OPTIMIZATION ON SCALAR FIELDS

Maximum and minimum values

Def: Let $f: A \to \mathbb{R}$ with $A \subseteq \mathbb{R}^n$ be a scalar field. Let $p \in A$ be a point in A. We say that

a) p maximum of $f \iff \forall x \in A : f(x) \leq f(p)$

b) p minimum of $f \iff \forall x \in A : f(x) \geq f(p)$

c) p local maximum of $f \iff$
$\iff \exists a \in (0, +\infty) : \forall x \in B(p, a) \cap A : f(x) \leq f(p)$

d) p local minimum of $f \iff$
$\iff \exists a \in (0, +\infty) : \forall x \in B(p, a) \cap A : f(x) \geq f(p)$

- If p is a local minimum or local maximum, we say that p is a local extremum.

Generalized Fermat theorem

The following result generalizes the Fermat theorem from Calculus 1.

Thm: Let $f: A \to \mathbb{R}$ with $A \subseteq \mathbb{R}^n$ be a scalar field. Then

$$\left.\begin{array}{l} p \in \operatorname{int}(A) \\ f \text{ differentiable on } p \\ p \text{ local extremum of } f \end{array}\right\} \Rightarrow \nabla f(p) = \mathbf{0}$$

Proof

With no loss of generality, assume that ~~$p \in \operatorname{int}(A)$,~~
p is a local maximum of f. Then,

$p \in \operatorname{int}(A) \Rightarrow \exists a \in (0, +\infty): B(p, a) \subseteq A$

Choose $a \in (0, +\infty)$ such that $B(p, a) \subseteq A$.

Furthermore:

p local maximum of $\Rightarrow$

$$\Rightarrow \exists b \in (0, +\infty): \forall x \in B(p, b) \cap A: f(x) \leq f(p).$$

Choose $b \in (0, +\infty)$ such that $\forall x \in B(p, b) \cap A: f(x) \leq f(p)$.

Let $\rho = \min\{a, b\}$ and define:

$\forall k \in [n]: \forall t \in (-\rho, \rho): g_k(t) = f(p + t e_k)$

with $e_1, e_2, \ldots, e_n$ the unit vectors of the coordinate system.

It follows that

$g_k(t) = f(p + t e_k) \leq f(p) = g(0), \ \forall t \in (-\rho, \rho) \Rightarrow$

$\Rightarrow t = 0$ local max of g_k (1)

We also have:

f differentiable on $p \Rightarrow \forall k \in [n]: g_k$ differentiable on 0. (2)

$0 \in (-\rho, \rho)$ (3)

From Eq. (1), (2), (3) via the Fermat theorem:

$\forall k \in [n]: g_k'(0) = 0 \Rightarrow$

$\Rightarrow \forall k \in [n]: \partial f(p) / \partial x_k = f'(p \mid e_k) = g_k'(0) = 0$

$\Rightarrow \nabla f(p) = \mathbf{0}$ $\square$

• It follows from the generalized Fermat theorem that local extrema $p \in A$ can occur only at points where at least one of the following conditions is satisfied:

a) $\nabla f(p) = 0$

b) f not differentiable at p

c) $p \in \partial A$ (p is on the boundary of A)

• If $p \in A$ satisfies one of these conditions, we say that p is a <u>stationary point</u> (or <u>critical point</u>) of f.

• To find the local min/max of a function; we

a) Find all stationary points

b) Use classification theorems to see if the point is a min or max.

Failure of 1st derivative test

We will now show that the 1st derivative test cannot be generalized to functions with 2 or more variables. This is bad news, since the 1st derivative test is the best method for functions of 1 variable.

Consider the following conjecture:

<u>Conjecture:</u> Let $f : A \to \mathbb{R}$ with $A \subseteq \mathbb{R}^n$ and let $p \in \operatorname{int} A$ with $\nabla f(p) = \mathbf{0}$. Then:

a) p local max $\Leftrightarrow \forall u \in \mathbb{R}^n : \exists a > 0 : g(t) = f(p + tu) \searrow (0, a)$

b) p local min $\Leftrightarrow \forall u \in \mathbb{R}^n : \exists a > 0 : g(t) = f(p + tu) \nearrow (0, a)$

Intuitively we expect this conjecture to be true. However, we will now exhibit a <u>counter example</u>: disproving the conjecture.

COUNTEREXAMPLE

Let $f(x,y) = 3x^4 - 4x^2y + y^2, \forall (x,y) \in \mathbb{R}^2$. Then $(0,0)$ is a stationary point and satisfies the conditions of the conjecture but it is NOT a local max or min of f.

<u>Proof</u>

▸ We show that $(0,0)$ is a stationary point.

$$\left.\begin{array}{l} \partial f/\partial x = (\partial/\partial x)(3x^4 - 4x^2y + y^2) = 12x^3 - 8xy \\ \partial f/\partial y = (\partial/\partial y)(3x^4 - 4x^2y + y^2) = -4x^2 + 2y \end{array}\right\} \Rightarrow$$

$$\Rightarrow \left.\frac{\partial f}{\partial x}\right|_{(0,0)} = 0 \wedge \left.\frac{\partial f}{\partial y}\right|_{(0,0)} = 0 \Rightarrow \nabla f(0,0) = \mathbf{0}$$

$\Rightarrow (0,0)$ is a stationary point.

▸ We show that the conditions of conjecture are satisfied.

Let $u = (a,b) \in \mathbb{R}^2 - \{\mathbf{0}\}$ with $a \neq 0 \vee b \neq 0$ be given.

We define:

$$\begin{aligned} g(t) &= f((0,0) + tu) = f((0,0) + t(a,b)) = f(ta, tb) = \\ &= 3(at)^4 - 4(at)^2(bt) + (bt)^2 = \\ &= 3a^4t^4 - 4a^2bt^3 + b^2t^2, \forall t \in [0, +\infty) \Rightarrow \end{aligned}$$

$$\Rightarrow g'(t) = 12a^4t^3 - 12a^2bt^2 + 2b^2t =$$

$= 2t\,(6a^4t^2 - 6a^2bt + b^2)$

$= 2t\,\varphi(t|a,b)$, $\forall t\in[0,+\infty)$

with $\varphi(t|a,b) = 6a^4t^2 - 6a^2bt + b^2$, $\forall t\in[0,+\infty)$.

We distinguish between the following cases.

Case 1: Assume that $b\neq 0$. Then since $\varphi(0|a,b) = b^2 > 0$, it follows that there is some interval $(0,t_0)$ such that

$(\forall t\in(0,t_0): \varphi(t|a,b) > 0) \Rightarrow$

$\Rightarrow \forall t\in(0,t_0): g'(t) = 2t\,\varphi(t|a,b) > 0$

$\Rightarrow g(t) = f((0,0)+tu) \nearrow (0,t_0)$.

Case 2: Assume that $b=0$. Then

$g'(t) = 2t\,\varphi(t|a,0) = 2t\,(6a^4t^2 - 0 + 0) =$

$= 12a^4t^3 \geq 0$, $\forall t\in(0,+\infty) \Rightarrow$

$\Rightarrow g(t) = f((0,0)+tu) \nearrow (0,+\infty)$

We conclude that in both cases the conditions of the conjecture are satisfied, and the conjecture would imply that $(0,0)$ is a local minimum.

► We show that $(0,0)$ is not a local minimum. First, we note that $f(0,0)=0$. Then we consider the values of f along the curve (c): $y=2x^2$. We see that

$f(x,2x^2) = 3x^4 - 4x^2(2x^2) + (2x^2)^2 =$

$= 3x^4 - 8x^4 + 4x^4 = -x^4 < 0$, $\forall x\in(0,+\infty)$

$\Rightarrow \forall r>0 : \exists x\in B((0,0),r) : f(x) < 0 = f(\mathbf{0})$

$\Rightarrow \mathbf{0}$ not local min of f !

Second derivative test

Let $f: A \to \mathbb{R}$ with $A \subseteq \mathbb{R}^2$ and let $p \in \operatorname{int}(A)$ with $\nabla f(p) = \mathbf{0}$. Define

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

Then:

a) $D > 0 \wedge f_{xx}(p) > 0 \Rightarrow p$ local minimum

b) $D > 0 \wedge f_{xx}(p) < 0 \Rightarrow p$ local maximum

c) $D < 0 \Rightarrow p$ saddle point

d) $D = 0 \Rightarrow$ inconclusive.

EXAMPLES

a) Classify all stationary points of $f(x,y) = x^3 + y^3 - 3xy$.

Solution

$$f_x(x,y) = (\partial/\partial x)(x^3+y^3-3xy) = 3x^2 - 3y$$
$$f_y(x,y) = (\partial/\partial y)(x^3+y^3-3xy) = 3y^2 - 3x$$

(x,y) stationary point $\Leftrightarrow \nabla f(x,y) = \mathbf{0} \Leftrightarrow \begin{cases} f_x(x,y) = 0 \\ f_y(x,y) = 0 \end{cases}$

$$\Leftrightarrow \begin{cases} 3x^2 - 3y = 0 \\ 3y^2 - 3x = 0 \end{cases} \Leftrightarrow \begin{cases} x^2 - y = 0 \\ y^2 - x = 0 \end{cases} \Leftrightarrow \begin{cases} y = x^2 \\ x^4 - x = 0 \end{cases} \Leftrightarrow$$

$$\Leftrightarrow \begin{cases} y = x^2 \\ x(x-1)(x^2+x+1) = 0 \end{cases} \Leftrightarrow \begin{cases} y = x^2 \\ x = 0 \end{cases} \vee \begin{cases} y = x^2 \\ x = 1 \end{cases} \Leftrightarrow$$

$$\Leftrightarrow \begin{cases} x=0 \\ y=0 \end{cases} \vee \begin{cases} x=1 \\ y=1 \end{cases}$$

Thus $(0,0)$ and $(1,1)$ are stationary points.

Evaluate:

$f_{xx}(x,y) = (\partial/\partial x)(3x^2-3y) = 6x$

$f_{xy}(x,y) = (\partial/\partial x)(3y^2-3x) = -3$

$f_{yy}(x,y) = (\partial/\partial y)(3y^2-3x) = 6y$

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

$= (6x)(6y) - (-3)^2 = 36xy - 9 = 9(4xy-1)$

For $(x,y) = (0,0)$:

$f_{xx}(0,0) = 0$

$D(0,0) = 9(4\cdot 0 - 1) = -9 < 0 \Rightarrow (0,0)$ is a saddle point.

For $(x,y) = (1,1)$:

$$\left.\begin{array}{l} D(1,1) = 9(4\cdot 1\cdot 1 - 1) = 9\cdot 3 = 27 > 0 \\ f_{xx}(1,1) = 6\cdot 1 = 6 > 0 \end{array}\right\} \Rightarrow (1,1) \text{ is a local minimum.}$$

↳ If $D=0$, then the 2nd derivative test is inconclusive. In that case, if the stationary point is a saddle point, we could be able to identify it as such using the following 1st derivative test. Note however that for local min or max, the 1st derivative test does not work as we have shown previously.

1st derivative test for saddle points

Let $f: A \to \mathbb{R}$ with $A \subseteq \mathbb{R}^n$ and let $p \in \text{int}(A)$ with $\nabla f(p) = 0$. Then:

$$\left[\exists a, b \in \mathbb{R}^2 : \exists t_1, t_2 \in (0, +\infty) : \begin{cases} g_1(t) = f(p+at) \nearrow (0, t_1) \\ g_2(t) = f(p+bt) \searrow (0, t_2) \end{cases}\right] \Rightarrow$$

$\Rightarrow$ p saddle point of f.

EXAMPLE

Show that $f(x,y) = x^4 y^7$ has a saddle point at $(0,0)$.

Solution

$$\left.\begin{aligned} f_x(x,y) &= (\partial/\partial x)(x^4y^7) = 4x^3y^7 \\ f_y(x,y) &= (\partial/\partial y)(x^4y^7) = 7x^4y^6 \end{aligned}\right\} \Rightarrow$$

$$\Rightarrow \nabla f(0,0) = (f_x(0,0), f_y(0,0)) =$$

$$= (4 \cdot 0 \cdot 0, 7 \cdot 00) = (0,0) = \mathbf{0} \Rightarrow$$

$\Rightarrow$ $(0,0)$ is a stationary point.

- Let $a, b \in \mathbb{R} - \{0\}$ be given. Define:

$$g(t) = f((0,0) + t(a,b)) = f(at, bt) = (at)^4 (bt)^7 =$$

$$= a^4 b^7 t^{11}, \forall t \in (0, +\infty) \Rightarrow$$

$$\Rightarrow g'(t) = 11 a^4 b^7 t^{10}, \forall t \in (0, +\infty).$$

For $a = b = 1$: $g'(t) = 11 t^{10}, \forall t \in (0, +\infty) \Rightarrow$

$$\Rightarrow g'(t) > 0, \forall t \in (0, +\infty) \Rightarrow g \nearrow (0, +\infty).$$

For $a=1$ and $b=-1$: $g'(t) = 11 \cdot 1^4 \cdot (-1)^7 t^{10} = -11t^{10}, \forall t \in (0,+\infty)$
$\Rightarrow g'(t) < 0, \forall t \in (0,+\infty) \Rightarrow g \searrow (0,+\infty)$
It follows that $(0,0)$ is a saddle point.

$\mapsto$ To see why the second derivative test fails, we note that:

$$f_{xx}(x,y) = (\partial/\partial x)(4x^3y^7) = 12x^2y^7$$
$$f_{xy}(x,y) = (\partial/\partial x)(7x^4y^6) = 28x^3y^6$$
$$f_{yy}(x,y) = (\partial/\partial y)(7x^4y^6) = 42x^4y^5$$

and therefore:

$$D(x,y) = f_{xx}(x,y)f_{yy}(x,y) - [f_{xy}(x,y)]^2 =$$
$$= (12x^2y^7)(42x^4y^5) - (28x^3y^6)^2 =$$
$$= 504x^6y^{12} - 784x^6y^{12} = -280x^6y^{12} \Rightarrow$$
$$\Rightarrow D(0,0) = -280 \cdot 0^6 \cdot 0^{12} = 0 \Rightarrow$$
$\Rightarrow$ 2nd derivative test inconclusive with respect to $(0,0)$.

EXERCISES

① Find all the local min and max for the following functions

a) $f(x,y) = x^2 + 3y^2$

b) $f(x,y) = 2x^2 - 5y^2$

c) $f(x,y) = x^2 + 2x + y^2 - 4y$

d) $f(x,y) = x^2 - y^2 + 8(x+y)$

e) $f(x,y) = x^2 + 2y^2 + xy - 2x - 4y$

f) $f(x,y) = 3x^2 + 2y^2 - 2xy + x - 3y$

g) $f(x,y) = x^3 - 3xy^2 + y^3$

h) $f(x,y) = x^3 + y^3 - 3a^3xy$

i) $f(x,y) = \sin x \sin y \sin(x+y)$

② Similarly for the functions

a) $f(x,y) = (x-1)^3(y-2)^4$

b) $f(x,y) = x^2y^2(x-1)^5$

③ Let $f(x,y) = ax^2 + 2bxy + cy^2 + 2dx + 2ey + f$ with $a>0$ and $b^2 < ac$. Show that

a) f has a local minimum at some point (x_1, y_1)

b) $f(x_1, y_1) = dx_1 + ey_1 + f$.

Constrained Optimization

• The problem: Let $f: A \to \mathbb{R}$ and $g_k: A \to \mathbb{R}$ for $k \in [m]$ with $A \subseteq \mathbb{R}^n$. The problem is to identify the local min/max of

$$z = f(x),\ \forall x \in A$$

under the constraints

$$\forall k \in [m]: g_k(x) = 0$$

Lagrange multiplier theorem

Def: Let $u_1, u_2, \dots, u_m \in \mathbb{R}^n$ be m vectors. We say that

$u_1, u_2, \dots, u_m$ are linearly independent $\iff \forall \lambda_1, \lambda_2, \dots, \lambda_m \in \mathbb{R}: (\lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_m u_m = \mathbf{0} \Rightarrow$
$\Rightarrow \lambda_1 = \lambda_2 = \dots = \lambda_m = 0)$

Thm: Consider the constrained optimization problem

$$\begin{cases} z = f(x),\ \forall x \in A \\ \forall k \in [m]: g_k(x) = 0 \end{cases}$$

with $f: A \to \mathbb{R}$ and $g_k: A \to \mathbb{R}$.

Assume that:

a) $f, g_1, g_2, \dots, g_m$ are differentiable at $p \in A$.

b) $\forall k \in [m]: g_k(p) = 0$

c) $\nabla g_1(p), \nabla g_2(p), \dots, \nabla g_m(p)$ are linearly independent

d) p is a local min or max of the constrained problem.

Then:

$$\exists \lambda_1, \lambda_2, \ldots, \lambda_m \in \mathbb{R}: \nabla f(p) = \sum_{k=1}^{m} \lambda_k \nabla g_k(p)$$

↳ The Lagrange multipliers method gives the stationary points of the constrained optimization problem but does not establish whether these points are local min or max. Doing so is very difficult, but there are occasional tricks we can try.

▸ <u>When constraints yield a finite curve.</u>

- Recall from single variable calculus that for any function $f: A \to \mathbb{R}$ with $A \subseteq \mathbb{R}$; and $a, b \in A$:

  f continuous in $[a, b] \Rightarrow f$ bounded in $[a, b]$.
- Let S be the set defined by the problem constraints:

  $S = \{ x \in \mathbb{R}^n \mid \forall k \in [m] : g_k(x) = 0 \}$

<u>Def</u>: We say that S is a finite curve if we can define a parameterization $\gamma: [0,1] \longrightarrow S$ such that $\gamma([0,1]) = S$ and γ continuous in $[0,1]$.

Now we can parameterize the function f on S by defining $g(t) = f(\gamma(t))$. Then, it follows that:

$$\left.\begin{array}{l} f \text{ continuous in } A \Rightarrow f \text{ continuous in } \mathcal{S} \\ \qquad\qquad\qquad\quad\; \gamma \text{ continuous in } [0,1] \end{array}\right\} \Rightarrow$$

$$\Rightarrow g \text{ continuous in } [0,1] \Rightarrow$$

$$\Rightarrow \exists\, t_1, t_2 \in [0,1] : \forall t \in [0,1] : g(t_1) \leq g(t) \leq g(t_2).$$

The points $\gamma(t_1), \gamma(t_2)$ will show up as stationary points of the Lagrange multiplier method solutions and are correspondingly the minimum and maximum of f under the given constraints.

- <u>To conclude:</u> If $\mathcal{S}$ is a finite curve, then the constrained optimization has a maximum and a minimum among the existing stationary points. To identify the minimum and maximum, we simply evaluate the function for all stationary points, and choose the stationary points that give the minimum and maximum values.

EXAMPLE

Find the minimum and maximum value of
$f(x,y,z) = x+2y+3z$
under the constraints

$$\begin{cases} x^2+y^2=2 \\ y+z=1 \end{cases}$$

Solution

Define
$$g_1(x,y,z) = 2-x^2-y^2$$
$$g_2(x,y,z) = y+z-1$$

Note that:

$\nabla f(x,y,z) = (1,2,3)$

$\nabla g_1(x,y,z) = (-2x,-2y,0)$

$\nabla g_2(x,y,z) = (0,1,1)$

▶ Linear independence

$$\nabla g_1(x,y,z) \times \nabla g_2(x,y,z) = (-2x,-2y,0) \times (0,1,1) =$$

$$= \begin{vmatrix} i & j & k \\ -2x & -2y & 0 \\ 0 & 1 & 1 \end{vmatrix} = \begin{vmatrix} i & j & -j+k \\ -2x & -2y & 2y \\ 0 & 1 & 0 \end{vmatrix} =$$

(-1)

$$= -1 \cdot \begin{vmatrix} i & -j+k \\ -2x & 2y \end{vmatrix} = -\left(2yi - (-2x)(-j+k)\right) =$$

$= -2x(-j+k) - 2yi$

therefore

$\nabla g_1(x,y,z), \nabla g_2(x,y,z)$ linearly independent $\Leftrightarrow$

$\Leftrightarrow \nabla g_1(x,y,z) \times \nabla g_2(x,y,z) \neq \mathbf{0} \Leftrightarrow -2x(-j+k) - 2yi \neq 0 \Leftrightarrow$

$\Leftrightarrow \underline{\underline{x \neq 0 \vee y \neq 0}}$

↳ Recall that $\nabla g_1, \nabla g_2$ need to be linearly independent to apply the Lagrange multiplier theorem, when evaluated at the stationary points. For two vectors we use the criterio:

$$\boxed{\begin{array}{l} \text{For } u,v \in \mathbb{R}^3: \\ u,v \text{ linearly independent} \Leftrightarrow u \times v \neq \mathbf{0} \end{array}}$$

▶ <u>Stationary points</u>

(x,y,z) is a stationary point $\Leftrightarrow$

$$\Leftrightarrow \begin{cases} \nabla f(x,y,z) = \lambda_1 \nabla g_1(x,y,z) + \lambda_2 \nabla g_2(x,y,z) \\ x^2+y^2 = 2 \\ y+z = 1 \end{cases} \Leftrightarrow$$

$$\Leftrightarrow \begin{cases} (1,2,3) = \lambda_1(-2x,-2y,0) + \lambda_2(0,1,1) \\ x^2+y^2 = 2 \\ y+z = 1 \end{cases} \Leftrightarrow$$

$$\begin{cases} 1 = -2\lambda_1 x \\ 2 = -2\lambda_1 y + \lambda_2 \\ 3 = \lambda_2 \\ x^2+y^2 = 2 \\ y+z = 1 \end{cases} \Leftrightarrow \begin{cases} 1 = -2\lambda_1 x \\ 2 = -2\lambda_1 y + 3 \\ x^2+y^2 = 2 \\ y+z = 1 \end{cases} \Leftrightarrow \begin{cases} 2\lambda_1 x = -1 & (1) \\ 2\lambda_1 y = 1 & (2) \\ x^2+y^2 = 2 & (3) \\ y+z = 1 & (4) \end{cases}$$

If $\lambda_1 = 0$, then equations (1) and (2) are inconsistent.
So, assume $\lambda_1 \neq 0$. Then:

$(1) \Leftrightarrow x = \frac{-1}{2\lambda_1}$ and $(2) \Leftrightarrow y = \frac{1}{2\lambda_1}$ and $(4) \Leftrightarrow z = 1-y$

It follows that

$$(3) \Leftrightarrow x^2+y^2 = 2 \Leftrightarrow \left(\frac{-1}{2\lambda_1}\right)^2 + \left(\frac{1}{2\lambda_1}\right)^2 = 2 \Leftrightarrow$$

$$\Leftrightarrow 2\frac{1}{4\lambda_1^2} = 2 \Leftrightarrow 4\lambda_1^2 = 1 \Leftrightarrow 2\lambda_1 = 1 \vee 2\lambda_1 = -1 \Leftrightarrow$$

$$\Leftrightarrow \lambda_1 = \frac{1}{2} \vee \lambda_1 = \frac{-1}{2}.$$

Thus the system of equations (1), (2), (3), (4) is equivalent to:

$$\begin{cases} x = -1/(2\lambda_1) \\ y = 1/(2\lambda_1) \\ \lambda_1 = 1/2 \vee \lambda_1 = -1/2 \\ z = 1-y \end{cases} \Leftrightarrow$$

$$\begin{cases} x = (-1/2)\cdot 2 = -1 \\ y = (1/2)\cdot 2 = 1 \\ z = 1 - 1 = 0 \end{cases} \vee \begin{cases} x = (-1/2)(-2) = 1 \\ y = (+1/2)(-2) = -1 \\ z = 1-(-1) = 2 \end{cases} \Leftrightarrow$$

$$\Leftrightarrow (x,y,z) = (-1,1,0) \vee (x,y,z) = (1,-1,2)$$

Note that both stationary points satisfy the linear independence condition.

- The set

  $A = \{(x,y,z) \in \mathbb{R}^3 \mid x^2+y^2=2 \wedge y+z=1\}$

  is the intersection of a cylindrical surface and a plane, therefore it is a finite curve. Consequently:

  $f(-1,1,0) = (-1)+2\cdot 1+3\cdot 0 = -1+2 = 1$

  $f(1,-1,2) = (1)+2(-1)+3\cdot 2 = 1-2+6 = 5$

  thus $(-1,1,0)$ is the minimum and

  $(1,-1,2)$ is the maximum.

EXAMPLES

a) Find the stationary points of
$f(x,y,z) = xyz$ with $x,y,z \in (0,+\infty)$
under the constraint
$xy+yz+zx=1$.

Solution

Define $g(x,y,z) = xy+yz+zx-1$. Note that

$$f_x(x,y,z) = (\partial/\partial x)(xyz) = yz \quad (1)$$

$$f_y(x,y,z) = (\partial/\partial y)(xyz) = zx \quad (2)$$

$$f_z(x,y,z) = (\partial/\partial z)(xyz) = xy \quad (3)$$

From (1),(2),(3): $\nabla f(x,y,z) = (yz, zx, xy)$.

$$g_x(x,y,z) = (\partial/\partial x)(xy+yz+zx-1) = y+z \quad (4)$$

$$g_y(x,y,z) = (\partial/\partial y)(xy+yz+zx-1) = z+x \quad (5)$$

$$g_z(x,y,z) = (\partial/\partial z)(xy+yz+zx-1) = x+y. \quad (6)$$

From (4),(5),(6): $\nabla g(x,y,z) = (y+z, z+x, x+y)$.

It follows that:

(x,y,z) is a stationary point $\Leftrightarrow \begin{cases} \nabla f(x,y,z) = \lambda \nabla g(x,y,z) \\ g(x,y,z) = 0 \end{cases}$

$$\Leftrightarrow \begin{cases} (yz, zx, xy) = \lambda(y+z, z+x, x+y) \\ xy+yz+zx = 1 \end{cases} \Leftrightarrow$$

$$\Leftrightarrow \begin{cases} xy = \lambda(x+y) \\ yz = \lambda(y+z) \\ zx = \lambda(z+x) \\ xy+yz+zx = 1 \end{cases} \Leftrightarrow \begin{cases} xyz = \lambda z(x+y) & (7) \\ xyz = \lambda x(y+z) & (8) \\ xyz = \lambda y(z+x) & (9) \\ xy+yz+zx = 1 & (10). \end{cases}$$

From (7) and (8):

$$\lambda z(x+y) = \lambda x(y+z) \Leftrightarrow \lambda[x(y+z) - z(x+y)] = 0 \Leftrightarrow$$
$$\Leftrightarrow \lambda(xy+xz-zx-zy) = 0 \Leftrightarrow \lambda(xy-zy) = 0 \Leftrightarrow$$
$$\Leftrightarrow \lambda y(x-z) = 0 \Leftrightarrow \lambda = 0 \vee y = 0 \vee x-z = 0 \Leftrightarrow$$
$$\Leftrightarrow \lambda = 0 \vee y = 0 \vee x = z. \qquad (11)$$

We note that for $\lambda = 0$:

$$\begin{cases} xy = 0 \\ yz = 0 \\ zx = 0 \\ xy+yz+zx = 1 \end{cases} \Leftrightarrow \begin{cases} xy = 0 \\ yz = 0 \\ zx = 0 \\ 0+0+0 = 1 \leftarrow \text{inconsistent.} \end{cases}$$

therefore $\lambda \neq 0$.

We have also assumed $y > 0 \Rightarrow y \neq 0$.

It follows that $(11) \Leftrightarrow x = z$.

Since the system remains invariant under the transformation $x \to y \to z \to x$, we can similarly show that $y = x$ and $z = y$. Thus, the system defined by (7), (8), (9), (10) is equivalent to

$$\begin{cases} x=y \\ y=z \\ z=x \\ xy+yz+zx=1 \end{cases} \Leftrightarrow \begin{cases} x=y=z \\ x^2+x^2+x^2=1 \end{cases} \Leftrightarrow \begin{cases} x=y=z \\ 3x^2=1 \end{cases} \Leftrightarrow$$

$$\Leftrightarrow \begin{cases} x=y=z \\ x^2=1/3 \end{cases} \Leftrightarrow x=y=z=\frac{1}{\sqrt{3}} \quad \text{(because } x>0\text{)}.$$

Therefore : the stationary point is:

$$(x,y,z)=\left(\frac{1}{\sqrt{3}},\frac{1}{\sqrt{3}},\frac{1}{\sqrt{3}}\right)$$

EXERCISES

④ Use the method of Lagrange multipliers to find the minimum or maximum point of the following functions subject to the given constants.

a) $f(x,y) = 3x + 4y$ with $x^2 + y^2 = 9$

b) $f(x,y) = 2xy$ with $2x^2 + y^2 = 16$

c) $f(x,y) = x^2 + y^2$ with $3x + 2y = 10$

d) $f(x,y) = x^2y + x + y$ with $xy = 9$

e) $f(x,y) = x^2 + y^2$ with $x^4 + y^4 = 1$

f) $f(x,y) = x^2 y^2$ with $x^2 + y^2 = 4$

⑤ Similarly for the following functions (two constraints)

a) $f(x,y,z) = x + y + z$

with $\begin{cases} x^2 + y^2 + z^2 = 9 \\ x^2 + y^2 + 16z^2 = 36 \end{cases}$

b) $f(x,y,z) = x + 2y + z$

with $\begin{cases} 2x + z = 4 \\ x^2 + y^2 = 1 \end{cases}$

c) $f(x,y,z) = x^2 + y^2 + z^2$

with $\begin{cases} x + 2y + z = 3 \\ x - y = 4 \end{cases}$

⑥ Show that the point on the line $(\ell): ax+by=c$ closest to the origin has coordinates (x_0, y_0) with

$$x_0 = \frac{ac}{a^2+b^2} \wedge y_0 = \frac{bc}{a^2+b^2}$$

⑦ Find the maximum point of
$f(x,y) = x^a y^b$, $\forall (x,y) \in [0,+\infty) \times [0,+\infty)$
on the unit circle
$(c): x^2+y^2=1$
with $a, b \in (0,+\infty)$

⑧ Find the maximum point of
$f(x,y,z) = x^a y^b z^c$, $\forall (x,y,z) \in [0,+\infty) \times [0,+\infty) \times [0,+\infty)$
on the unit ~~circle~~ sphere
$(c): x^2+y^2+z^2=1$
with $a, b, c \in (0,+\infty)$.

⑨ Consider scalar fields $f: \mathbb{R}^2 \to \mathbb{R}$ and $g: \mathbb{R}^2 \to \mathbb{R}$ and assume that $f(x,y)$ has a maximum under the constraint $g(x,y)=a$ at the point $(x(a), y(a))$ which is a function of the <u>budget variable</u> $a \in \mathbb{R}$.
Show that:

a) $\forall a \in \mathbb{R}: \nabla g(x(a), y(a)) \cdot (x'(a), y'(a)) = 1$

b) The Lagrange multiplier $\lambda(a)$ corresponding to the point $(x(a), y(a))$ is given by:

$$\lambda(a) = \frac{d}{da} f(x(a), y(a))$$

↳ The problem gives a physical interpretation of the Lagrange multiplier $\lambda(a)$. It shows how fast the maximum value $f(x(a), y(a))$ of the scalar field f increases as we increase the budget value a. A similar result can be obtained for the most general constrained optimization problem.

⑩ Boltzmann distribution

Consider the entropy $S(x_1, x_2, \ldots, x_n)$ defined as

$$S(x_1, x_2, \ldots, x_n) = \sum_{a=1}^{n} x_a \ln x_a =$$

$$= x_1 \ln x_1 + x_2 \ln x_2 + \cdots + x_n \ln x_n$$

with $x_1, x_2, \ldots, x_n \in (0, +\infty)$ subject to the constraints

$$\begin{cases} \sum_{a=1}^{n} x_a = N \\ \sum_{a=1}^{n} E_a x_a = E \end{cases}$$

with $E_1, E_2, \ldots, E_n, E, A \in (0, +\infty)$. Show that the entropy is maximized when

$$\forall a \in [n] : x_a = \frac{\exp(\mu E_a)}{A}$$

with $A = (1/N) \sum_{a=1}^{n} \exp(\mu E_a)$

and find the value of the constant μ.

▼ Optimization problems on a bounded set

Def: Let $A \subseteq \mathbb{R}^n$. We say that
A bounded $\iff \exists p \in \mathbb{R}^n : \exists a \in (0, +\infty) : A \subseteq B(p, a)$

Thm: Let $f: A \to \mathbb{R}$ with $A \subseteq \mathbb{R}^n$. Assume that:
a) f continuous at A
b) A is closed (i.e. $\partial A \subseteq A$)
c) A is bounded.
Then:
$\exists p_1, p_2 \in A : \forall x \in A : f(p_1) \leq f(x) \leq f(p_2)$.

↳ This theorem is called the extremum value theorem and it generalizes the extremum value theorem of single-variable calculus.

▸ The problem: Let $f: \mathbb{R}^2 \to \mathbb{R}$ be a function and we want to optimize it under the constraint $g(x, y) \leq 0$. We assume that the set

$$S = \{(x, y) \in \mathbb{R}^2 \mid g(x, y) \leq 0\}$$

is closed and bounded.

EXAMPLE

Find the minimum and maximum of the scalar field
$f(x,y) = x^2+y^2+2$
under the constraint $x^2+y^2/4 \leq 1$.

Solution

Define $S = \{(x,y) \in \mathbb{R}^2 \mid x^2+y^2/4 \leq 1\}$.

- <u>Interior stationary points</u>

$\partial f/\partial x = (\partial/\partial x)(x^2+y^2+2) = 2x$
$\partial f/\partial y = (\partial/\partial y)(x^2+y^2+2) = 2y$
(x,y) stationary point on $\mathrm{int}(S) \Leftrightarrow \nabla f(x,y) = \mathbf{0} \Leftrightarrow$
$\Leftrightarrow \begin{cases} \partial f/\partial x = 0 \\ \partial f/\partial y = 0 \end{cases} \Leftrightarrow (2x,2y) = (0,0) \Leftrightarrow (x,y) = (0,0) \in \mathrm{int}(S)$
$\leftarrow$ accepted.

- <u>Boundary stationary points</u>

Note that $\partial S = \{(x,y) \in \mathbb{R}^2 \mid x^2+y^2/4 = 1\}$.
Define $g(x,y) = x^2+y^2/4-1$. Then
$\begin{cases} \partial g/\partial x = 2x \\ \partial g/\partial y = 2y/4 = y/2 \end{cases} \Rightarrow \nabla g(x,y) = (2x, y/2)$
Let $(x,y) \in \partial S$ be given. Then:
$(x,y) \in \partial S \Rightarrow x^2+y^2/4 = 1 \Rightarrow (x,y) \neq (0,0) \Rightarrow$
$\Rightarrow \nabla g(x,y) = (2x, y/2) \neq (0,0) \Rightarrow$
$\Rightarrow \nabla g(x,y)$ linearly independent.
and thus: $\forall (x,y) \in \partial S : \nabla g(x,y)$ linearly independent.
It follows that

$$(x,y) \text{ stationary point} \Leftrightarrow \begin{cases} \nabla f(x,y) = \lambda \nabla g(x,y) \\ g(x,y) = 0 \end{cases} \Leftrightarrow$$

$$\Leftrightarrow \begin{cases} (2x, 2y) = \lambda (2x, y/2) \\ x^2 + y^2/4 = 1 \end{cases} \Leftrightarrow \begin{cases} 2x = 2\lambda x \\ 2y = \lambda y/2 \\ 4x^2 + y^2 = 4 \end{cases} \Leftrightarrow$$

$$\Leftrightarrow \begin{cases} \lambda x - x = 0 \\ \lambda y - 4y = 0 \\ 4x^2 + y^2 = 4 \end{cases} \Leftrightarrow \begin{cases} (\lambda - 1)x = 0 \\ (\lambda - 4)y = 0 \\ 4x^2 + y^2 = 4 \end{cases} \Leftrightarrow$$

$$\Leftrightarrow \begin{cases} \lambda = 1 \\ (1-4)y = 0 \\ 4x^2 + y^2 = 4 \end{cases} \vee \begin{cases} x = 0 \\ (\lambda - 4)y = 0 \\ 0 + y^2 = 4 \end{cases} \Leftrightarrow$$

$$\Leftrightarrow \begin{cases} \lambda = 1 \\ y = 0 \\ 4x^2 = 4 \end{cases} \vee \begin{cases} x = 0 \\ \lambda = 4 \\ y^2 = 4 \end{cases} \Leftrightarrow$$

$$\Leftrightarrow \begin{cases} \lambda = 1 \\ y = 0 \\ x = 1 \end{cases} \vee \begin{cases} \lambda = 1 \\ y = 0 \\ x = -1 \end{cases} \vee \begin{cases} x = 0 \\ \lambda = 4 \\ y = 2 \end{cases} \vee \begin{cases} x = 0 \\ \lambda = 4 \\ y = -2 \end{cases} \Leftrightarrow$$

$$\Leftrightarrow (x,y) \in \{(1,0), (-1,0), (0,2), (0,-2)\}.$$

▸ <u>Point classification</u>

$f(0,0) = 2 + 0^2 + 0^2 = 2$

$f(1,0) = 2 + 1^2 + 0^2 = 3$

$f(-1,0) = 2 + (-1)^2 + 0^2 = 3$

$f(0,2) = 2 + 0^2 + 2^2 = 6$

$f(0,-2) = 2 + 0^2 + (-2)^2 = 6$

Since S is closed and bounded it follows that f has minimum at $(0,0)$ and maximum at $(0,2)$ and $(0,-2)$.

$\mapsto$ The boundary stationary points can also be found by parameterizing the boundary ∂S as follows.

<u>2nd method</u>: We parameterize ∂S with

$$\begin{cases} x(t) = \cos t \\ y(t) = 2\sin t \end{cases}, \quad \forall t \in [0, 2\pi)$$

It follows that:

$$g(t) \equiv f(x(t), y(t)) = f(\cos t, 2\sin t) =$$
$$= 2 + \cos^2 t + 4\sin^2 t = 2 + (\cos^2 t + \sin^2 t) + 3\sin^2 t$$
$$= 2 + 1 + 3\sin^2 t = 3 + 3\sin^2 t \Rightarrow$$
$$\Rightarrow g'(t) = 3(\sin^2 t)' = 3 \cdot 2\sin t \, (\sin t)' = 6 \sin t \cos t =$$
$$= 3(2\sin t \cos t) = 3\sin(2t)$$

Solve:

$$g'(t) = 0 \Leftrightarrow 3\sin(2t) = 0 \Leftrightarrow \sin(2t) = 0 \Leftrightarrow 2t = k\pi \Leftrightarrow$$
$$\Leftrightarrow t = \frac{k\pi}{2}$$

Since

$$t \in [0, 2\pi) \Leftrightarrow 0 \leq t < 2\pi \Leftrightarrow 0 \leq \frac{k\pi}{2} < 2\pi \Leftrightarrow$$
$$\Leftrightarrow 0 \leq k/2 < 2 \Leftrightarrow 0 \leq k < 4 \Leftrightarrow$$
$$\Leftrightarrow k \in \{0, 1, 2, 3\}$$

Thus we find 4 stationary points:

$t = 0$: $(x, y) = (\cos 0, 2\sin 0) = (1, 0)$

$t = \pi/2$: $(x, y) = (\cos(\pi/2), 2\sin(\pi/2)) = (0, 2)$

$t = \pi$: $(x, y) = (\cos \pi, 2\sin \pi) = (-1, 0)$

$t = 3\pi/2$: $(x, y) = (\cos(3\pi/2), 2\sin(3\pi/2)) = (0, -2)$

EXERCISES

(11) Find the minimum and maximum of the following functions constrained on the set S

a) $f(x,y) = 1+x+y$

on $S = \{(x,y) \in \mathbb{R}^2 \mid x^2+y^2 \leq 1\}$

b) $f(x,y) = x+y-2xy$

on $S = \{(x,y) \in \mathbb{R}^2 \mid x^2+y^2 \leq 4\}$

c) $f(x,y) = x^2+y^2+3x-xy$

on $S = \{(x,y) \in \mathbb{R}^2 \mid x^2+y^2 \leq 9\}$

d) $f(x,y) = \dfrac{x}{y^2+1}$ on $S = \{(x,y) \in \mathbb{R}^2 \mid x^2/4 + y^2/9 \leq 1\}$

e) $f(x,y) = (1+x+y)^2$ on $S = \{(x,y) \in \mathbb{R}^2 \mid x^2/4 + y^2/16 \leq 1\}$