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*

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

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 ***positive 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 $f : K \longrightarrow \mathbb{R}$ 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 $Q(\vec{h}) = \vec{h}^t Hf(\vec{a})\vec{h}$ is not zero and is neither positive nor negative definite, then $\vec{a}$ is a saddle point.*

Proof. Immediate from Taylor's Theorem. □

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 $Q(\vec{h}) = h^t Ah$.*

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 $d_2 < 0$.

21. Maxima and minima: II

To see how to maximise and minimise a function on the boundary, let's conside a concrete example.

Let

$$K = \{\, (x, y) \,|\, x^2 + y^2 \leq 2 \,\}.$$

Then K is compact. Let

$$f \colon K \longrightarrow \mathbb{R},$$

be the function $f(x, y) = xy$. Then f is continuous and so f achieves its maximum and minimum.

I. Let's first consider the interior points. Then

$$\nabla f(x, y) = (y, x),$$

so that $(0, 0)$ is the only critical point. The Hessian of f is

$$Hf(x, y) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

$d_1 = 0$ and $d_2 = -1 \neq 0$ so that $(0, 0)$ is a saddle point.

It follows that the maxima and minima of f are on the boundary, that is, the set of points

$$C = \{\, (x, y) \,|\, x^2 + y^2 = 2 \,\}.$$

II. Let $g \colon \mathbb{R}^2 \longrightarrow \mathbb{R}$ be the function $g(x, y) = x^2 + y^2$. Then the circle C is a level curve of g. The original problem asks to maximise and minimise

$$f(x, y) = xy \qquad \text{subject to} \qquad g(x, y) = x^2 + y^2 = 2.$$

One way to proceed is to use the second equation to eliminate a variable. The method of Lagrange multipliers does exactly the opposite. Instead of eliminating a variable we add one more variable, traditionally called λ. So now let's maximise and minimise

$$h(x, y, \lambda) = f(x, y) - \lambda(g(x, y) - 2) = xy - \lambda(x^2 + y^2 - 2).$$

We find the critical points of $h(x, y, \lambda)$:

$$\begin{aligned} y &= 2\lambda x \\ x &= 2\lambda y \\ 2 &= x^2 + y^2. \end{aligned}$$

First note that if $x = 0$ then $y = 0$ and $x^2 + y^2 = 0 \neq 2$, impossible. So $x \neq 0$. Similarly one can check that $y \neq 0$ and $\lambda \neq 0$. Divide the first equation by the second:

$$\frac{y}{x} = \frac{x}{y},$$

so that $y^2 = x^2$. As $x^2 + y^2 = 2$ it follows that $x^2 = y^2 = 1$. So $x = \pm 1$ and $y = \pm 1$. This gives four potential points $(1, 1)$, $(-1, 1)$, $(1, -1)$, $(-1, -1)$. Then the maximum value of f is 1, and this occurs at the first and the last point. The minimum value of f is -1, and this occurs at the second and the third point.

One can also try to parametrise the boundary:

$$\vec{r}(t) = \sqrt{2}(\cos t, \sin t).$$

So we maximise the composition

$$h\colon [0, 2\pi] \longrightarrow \mathbb{R},$$

where $h(t) = 2 \cos t \sin t$. As $I = [0, 2\pi]$ is compact, h has a maximum and minimum on I. When $h'(t) = 0$, we get

$$\cos^2 t - \sin^2 t = 0.$$

Note that the LHS is $\cos 2t$, so we want

$$\cos 2t = 0.$$

It follows that $2t = \pi/2 + 2m\pi$, so that

$$t = \pi/4, \qquad 3\pi/4, \qquad 5\pi/4, \qquad \text{and} \qquad 7\pi/4.$$

These give the four points we had before.

What is the closest point to the origin on the surface

$$F = \{\, (x, y, z) \in \mathbb{R}^3 \,|\, x \geq 0, y \geq 0, z \geq 0, xyz = p \,\}?$$

So we want to minimise the distance to the origin on F. The first trick is to minimise the square of the distance. In other words, we are trying to minimise $f(x, y, z) = x^2 + y^2 + z^2$ on the surface

$$F = \{\, (x, y, z) \in \mathbb{R}^3 \,|\, x \geq 0, y \geq 0, z \geq 0, xyz = p \,\}.$$

In words, given three numbers $x \geq$, $y \geq 0$ and $z \geq 0$ whose product is $p > 0$, what is the minimum value of $x^2 + y^2 + z^2$?

Now F is closed but it is not bounded, so it is not even clear that the minimum exists.

Let's use the method of Lagrange multipliers. Let

$$h\colon \mathbb{R}^4 \longrightarrow \mathbb{R},$$

be the function

$$h(x, y, z, \lambda) = x^2 + y^2 + z^2 - \lambda(xyz - p).$$

We look for the critical points of h:

$$\begin{aligned} 2x &= \lambda yz \\ 2y &= \lambda xz \\ 2z &= \lambda xy \\ p &= xyz. \end{aligned}$$

Once again, it is not possible for any of the variables to be zero. Taking the product of the first three equations, we get

$$8(xyz) = \lambda^3(x^2y^2z^2).$$

So, dividing by xyz and using the last equation, we get

$$8 = \lambda^3 p,$$

that is

$$\lambda = \frac{2}{p^{1/3}}.$$

Taking the product of the first two equations, and dividing by xy, we get

$$4 = \lambda^2 z^2,$$

so that

$$z = p^{1/3}.$$

So $h(x, y, z, \lambda)$ has a critical point at

$$(x, y, z, \lambda) = (p^{1/3}, p^{1/3}, p^{1/3}, \frac{2}{p^{1/3}}).$$

We check that the point

$$(x, y, z) = (p^{1/3}, p^{1/3}, p^{1/3}),$$

is a minimum of $x^2 + y^2 + z^2$ subject to the constraint $xyz = p$. At this point the sum of the squares is

$$3p^{2/3}.$$

Suppose that $x \geq \sqrt{3}p^{1/3}$. Then the sum of the squares is at least $3p^{2/3}$. Similarly if $y \geq \sqrt{3}p^{1/3}$ or $z \geq \sqrt{3}p^{1/3}$. On the other hand, the set

$$K = \{ (x, y, z) \in \mathbb{R}^3 \mid x \in [0, \sqrt{3}p^{1/3}], y \in [0, \sqrt{3}p^{1/3}], z \in [0, \sqrt{3}p^{1/3}], xyz = p \},$$

is closed and bounded, so that f achieves it minimum on this set, which we have already decided is at

$$(x, y, z) = (p^{1/3}, p^{1/3}, p^{1/3}),$$

since f is larger on the boundary. Putting all of this together, the point

$$(x, y, z) = (p^{1/3}, p^{1/3}, p^{1/3}),$$

is a point where the sum of the squares is a minimum.

Here is another such problem. Find the closest point to the origin which also belongs to the cone

$$x^2 + y^2 = z^2,$$

and to the plane

$$x + y + z = 3.$$

As before, we minimise $f(x, y, z) = x^2 + y^2 + z^2$ subject to $g_1(x, y, z) = x^2 + y^2 - z^2 = 0$ and $g_2(x, y, z) = x + y + z = 3$. Introduce a new function, with two new variables λ_1 and λ_2,

$$h\colon \mathbb{R}^5 \longrightarrow \mathbb{R},$$

given by

$$\begin{aligned} h(x, y, z, \lambda_1, \lambda_2) &= f(x, y, z) - \lambda_1 g_1(x, y, z) - \lambda_2 g_2(x, y, z) \\ &= x^2 + y^2 + z^2 - \lambda_1(x^2 + y^2 - z^2) - \lambda_2(x + y + z - 3). \end{aligned}$$

We find the critical points of h:

$$\begin{aligned} 2x &= 2\lambda_1 x + \lambda_2 \\ 2y &= 2\lambda_1 y + \lambda_2 \\ 2z &= -2\lambda_1 z + \lambda_2 \\ z^2 &= x^2 + y^2 \\ 3 &= x + y + z. \end{aligned}$$

Suppose we substract the first equation from the second:

$$y - x = \lambda_1(y - x).$$

So either $x = y$ or $\lambda_1 = 1$. Suppose $x \neq y$. Then $\lambda_1 = 1$ and $\lambda_2 = 0$. In this case $z = -z$, so that $z = 0$. But then $x^2 + y^2 = 0$ and so $x = y = 0$, which is not possible.

It follows that $x = y$, in which case $z = \pm\sqrt{2}x$ and

$$(2 \pm \sqrt{2})x = 3.$$

So

$$x = \frac{3}{2 \pm \sqrt{2}} = \frac{3(2 \mp \sqrt{2})}{2}.$$

This gives us two critical points:

$$\begin{aligned} P_1 &= (\frac{3(2 - \sqrt{2})}{2}, \frac{3(2 - \sqrt{2})}{2}, \frac{3\sqrt{2}(2 - \sqrt{2})}{2}) \\ P_2 &= (\frac{3(2 + \sqrt{2})}{2}, \frac{3(2 + \sqrt{2})}{2}, -\frac{3\sqrt{2}(2 - \sqrt{2})}{2}). \end{aligned}$$

Of the two, clearly the first is closest to the origin.

To finish, we had better show that this point is the closest to the origin on the whole locus

$$F = \{ (x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 = z^2, x + y + z = 3 \}.$$

Let

$$K = \{ (x, y, z) \in F \mid x^2 + y^2 + z^2 \leq 25 \}.$$

Then K is closed and bounded, whence compact. So f achieves its minimum somewhere on K, and so it must achieve its minimum at $P = P_1$. Clearly outside f is at least 25 on $F \setminus K$, and so f is a minimum at P_1 on the whole of F.