) + \underbrace{f_x(a,b)h + f_y(a,b)k}_{\text{is 0 if }(a,b)\text{ is a critical point}} + \frac{\boxed{f_{xx}(a,b)h^2 + 2f_{xy}(a,b)hk + f_{yy}(a,b)k^2}}{2!},$$

is 0 if (a,b) is a critical point — we need the "sign" of the numerator

Definition: A function $Q:\mathbb{R}^n \to \mathbb{R}$ is a *quadratic form* if it is homogeneous of degree two i.e. a linear combination of $x_i x_j$. A quadratic form Q is:

- *positive definite* if $Q(\mathbf{x}) > 0$ for all $\mathbf{x} \neq \mathbf{0}$; $\implies$ minimum
- *negative definite* if $Q(\mathbf{x}) < 0$ for all $\mathbf{x} \neq \mathbf{0}$; $\implies$ maximum
- *indefinite* if $Q(\mathbf{x}) > 0$ for some $\mathbf{x} \neq \mathbf{0}$, and $Q(\mathbf{x}) < 0$ for some other $\mathbf{x} \neq \mathbf{0}$. $\implies$ not maximum nor minimum

A quadratic form can be at most one of the three types. But it is possible to be none of the three types, e.g. $Q(h,k) = h^2$. (see later)

Definition: A quadratic form Q is:

- *positive definite* if $Q(\mathbf{x}) > 0$ for all $\mathbf{x} \neq \mathbf{0}$; $\implies$ minimum
- *negative definite* if $Q(\mathbf{x}) < 0$ for all $\mathbf{x} \neq \mathbf{0}$; $\implies$ maximum
- *indefinite* if $Q(\mathbf{x}) > 0$ for some $\mathbf{x} \neq \mathbf{0}$, and $Q(\mathbf{x}) < 0$ for some other $\mathbf{x} \neq \mathbf{0}$. $\implies$ not maximum nor minimum

Let's start with the 2-variable case: any 2-variable quadratic form has the form $Ah^2 + 2Bhk + Ck^2$. (We are interested in $f_{xx}(a,b)h^2 + 2f_{xy}(a,b)hk + f_{yy}(a,b)k^2$.)
In the previous example where $B=0$, we can quickly tell the definiteness from the signs of A and C. In the general case, we will have to complete the square:

$$Ah^2 + 2Bhk + Ck^2 = A\left(h + \frac{B}{A}k\right)^2 + \frac{AC-B^2}{A}k^2 \qquad \det\begin{pmatrix} A & B \\ B & C\end{pmatrix}$$

So $Q(x)$ is positive definite
if A and $\frac{AC-B^2}{A}$ are both positive, i.e. $A>0$ and $\boxed{AC-B^2} > 0$;
$Q(x)$ is negative definite
if A and $\frac{AC-B^2}{A}$ are both negative, i.e. $A<0$ and $AC-B^2>0$;
$Q(x)$ is indefinite if A and $\frac{AC-B^2}{A}$ have different signs, i.e. $A \neq 0$ and $AC-B^2<0$.

To phrase this in a way that will extend to functions of more than 2 variables:
Definition: The *Hessian matrix* $\mathcal{H}(\mathbf{a})$ of $f:\mathbb{R}^n \to \mathbb{R}$ at a point $\mathbf{a}$ in $\mathbb{R}^n$ is the $n \times n$ matrix with $\dfrac{\partial^2 f}{\partial x_j x_i}(\mathbf{a})$ in row i and column j.
Example: For a 2-variable function, the Hessian matrix is $\mathcal{H}(a,b) = \begin{pmatrix} f_{xx}(a,b) & f_{xy}(a,b) \\ f_{yx}(a,b) & f_{yy}(a,b)\end{pmatrix}$.
Let $D_i(\mathbf{a})$ denote the determinant of the $i \times i$ matrix containing only the first i rows and the first i columns of $\mathcal{H}(\mathbf{a})$.
For a 2-variable function, $D_1(a,b) = f_{xx}(a,b)$ and $D_2(a,b) = \det\mathcal{H}(a,b)$. We saw previously that D_1 and D_2/D_1 are the coefficients after completing the square.
Theorem: Second Derivative Test for 2-variable functions: Let (a,b) be a critical point of $f:\mathbb{R}^2 \to \mathbb{R}$, and suppose all second-order partial derivatives of f are continuous near (a,b).
If $D_1(a,b), D_2(a,b) > 0$, then $\mathcal{H}(a,b)$ is positive definite and (a,b) is a minimum.
If $D_1(a,b) < 0, D_2(a,b) > 0$, then $\mathcal{H}(a,b)$ is negative definite and (a,b) is a maximum.
If $D_2(a,b) \neq 0$ and the above conditions do not hold, i.e. if $D_2(a,b) < 0$, then $\mathcal{H}(a,b)$ is indefinite and (a,b) is not a minimum or maximum, i.e. a *saddle point*.
If $D_2(a,b) = 0$, then the second derivative test is inconclusive; we need more info.

Example: Show that $(0,0)$ is a critical point of $f(x,y) = xy + y^2e^x - 3x^2$, and determine if it is a minimum, maximum or neither.

Now we describe the second derivative test in general.

Recall that the Hessian matrix $\mathcal{H}(\mathbf{a})$ has $\frac{\partial^2 f}{\partial x_j x_i}(\mathbf{a})$ in row i and column j. We are interested in the definiteness of the associated quadratic form $\sum_{i,j} \frac{\partial^2 f}{\partial x_j x_i}(\mathbf{a}) h_i h_j$.

Recall that $D_i(\mathbf{a})$ is the determinant of the $i \times i$ matrix containing only the first i rows and the first i columns of $\mathcal{H}(\mathbf{a})$. If none of the D_i are 0, then there is a way to complete the square in the quadratic form so the coefficients are D_1, D_2/D_1, D_3/D_2, ..., D_n/D_{n-1}. (The case where some $D_i = 0$ needs a different argument.)

Theorem 3: Second Derivative Test: (see also Theorem 8 §10.7) Let $\mathbf{a}$ be a critical point of $f : \mathbb{R}^n \to \mathbb{R}$, and suppose all second-order partial derivatives of f are continuous near $\mathbf{a}$.

If $D_i(\mathbf{a}) > 0$ for all i, then $\mathcal{H}(\mathbf{a})$ is positive definite and $\mathbf{a}$ is a minimum.

If $D_i(\mathbf{a}) < 0$ for all odd i and $D_i(\mathbf{a}) > 0$ for all even i, then $\mathcal{H}(\mathbf{a})$ is negative definite and $\mathbf{a}$ is a maximum.

If $D_n(\mathbf{a}) \neq 0$ and the above conditions do not hold, then $\mathcal{H}(\mathbf{a})$ is indefinite and $\mathbf{a}$ is not a maximum nor minimum, i.e. a saddle point.

If $D_n(\mathbf{a}) = 0$, then the second derivative test is inconclusive; we need more info.

Example: $(1,0,1)$ is a critical point of the function

$$f(x,y,z) = \frac{4x}{1+x^2+y} + yz - 2z^2 + 4z, \text{ and } \mathcal{H}(1,0,1) = \begin{pmatrix} -2 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & -4 \end{pmatrix}.$$

$\det \mathcal{H}(1,0,1) = 14 \neq 0$ so the second derivative test will have a conclusion.

$$D_1(1,0,1) = \left|-2\right| = -2 < 0$$

$$D_2(1,0,1) = \begin{vmatrix} -2 & 1 \\ 1 & 1 \end{vmatrix} = -3 < 0$$

$$D_3(1,0,1) = \begin{vmatrix} -2 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & -4 \end{vmatrix} = 14 > 0$$

The sign sequence of the D_i is $--+$, which is not $+++$ nor $-+-$, so $(1,0,1)$ is neither a minimum nor a maximum, i.e. a saddle point.

Non-examinable: definiteness by eigenvalues

Recall that the Hessian matrix $\mathcal{H}(\mathbf{a})$ has $\frac{\partial^2 f}{\partial x_j x_i}(\mathbf{a})$ in row i and column j. We are interested in the definiteness of the associated quadratic form $\sum_{i,j} \frac{\partial^2 f}{\partial x_j x_i}(\mathbf{a}) h_i h_j$.

By the equality of mixed partial derivatives, the Hessian matrix is a symmetric matrix. Recall from linear algebra that every symmetric matrix is diagonalisable with an orthonormal basis of eigenvectors. This basis gives a different way to complete the square in the quadratic form so the coefficients are precisely the eigenvalues. (The eigenvalues are not the numbers D_i/D_{i-1}.) So:

If all eigenvalues are positive, then the quadratic form is positive definite.

If all eigenvalues are negative, then the quadratic form is negative definite.

If there is at least one positive eigenvalue and at least one negative eigenvalue, then the quadratic form is indefinite.

This is slightly stronger than the D_i/D_{i-1} method: if a 3x3 symmetric matrix had one positive, one negative, and one zero eigenvalue, then the eigenvalue method says it is indefinite. But a zero eigenvalue means that the matrix is not invertible, so its determinant

is 0, so the D_i/D_{i-1} method is inconclusive.

Alternatives to the second derivative test

If $D_n(\mathbf{a}) = 0$, so the second derivative test is inconclusive, or if it is inconvenient to calculate second derivatives, then:

- We can show $\mathbf{a}$ is a saddle point by finding a direction (not a point!) where $f(\mathbf{x}) > f(\mathbf{a})$ and another direction where $f(\mathbf{x}) < f(\mathbf{a})$ (below, ex. sheet #18 Q2).
- We can show $\mathbf{a}$ is a $\begin{matrix}\text{maximum}\\ \text{minimum}\end{matrix}$ by showing that $\begin{matrix} f(\mathbf{x}) \leq f(\mathbf{a}) \\ f(\mathbf{x}) \geq f(\mathbf{a}) \end{matrix}$ for all $\mathbf{x}$ close to $\mathbf{a}$ (p13).

Example: Classify the critical point $(0,0)$ of $f(x,y) = x^2 + y^5$.

Example: Classify the critical point $(0,0)$ of $f(x,y) = x^2y^2 + x^3y^2$.

Algebraic manipulation such as factoring can also show that a certain point is a saddle point.

Example: Classify the critical point $(-1,0)$ of $f(x,y) = x^2y^2 + x^3y^2$.