20. Maxima and minima: I

Recall that $K \subset \mathbb{R}^n$ is **closed** if the complement is open. Recall also that this is equivalent to saying that K contains all of its limit points.

Definition 20.1. *We that $K \subset \mathbb{R}^n$ is **bounded** if there is a real number M such that*

$$\|x\| \leq M,$$

for all $x \in K$.

*We say that K is **compact** if K is closed and bounded.*

Note that K is bounded if and only if

$$K \subset B_M(O),$$

where O is the origin.

Example 20.2.

(1) $[a, b] \subset \mathbb{R}$ *is compact.*
(2) $(a, b] \subset \mathbb{R}$ *is bounded but not closed (whence not compact).*
(3) $[a, \infty) \subset \mathbb{R}$ *is closed but not bounded (whence not compact).*
(4)
$$K = \{\, x \in \mathbb{R}^n \mid \|x\| \leq M \,\},$$
is compact.
(5)
$$K = \{\, x \in \mathbb{R}^n \mid \|x\| < M \,\},$$
is bounded but not closed.

Theorem 20.3. *Let $f\colon K \longrightarrow \mathbb{R}$ be a continuous function.*

If K is compact then there are two points $\vec{a}_{\min}$ and $\vec{a}_{\max}$ in K such that

$$f(\vec{a}_{\min}) \leq f(\vec{x}) \leq f(\vec{a}_{\max}).$$

The proof of (20.3) is highly non-trivial.

Definition 20.4. *Let $K \subset \mathbb{R}^n$. We say that $\vec{a} \in K$ is an **interior point** if there is an open ball containing $\vec{a}$ which is contained in K.*

*Otherwise $\vec{a}$ is a **boundary point** of K.*

Example 20.5. *If $K = [a, b] \subset \mathbb{R}$ then every point $a < x < b$ is an interior point and a and b are boundary points.*

To find the maxima and minima of $f\colon K \longrightarrow \mathbb{R}$ we break the problem into two pieces:

I. The interior points. Use the derivative test, this lecture.
II. The boundary points. Use Lagrange multipliers, see lecture 21.

Notice that the boundary can have a rather complicated structure in higher dimensions. For example, in $\mathbb{R}^3$ the closed unit ball is compact. The boundary is the set of points in the open unit ball and the set of boundary points is the set of points on the unit sphere.

Definition 20.6. *Let $f\colon K \longrightarrow \mathbb{R}$ be a function and let $\vec{a} \in K$ be an interior point. We say that f has a* ***local minimum*** *at $\vec{a}$ if there is an open ball $U = B_\delta(\vec{a})$ centred at $\vec{a}$ contained in K such that*

$$f(\vec{a}) \leq f(\vec{x}),$$

for all $\vec{x} \in U$.

We can define ***local maxima*** *in a similar fashion.*

Definition 20.7. *Let $f\colon K \longrightarrow \mathbb{R}$ be a differentiable function. We say that a point $\vec{a} \in K$ in the interior of K is a* ***critical point*** *if $Df(\vec{a}) = \vec{0}$.*

Proposition 20.8. *Let $K \subset \mathbb{R}^n$ be a compact set and let $f\colon K \longrightarrow \mathbb{R}$ be a differentiable function. Let $\vec{a} \in K$ be an interior point.*

If $\vec{a}$ is a local maximum, then $\vec{a}$ is a critical point.

Proof.

$$\frac{\partial f}{\partial x_i}(\vec{a}) = \lim_{h \downarrow 0} \frac{f(\vec{a} + h\hat{e}_i) - f(\hat{e}_i)}{h} \leq 0$$
$$\frac{\partial f}{\partial x_i}(\vec{a}) = \lim_{h \uparrow 0} \frac{f(\vec{a} + h\hat{e}_i) - f(\hat{e}_i)}{h} \geq 0. \qquad \square$$

Recall the one variable second derivative test:

(i) If $f'(a) = 0$ and $f''(a) < 0$, then a is a local maximum of f.
(ii) If $f'(a) = 0$ and $f''(a) > 0$, then a is a local minimum of f.
(iii) If $f'(a) = 0$ and $f''(a) = 0$, then we don't know.

The reason why the second derivative works follows from Taylor's Theorem with remainder, applied to the second Taylor polynomial.

To figure out the multi-variable form of the second derivative test, we need to consider the multi-variable second Taylor polynomial:

$$P_{\vec{a},2}f(\vec{x}) = f(\vec{a}) + \nabla f(\vec{a}) \cdot \vec{h} + \frac{1}{2}\vec{h}^t Hf(\vec{a})\vec{h}.$$

Recall that

$$\vec{h} = (h_1, h_2, \dots, h_n) \qquad \text{and} \qquad Hf(\vec{a}) = (\frac{\partial^2 f}{\partial x_i \partial x_j}(\vec{a})).$$

The important term is then

$$Q(\vec{h}) = h^t Hf(\vec{a})h.$$

Definition 20.9. *If A is a symmetric $n \times n$ matrix, then the function*

$$Q(\vec{x}) = \vec{x}^t A \vec{x},$$

is called a ***symmetric quadratic form****.*

We say that Q is ***positive definite*** *if $\vec{x} \neq 0$ implies that $Q(\vec{x}) > 0$.*

We say that Q is ***negative definite*** *if $\vec{x} \neq 0$ implies that $Q(\vec{x}) < 0$.*

Example 20.10. *If $A = I_2$ then*

$$Q(x, y) = (x, y) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = x^2 + y^2,$$

which is positive definite. If $A = -I_2$ then $Q(x, y) = -x^2 - y^2$ is negative definite. Finally if

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

then $Q(x, y) = x^2 - y^2$ is neither positive nor negative definite.

Proposition 20.11. *If $\vec{a} \in K \subset \mathbb{R}^n$ is an interior point and $f\colon K \longrightarrow \mathbb{R}$ is $\mathcal{C}^3$ and $\vec{a}$ is a critical point, then*

1. *If $Q(\vec{h}) = \vec{h}^t Hf(\vec{a})\vec{h}$ is positive definite, then $\vec{a}$ is a minimum.*
2. *If $Q(\vec{h}) = \vec{h}^t Hf(\vec{a})\vec{h}$ is negative definite, then $\vec{a}$ is a maximum.*
3. *If the determinant of $Q(\vec{h}) = \vec{h}^t Hf(\vec{a})\vec{h}$ is not zero and $Q(\vec{h})$ is neither positive nor negative definite, then $\vec{a}$ is a saddle point.*

Proof. Immediate from Taylor's Theorem. $\square$

Proposition 20.12. *If A is a $n \times n$ matrix, then let d_i be the determinant of the upper left $i \times i$ submatrix. Let $Q(\vec{h}) = h^t A h$.*

1. *If $d_i > 0$ for all i, then Q is positive definite.*
2. *If $d_i > 0$ for i even and $d_i < 0$ for i odd, then Q is negative definite.*

Let's consider the 2×2 case.

$$A = \begin{pmatrix} a & b \\ b & d \end{pmatrix}$$

In this case

$$Q(x, y) = ax^2 + 2bxy + cy^2.$$

Assume that $d_1 = a > 0$. Let's complete the square. $a = \alpha^2$, some $\alpha > 0$.

$$Q(x, y) = (\alpha x + b/\alpha y)^2 + (d - b^2/\alpha^2)y^2 = (\alpha x + b/\alpha y)^2 + (ad - b^2)/ay^2.$$

In this case $d_1 = a > 0$ and $d_2 = ad - b^2$. So the coefficient of y^2 is positive if $d_2 > 0$ and negative if $d_2 < 0$.