§ 13 Extreme Values

Definition 13.1

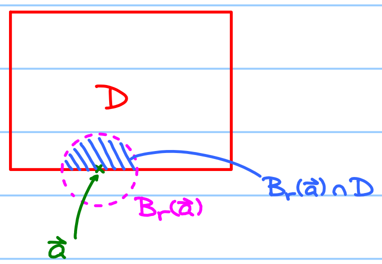
Let $f: D \to \mathbb{R}$ be a function where $D \subseteq \mathbb{R}^n$ and let $\vec{a} \in D$.

a) If $f(\vec{x}) \leq f(\vec{a})$ $(f(\vec{x}) \geq f(\vec{a}))$ for all $\vec{x} \in D$,

then f is said to have a global / an absolute maximum (minimum) at $\vec{a}$.

b) If there exists $r>0$ such that $f(\vec{x}) \leq f(\vec{a})$ $(f(\vec{x}) \geq f(\vec{a}))$ for all $\vec{x} \in B_r(\vec{a}) \cap D$,

then f is said to have a local maximum (minimum) at $\vec{a}$.



Definition 13.2

A point $\vec{a}$ is said to be a stationary point if $\nabla f(\vec{a}) = \vec{0}$.

Proposition 13.1

Let D be an open subset in $\mathbb{R}^n$, let $\vec{a} \in D$ and let $f: D \to \mathbb{R}$ be a differentiable function.

If f attains a local maximum or minimum at $\vec{a}$, then $\nabla f(\vec{a}) = \vec{0}$, i.e. $\vec{a}$ is a stationary point.

proof:

It is equivalent to show $\frac{\partial f}{\partial x_i}(\vec{a}) = 0$ for all $i = 1, 2, \cdots, n$.

Since f attains a local maximum or minimum at $\vec{a}$,

in particular f attains a local maximum or minimum at $\vec{a}$ along the x_i-direction.

Then, $\frac{\partial f}{\partial x_i}(\vec{a}) = 0$ follows from single variable calculus

Example 13.1

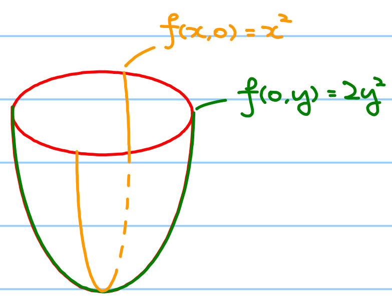
$z = f(x,y) = x^2 + 2y^2$

We have $\nabla f(x,y) = (2x, 4y)$.

$\nabla f(x,y) = 0$ if and only if $x = y = 0$.

Note that $f(x,y) = x^2 + 2y^2 \geq 0 = f(0,0)$,

so f has at local (in fact a global) minimum at $(0,0)$.



However, the converse of proposition 13.1 is NOT true

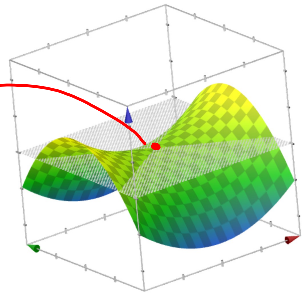
Example 13.2

$z = f(x,y) = x^2 - 2y^2$

We have $\nabla f(x,y) = (2x, -4y)$

$\nabla f(x,y) = 0$ if and only if $x = y = 0$.

(0,0) is a saddle point



However, f is not having a maximum or minimum at $(0,0)$.

Think: $f(x,0) = x^2 \geq 0 = f(0,0)$ while $f(0,y) = -2y^2 \leq 0 = f(0,0)$.

Definition 13.2

A stationary point is said to be a saddle point if it does not give rise to an extreme value

Second Derivative Check for Two Variables

Question: If $\frac{\partial f}{\partial x}(a_1,a_2) = \frac{\partial f}{\partial y}(a_1,a_2) = 0$, how to determine whether f has a maximum or minimum (or even a saddle point)?

Assume that $f(x,y)$ and its first and second partial derivatives are continuous on an open disk centered at $\vec{a} = (a_1,a_2)$. Then on that open disk, we have

$$f(x,y) \sim P_2(x,y) = f(\vec{a}) + \cancelto{0}{\nabla f(\vec{a})}\,(\vec{x}-\vec{a}) + \frac{1}{2}(\vec{x}-\vec{a})^T H(\vec{a})(\vec{x}-\vec{a})$$

$$= f(\vec{a}) + \frac{1}{2}(\vec{x}-\vec{a})^T H(\vec{a})(\vec{x}-\vec{a}) \qquad \text{where } H(\vec{x}) = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix}$$

Theorem 11.2 (Second Derivative Test)

Suppose that $f(x,y)$ and its first and second partial derivatives are continuous on an open disk centered at $\vec{a} = (a_1,a_2)$, and that $\frac{\partial f}{\partial x}(\vec{a}) = \frac{\partial f}{\partial y}(\vec{a}) = 0$. Then,

1) if $\det(H(\vec{a})) > 0$ and $\frac{\partial^2 f}{\partial x^2}(\vec{a}) < 0$, then f has a local maximum at $\vec{a}$.
2) if $\det(H(\vec{a})) > 0$ and $\frac{\partial^2 f}{\partial x^2}(\vec{a}) > 0$, then f has a local minimum at $\vec{a}$
3) if $\det(H(\vec{a})) < 0$, then f has a saddle point at $\vec{a}$.
4) if $\det(H(\vec{a})) = 0$, then there is NO conclusion.

Idea of proof:

Fact. $H(\vec{a})$ is a symmetric matrix, so there exists an orthogonal matrix $Q \in M_2(\mathbb{R})$, (i.e. $Q^TQ = Q^TQ = I$) such that

$$QH(\vec{a})Q^{-1} = D = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$$

where λ_1, λ_2 are eigenvalues of $H(\vec{a})$.

Also, note that $Q^{-1} = Q^T$, so $H(\vec{a}) = Q^{-1}DQ = Q^TDQ$

Then, $f(\vec{x}) \sim f(\vec{a}) + \frac{1}{2}(\vec{x}-\vec{a})^T H(\vec{a})(\vec{x}-\vec{a})$

$$= f(\vec{a}) + \frac{1}{2}(\vec{x}-\vec{a})^T Q^TDQ(\vec{x}-\vec{a})$$

$$= f(\vec{a}) + \frac{1}{2}\vec{u}^TD\vec{u} \qquad \text{where } \vec{u} = \begin{bmatrix} u \\ v \end{bmatrix} = Q(\vec{x}-\vec{a})$$

$$= f(\vec{a}) + \frac{1}{2}\lambda_1 u^2 + \frac{1}{2}\lambda_2 v^2$$

We perform a change of coordinates from (x,y) to (u,v), then we can see what we need to know is the signs of those eigenvalues

λ_1 and λ_2 are solution of the equation $\det(H(\vec{a}) - \lambda I) = 0$

$$\begin{vmatrix} \frac{\partial^2 f}{\partial x^2}(a_1,a_2) - \lambda & \frac{\partial^2 f}{\partial x \partial y}(a_1,a_2) \\ \frac{\partial^2 f}{\partial y \partial x}(a_1,a_2) & \frac{\partial^2 f}{\partial y^2}(a_1,a_2) - \lambda \end{vmatrix} = 0$$

$$\lambda^2 - \left(\frac{\partial^2 f}{\partial x^2}(a_1,a_2) + \frac{\partial^2 f}{\partial y^2}(a_1,a_2)\right)\lambda + \frac{\partial^2 f}{\partial x^2}(\vec{a})\frac{\partial^2 f}{\partial y^2}(\vec{a}) - \left(\frac{\partial^2 f}{\partial x \partial y}(\vec{a})\right)^2 = 0$$

$$\lambda^2 - \left(\frac{\partial^2 f}{\partial x^2}(a_1,a_2) + \frac{\partial^2 f}{\partial y^2}(a_1,a_2)\right)\lambda + \det(H(\vec{a})) = 0$$

$$\therefore \quad \lambda_1 + \lambda_2 = \frac{\partial^2 f}{\partial x^2}(a_1,a_2) + \frac{\partial^2 f}{\partial y^2}(a_1,a_2) \qquad \lambda_1\lambda_2 = \det(H(\vec{a}))$$

- If $\det(H(\vec{a})) = \lambda_1\lambda_2 < 0$, then λ_1 and λ_2 have opposite sign (saddle point)
- If $\det(H(\vec{a})) = \lambda_1\lambda_2 > 0$, then λ_1 and λ_2 have same sign

  Also, $\det(H(\vec{a})) = \frac{\partial^2 f}{\partial x^2}(\vec{a})\frac{\partial^2 f}{\partial y^2}(\vec{a}) - \left(\frac{\partial^2 f}{\partial x \partial y}(\vec{a})\right)^2 > 0 \Rightarrow \frac{\partial^2 f}{\partial x^2}(\vec{a})\frac{\partial^2 f}{\partial y^2}(\vec{a}) > 0$

  If $\frac{\partial^2 f}{\partial x^2}(\vec{a}) > 0$, then $\frac{\partial^2 f}{\partial y^2}(\vec{a}) > 0$ as well. Then $\lambda_1 + \lambda_2 = \frac{\partial^2 f}{\partial x^2}(a_1,a_2) + \frac{\partial^2 f}{\partial y^2}(a_1,a_2) > 0$

  $\therefore \lambda_1, \lambda_2 > 0$ (minimum)

  If $\frac{\partial^2 f}{\partial x^2}(\vec{a}) < 0$, then $\frac{\partial^2 f}{\partial y^2}(\vec{a}) < 0$ as well. Then $\lambda_1 + \lambda_2 = \frac{\partial^2 f}{\partial x^2}(a_1,a_2) + \frac{\partial^2 f}{\partial y^2}(a_1,a_2) < 0$

  $\therefore \lambda_1, \lambda_2 < 0$ (maximum)

- If $\det(H(\vec{a})) = \lambda_1\lambda_2 = 0$, then at least one of λ_1 and $\lambda_2 = 0$.

  We have to look at higher order terms (No conclusion)

Example 13.3

Let $f(x,y) = x^3+y^3-3xy$. Then, $\nabla f(x,y) = (3x^2-3y, 3y^2-3x)$

If $\nabla f(x,y) = 0$ $\begin{cases} 3x^2-3y=0 \\ 3y^2-3x=0 \end{cases} \Rightarrow \begin{cases} x^2 = y \\ y^2 = x \end{cases}$

$\therefore (x,y) = (0,0)$ or $(x,y) = (1,1)$

$$H(x,y) = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix} = \begin{bmatrix} 6x & -3 \\ -3 & 6y \end{bmatrix}$$

$\det(H(0,0)) = \begin{vmatrix} 0 & -3 \\ -3 & 0 \end{vmatrix} = -9 < 0$ $\therefore f$ has a saddle point at $(0,0)$.

$\det(H(1,1)) = \begin{vmatrix} 6 & -3 \\ -3 & 6 \end{vmatrix} = 27 > 0$ and $\frac{\partial^2 f}{\partial x^2}(1,1) = 6 > 0$ $\therefore f$ has a minimum point at $(1,1)$.

Further exercise :

$H(0,0) = \begin{bmatrix} 0 & -3 \\ -3 & 0 \end{bmatrix}$. Let $Q = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}$.

Show that $QQ^T = Q^TQ = I$ and $QH(0,0)Q^{-1} = \begin{bmatrix} -3 & 0 \\ 0 & 3 \end{bmatrix}$

Let $\vec{u} = \begin{bmatrix} u \\ v \end{bmatrix} = Q\vec{x} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$ i.e. $\vec{x} = \begin{bmatrix} x \\ y \end{bmatrix} = Q^{-1}\vec{u} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}}u + \frac{1}{\sqrt{2}}v \\ \frac{1}{\sqrt{2}}u - \frac{1}{\sqrt{2}}v \end{bmatrix}$

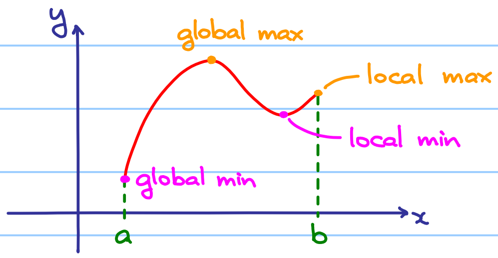
Then, by putting $x = \frac{1}{\sqrt{2}}u + \frac{1}{\sqrt{2}}v$ and $y = \frac{1}{\sqrt{2}}u - \frac{1}{\sqrt{2}}v$,

$$f(u,v) = \left(\tfrac{1}{\sqrt{2}}u + \tfrac{1}{\sqrt{2}}v\right)^3 + \left(\tfrac{1}{\sqrt{2}}u - \tfrac{1}{\sqrt{2}}v\right)^3 - 3\left(\tfrac{1}{\sqrt{2}}u + \tfrac{1}{\sqrt{2}}v\right)\left(\tfrac{1}{\sqrt{2}}u - \tfrac{1}{\sqrt{2}}v\right)$$

$$= -\tfrac{3}{2}u^2 + \tfrac{3}{2}v^2 + \tfrac{1}{\sqrt{2}}u^3 + \tfrac{3}{\sqrt{2}}uv^2$$

$\sim -\frac{3}{2}u^2 + \frac{3}{2}v^2$ around $(0,0)$ $\therefore f$ has a saddle point at $(0,0)$.

Recall: If $y = f(x)$ is defined on $[a,b]$

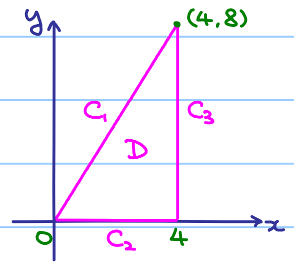


Note: To find global extreme values, we have to look at the boundary of the domain!

Example 13.4

Let $f(x,y) = xy - 2x - 3y$ defined on $D = \{(x,y) \in \mathbb{R}^2 : 0 \le x \le 4,\ 0 \le y \le 2x\}$.

Remark: D is compact in $\mathbb{R}^2$ and $f: D \to \mathbb{R}$ is continuous, so f attains absolute minimum and maximum at some points in D.



$f(x,y) = xy - 2x - 3y$ and so $\nabla f(x,y) = (y-2, x-3)$.

$\nabla f(x,y) = (0,0) \Rightarrow (x,y) = (2,3)$ Check: $(2,3) \in D$

$H(2,3) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \Rightarrow \det(H(2,3)) = -1 < 0$

$\therefore f$ has a saddle point at $(2,3)$.

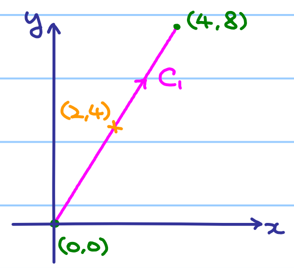
For the boundary,

C_1 is given by $y = 2x$

$C_1 : \gamma_1(t) = (x(t), y(t)) = (t, 2t)$ for $0 \le t \le 4$

$f_1(t) = f(\gamma_1(t)) = f(t, 2t) = 2t^2 - 8t$

$f_1'(t) = 4t - 8$, $f_1'(t) > 0$ if $t > 2$, $f_1'(t) < 0$ if $t < 2$,



When we go along C_1 from $(0,0)$ to $(4,8)$,

f is strictly increasing from $(0,0)$ to $(2,4)$;

f is strictly decreasing from $(2,4)$ to $(4,8)$;

$f(0,0) = 0$, $f(2,4) = -8$, $f(4,8) = 0$

$C_2 : \gamma_2(t) = (x(t), y(t)) = (t, 0)$ for $0 \le t \le 4$

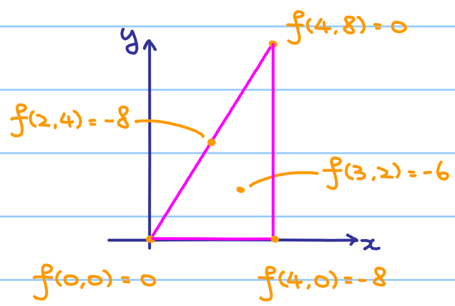
$f_2(t) = f(\gamma_2(t)) = f(t, 0) = -2t$, $f_2'(t) = -2 < 0$

When we go along C_2 from $(0,0)$ to $(0,4)$, f is strictly decreasing.

$C_3 : \gamma_3(t) = (x(t), y(t)) = (4, t)$ for $0 \le t \le 8$

$f_3(t) = f(\gamma_3(t)) = f(4, t) = t - 8$, $f_3'(t) = 1 > 0$

When we go along C_3 from $(4,0)$ to $(4,8)$, f is strictly increasing.



Global max of $f = 0$, it is attained at $(0,0)$ and $(4,8)$

Global max of $f = -8$, it is attained at $(4,0)$ and $(2,4)$

Example 13.5

Let $f(x,y) = (x-1)^2+(y-1)^2$ defined on $D = \{(x,y)\in\mathbb{R}^2 : x^2+y^2 \le 4\}$

D is the disk centered at $(0,0)$ with radius 2.

$\nabla f(x,y) = (2(x-1), 2(y-1))$, so $\nabla f(x,y) = 0$ if $(x,y) = (1,1)$.

$H(x,y) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$, so $\det(H(1,1)) = 4 > 0$ and $\frac{\partial^2 f}{\partial x^2}(1,1) = 2 > 0$

$\therefore$ f has a local minimum at $(1,1)$ and $f(1,1) = 0$

The boundary of D is the circle C given by $\gamma(t) = (2\cos t, 2\sin t)$ for $0 \le t < 2\pi$.

$f(t) = f(\gamma(t)) = (2\cos t - 1)^2 + (2\sin t - 1)^2$

$= 6 - 4\cos t - 4\sin t$

$f'(t) = 4\sin t - 4\cos t$

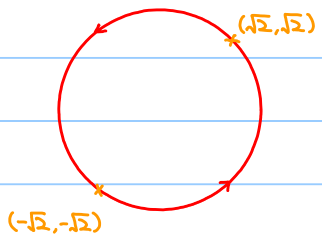
$f'(t) = 0$

$4\sin t - 4\cos t = 0$

$\tan t = 1$

$t = \frac{\pi}{4}, \frac{5\pi}{4}$

t	$0 \le t < \frac{\pi}{4}$	$\frac{\pi}{4} < t < \frac{5\pi}{4}$	$\frac{5\pi}{4} < t < 2\pi$
$f'(t)$	$-$	$+$	$-$



When we go along C,

f is strictly increasing from $(\sqrt{2},\sqrt{2})$ to $(-\sqrt{2},-\sqrt{2})$.

f is strictly decreasing from $(-\sqrt{2},-\sqrt{2})$ to $(\sqrt{2},\sqrt{2})$.

$f(\sqrt{2},\sqrt{2}) = 6-4\sqrt{2}$, $f(-\sqrt{2},-\sqrt{2}) = 6+4\sqrt{2}$

$\therefore$ f attains global max. at $(-\sqrt{2},-\sqrt{2})$ and $f(-\sqrt{2},-\sqrt{2}) = 6+4\sqrt{2}$

Comparing: $f(1,1) = 0 \le 6-4\sqrt{2} = f(\sqrt{2},\sqrt{2})$

$\therefore$ f attains global min. at $(1,1)$ and $f(1,1) = 0$

Second Derivative Check for General Cases

Let D be an open subset in $\mathbb{R}^n$, let $\vec{a} \in D$ and let $f: D \to \mathbb{R}$ be a C^2-function.

Suppose that $\nabla f(\vec{a}) = 0$, i.e. $\vec{a}$ is a stationary point, around the point $\vec{x} = \vec{a}$,

$$f(\vec{x}) \sim P_2(\vec{x}) = f(\vec{a}) + \cancelto{0}{\nabla f(\vec{a})}\,(\vec{x}-\vec{a}) + \frac{1}{2}(\vec{x}-\vec{a})^T H(\vec{a})(\vec{x}-\vec{a})$$

$$= f(\vec{a}) + \frac{1}{2}(\vec{x}-\vec{a})^T H(\vec{a})(\vec{x}-\vec{a})$$

$$\text{where } H(\vec{x}) = \left[\frac{\partial^2 f}{\partial x_i \partial x_j}\right]_{1 \le i,j \le n} = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix}$$

Fact. $H(\vec{a})$ is a symmetric matrix, so there exists an orthogonal matrix $Q \in M_n(\mathbb{R})$,

(i.e. $Q^TQ = Q^TQ = I$) such that

$$QH(\vec{a})Q^{-1} = D = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}$$

where $\lambda_1, \lambda_2, \cdots, \lambda_n$ are eigenvalues of $H(\vec{a})$.

Also, note that $Q^{-1} = Q^T$, so $H(\vec{a}) = Q^{-1}DQ = Q^TDQ$

Then,
$$\begin{aligned} f(\vec{x}) &\sim f(\vec{a}) + \frac{1}{2}(\vec{x}-\vec{a})^T H(\vec{a})(\vec{x}-\vec{a}) \\ &= f(\vec{a}) + \frac{1}{2}(\vec{x}-\vec{a})^T Q^T D Q(\vec{x}-\vec{a}) \\ &= f(\vec{a}) + \frac{1}{2}\vec{u}^T D \vec{u} \qquad \text{where } \vec{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} = Q(\vec{x}-\vec{a}) \\ &= f(\vec{a}) + \frac{1}{2}\sum_{i=1}^{n} \lambda_i u_i^2 \end{aligned}$$

Therefore,

- if $\lambda_i > 0$ for all $1 \le i \le n$, f attains a local minimum at $\vec{x} = \vec{a}$;
- if $\lambda_i < 0$ for all $1 \le i \le n$, f attains a local maximum at $\vec{x} = \vec{a}$;
- otherwise, $\vec{x} = \vec{a}$ is a saddle point.

Remark: if $\lambda_i > 0$ for all $1 \le i \le n$, $H(\vec{a})$ is said to be positive definite;
if $\lambda_i < 0$ for all $1 \le i \le n$, $H(\vec{a})$ is said to be negative definite.

With some knowledge in linear algebra, we have :

Proposition 13.2

Let $H_k(\vec{x}) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_k} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_k} \\ \vdots & \vdots & & \vdots \\ \frac{\partial^2 f}{\partial x_k \partial x_1} & \frac{\partial^2 f}{\partial x_k \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_k^2} \end{bmatrix}$ be $k \times k$-submatrix of $H(\vec{x})$.

1) $H(\vec{a})$ is positive definite $\Leftrightarrow \det(H_k(\vec{a})) > 0$ for all $1 \le k \le n$

2) $H(\vec{a})$ is negative definite $\Leftrightarrow \det(H_k(\vec{a})) \begin{cases} > 0 & \text{if } k \text{ is even} \\ < 0 & \text{if } k \text{ is odd} \end{cases}$

(or equivalently $(-1)^k \det(H_k(\vec{a})) > 0$ for all $1 \le k \le n$)

Think : When $n=2$,

- $H(\vec{a})$ is positive definite $\Leftrightarrow \frac{\partial^2 f}{\partial x^2}(\vec{a}) > 0$ and $\det(H(\vec{a})) > 0$
- $H(\vec{a})$ is negative definite $\Leftrightarrow \frac{\partial^2 f}{\partial x^2}(\vec{a}) < 0$ and $\det(H(\vec{a})) > 0$

It is just theorem 11.2 (Second Derivative Test)

§ 14 Lagrange Multipliers

Let $D \subseteq \mathbb{R}^n$ be an open subset and let $g: D \to \mathbb{R}$ be a smooth function.

Suppose that $L_c(g) = \{\vec{x} \in D : g(\vec{x}) = c\}$ and $\nabla g(\vec{x}) \neq \vec{0}$ for all points $\vec{x} \in L_c(g)$

Let $f: \mathbb{R}^n \to \mathbb{R}$ be a smooth function.

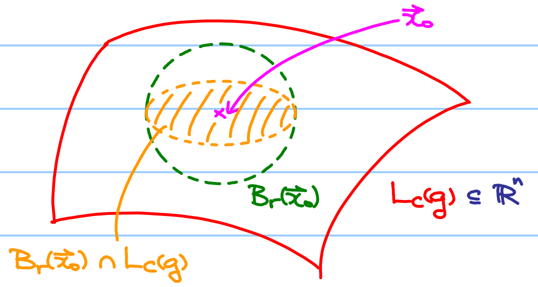
Question: How to find local extrema of f restricted on $L_c(g)$ (i.e. $f: L_c(g) \to \mathbb{R}$)?

Recall: Definition 13.1 gives

$f: L_c(g) \to \mathbb{R}$ attains local maximum (minimum) at $\vec{x}_0$ if

there exists open ball $B_r(\vec{x}_0)$ such that

$f(\vec{x}_0) \geq f(\vec{x})$ ($f(\vec{x}_0) \leq f(\vec{x})$) for all $\vec{x} \in B_r(\vec{x}_0) \cap L_c(g)$.



Remark:

Technique in the previous chapter may not work

If $f: L_c(g) \to \mathbb{R}$ attains local maximum at $\vec{x}_0$, it only means $f(\vec{x}_0) \geq f(\vec{x})$ for all $\vec{x} \in B_r(\vec{x}_0) \cap L_c(g)$,

but not $f(\vec{x}_0) \geq f(\vec{x})$ for all $\vec{x} \in B_r(\vec{x}_0)$, so we do not necessarily have $\nabla f(\vec{x}_0) = \vec{0}$

Theorem 14.1 (Lagrange Multiplier)

Suppose that $f: L_c(g) \to \mathbb{R}$ attains local extrema at $\vec{x}_0$, then $\nabla f(\vec{x}_0) = \lambda \nabla g(\vec{x}_0)$ for some $\lambda \in \mathbb{R}$,

i.e. $\nabla f(\vec{x}_0) \parallel \nabla g(\vec{x}_0)$ at local extreme points.

(Caution: The converse may not be true, i.e. even $\nabla f(\vec{x}_0) = \lambda \nabla g(\vec{x}_0)$ for some $\vec{x}_0 \in L_c(g)$, $\lambda \in \mathbb{R}$,

$f(\vec{x}_0)$ may not be a local extrema.)

proof:

Let $\gamma: (-\varepsilon, \varepsilon) \to D$ be a differentiable curve such that

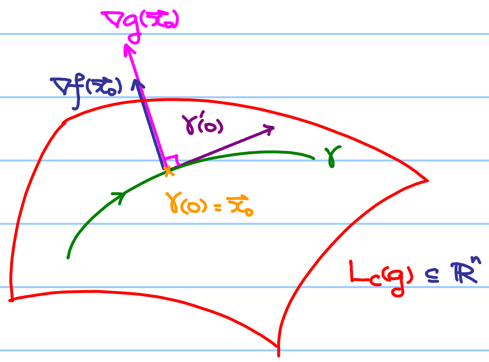
γ lies on $L_c(f)$, $\gamma(0) = \vec{x}_0$ and $\gamma'(0)$ is nonzero.

Then $f(\gamma(t))$ attains local extrema at $t = 0$, i.e.

$$\left.\frac{d}{dt}\right|_{t=0} f(\gamma(t)) = 0$$

$$\nabla f(\gamma(0)) \cdot \gamma'(0) = 0$$

$$\nabla f(\vec{x}_0) \cdot \gamma'(0) = 0$$



Therefore, $\nabla f(\vec{x}_0)$ is orthogonal to every vector tangent to $L_c(g)$ at $\vec{x}_0$,

and $\nabla f(\vec{x}_0)$ gives a normal of $L_c(g)$ at $\vec{x}_0$.

$\therefore$ $\nabla f(\vec{x}_0) \parallel \nabla g(\vec{x}_0)$, i.e. $\nabla f(\vec{x}_0) = \lambda \nabla g(\vec{x}_0)$ for some $\lambda \in \mathbb{R}$.

Example 14.1

Let $f:\mathbb{R}^2\to\mathbb{R}$ be a function defined by $f(x,y)=x^2y$.

Find the global extrema of f restricted on the unit circle $\mathcal{C}$ centered at the origin.

Remark: $\mathcal{C}$ is compact in $\mathbb{R}^2$.

Let $g(x,y)=x^2+y^2-1$. Then, $\mathcal{C}=L_0(g)=\{(x,y)\in\mathbb{R}^2 : x^2+y^2-1=0\}$

$\nabla f(x,y)=(2xy,x^2)$, $\nabla g(x,y)=(2x,2y)$

$$\begin{cases}\nabla f(x,y)=\lambda\nabla g(x,y)\\ x^2+y^2=1\end{cases}\Rightarrow\begin{cases}2xy=2x\lambda & -(1)\\ x^2=2y\lambda & -(2)\\ x^2+y^2=1 & -(3)\end{cases}$$

From (1), we have $y=\lambda$ or $x=0$.

- If $y=\lambda$, from (2), $x^2=2\lambda^2$. Put $y=\lambda$ and $x^2=2\lambda^2$ into (3), we have $3\lambda^2=1$, so $\lambda=\pm\frac{1}{\sqrt{3}}$.

  $\therefore$ f may have extreme values at $(\pm\sqrt{\frac{2}{3}},\frac{1}{\sqrt{3}})$, $(\pm\sqrt{\frac{2}{3}},-\frac{1}{\sqrt{3}})$.

- If $x=0$, then $y=\pm1$.

  $\therefore$ f may have extreme values at $(0,\pm1)$.

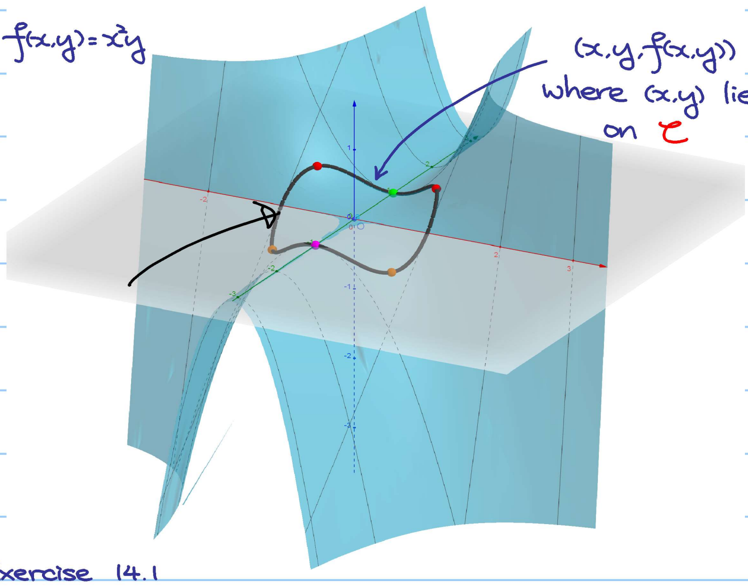
Abs max/min must be attained at these pts

$f(\pm\sqrt{\frac{2}{3}},\frac{1}{\sqrt{3}})=\frac{2}{3\sqrt{3}}$ global max

$f(\pm\sqrt{\frac{2}{3}},-\frac{1}{\sqrt{3}})=-\frac{2}{3\sqrt{3}}$ global min

$f(0,1)=0$??

$f(0,-1)=0$??



It may be hard to see $f(0,1)$ is a local minimum and $f(0,-1)$ is a local maximum unless we plot a graph.

Exercise 14.1

The unit circle $\mathcal{C}$ centered at the origin can be parametrized as

$\gamma(t)=(\cos t,\sin t)$, $t\in[0,2\pi]$.

Find the global extrema of f restricted on $\mathcal{C}$ is just finding global extrema of $f(\gamma(t))=f(\cos t,\sin t)=\cos^2 t\sin t$ which is just a single variable calculus problem.

Example 14.2

Find the point(s) on the surface $z^2 = xy+4$ closest to the origin

Let $g:\mathbb{R}^3 \to \mathbb{R}$ defined by $g(x,y,z) = xy+4-z^2$. Then $L_0(g)$ is the given surface.

The distance between a point (x,y,z) and the origin is $\sqrt{x^2+y^2+z^2}$.

Therefore, we have to minimize $\sqrt{x^2+y^2+z^2}$ on the surface $L_0(g)$.

However, it is also equivalent to minimize $f(x,y,z) = x^2+y^2+z^2$

(Technical point: The equation $z^2=xy+4$ always has a solution for z no matter how large x and y are, so $f(x,y,z)=x^2+y^2+z^2$ has no global maximum. Also, some arguments are needed to show the existence of global minimum.)

$\nabla f(x,y,z) = (2x,2y,2z)$ $\quad\quad$ $\nabla g(x,y,z) = (y,x,2z)$

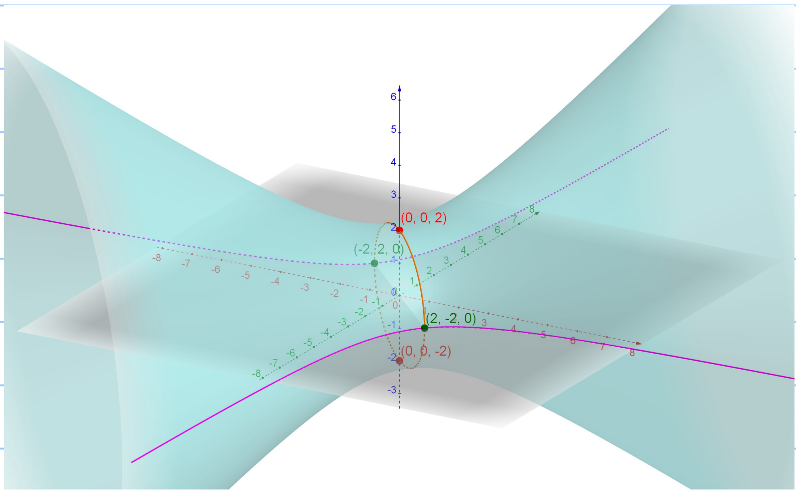
$$\begin{cases} \nabla f(x,y,z) = \lambda \nabla g(x,y,z) \\ z^2 = xy+4 \end{cases} \Rightarrow \begin{cases} 2x = \lambda y & -(1) \\ 2y = \lambda x & -(2) \\ 2z = 2\lambda z & -(3) \\ z^2 = xy+4 & -(4) \end{cases}$$

From (3), we have $\lambda = 1$ or $z=0$.

- If $\lambda=1$, from (1) and (2), we have $2x=y$ and $2y=x$, so $x=y=0$.
  Put $x=y=0$ into (4), we have $z=\pm 2$
  $\therefore$ f may have extreme values at $(0,0,\pm 2)$
- If $z=0$, from (4), $xy=-4$
  (1) × (2), we get
  $$\begin{aligned} 4xy &= \lambda^2 xy \\ \lambda^2 &= 4 \quad (\because xy=-4\neq 0) \\ \lambda &= \pm 2 \end{aligned}$$
  - If $\lambda=2$, from (1), we have $x=y$.
    Put it into $xy=-4$ and we can see $x^2=-4$ has no solution.
  - If $\lambda=-2$, from (1), we have $x=-y$
    Put it into $xy=-4$ and we have $x^2=4$ and so $x=\pm 2$. Then $y=\mp 2$.
    $\therefore$ f may have extreme values at $(\pm 2, \mp 2, 0)$.

$f(0,0,\pm 2) = 4$, $f(\pm 2,\mp 2, 0) = 8$

global min $\quad\quad$??



$f(0,0,\pm 2) = 4$, $f(\pm 2, \mp 2, 0) = 8$

global min ??

(saddle point in fact)

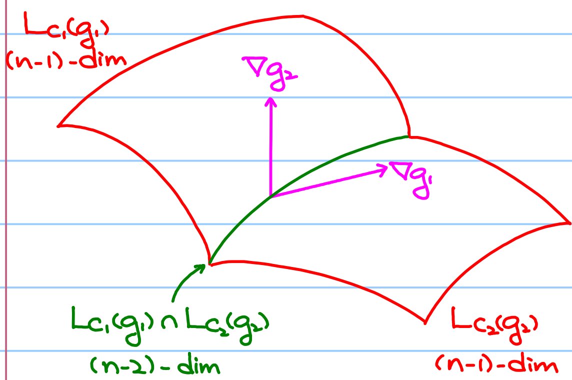
Lagrange Multipliers with More Constraints

Let $f, g_1, g_2 : \mathbb{R}^n \to \mathbb{R}$ be smooth functions.

Suppose that $L_{c_i}(g_i) = \{\vec{x} \in D : g_i(\vec{x}) = c_i\}$ and $\nabla g_i(\vec{x}) \neq \vec{0}$ for all points $\vec{x} \in L_{c_i}(g_i)$, for $i = 1, 2$.

Question: How to find local extrema of f restricted on $L_{c_1}(g_1) \cap L_{c_2}(g_2)$?

Assume $\nabla g_1(\vec{x})$ is not parallel to $\nabla g_2(\vec{x})$ for all $\vec{x} \in L_{c_1}(g_1) \cap L_{c_2}(g_2)$.



Note: ∇g_1 and ∇g_2 are normal to $L_{c_1}(g_1) \cap L_{c_2}(g_2)$.
Therefore, any vector normal to $L_{c_1}(g_1) \cap L_{c_2}(g_2)$ is just linear combination of ∇g_1 and ∇g_2.

Let γ be a smooth curve on $L_{c_1}(g_1) \cap L_{c_2}(g_2)$ such that $\gamma(0) = \vec{x}_0$ and $\gamma'(0)$ is nonzero.

If $f(\gamma(t))$ attains local extrema at $\vec{x}_0 = \gamma(0)$, then

$$\left.\frac{d}{dt}\right|_{t=0} f(\gamma(t)) = 0$$

$$\nabla f(\gamma(0)) \cdot \gamma'(0) = 0$$

$$\nabla f(\vec{x}_0) \cdot \gamma'(0) = 0$$

Therefore, $\nabla f(\vec{x}_0)$ is orthogonal to every vector tangent to $L_{c_1}(g_1) \cap L_{c_2}(g_2)$ at $\vec{x}_0$, and so $\nabla f(\vec{x}_0)$ is normal to $L_{c_1}(g_1) \cap L_{c_2}(g_2)$.

$\therefore$ $\nabla f(\vec{x}_0)$ is a linear combination of $\nabla g_1(\vec{x}_0)$ and $\nabla g_2(\vec{x}_0)$,

i.e. $\nabla f(\vec{x}_0) = \lambda \nabla g_1(\vec{x}_0) + \mu \nabla g_2(\vec{x}_0)$ for some $\lambda, \mu \in \mathbb{R}$.

Example 14.3

Let $f: \mathbb{R}^3 \to \mathbb{R}$ be a function defined by $f(x,y,z) = x^2 + 2y - z^2$

Find local extrema of f restricted on the line L defined by $\begin{cases} 2x - y = 0 \\ y + z = 0 \end{cases}$

Actually, L is the intersection of two planes:

Let $g_1(x,y,z) = 2x - y$ and $g_2(x,y,z) = y + z$, then $L = L_0(g_1) \cap L_0(g_2)$.

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

linearly independent

$$\begin{cases} \nabla f(x,y,z) = \lambda \nabla g_1(x,y,z) + \mu \nabla g_2(x,y,z) \\ 2x - y = 0 \\ y + z = 0 \end{cases} \Rightarrow \begin{cases} 2x = 2\lambda & \text{—(1)} \\ 2 = -\lambda + \mu & \text{—(2)} \\ -2z = \mu & \text{—(3)} \\ 2x - y = 0 & \text{—(4)} \\ y + z = 0 & \text{—(5)} \end{cases}$$

Exercise: $x = \frac{2}{3}$, $y = \frac{4}{3}$, $z = -\frac{4}{3}$, $\lambda = \frac{2}{3}$, $\mu = \frac{8}{3}$

$\therefore f(\frac{2}{3}, \frac{4}{3}, -\frac{4}{3}) = \frac{4}{3}$ is a local extreme point. (In fact, it is a local maximum.)

Exercise: Show that L is given by $\gamma(t) = (t, 2t, -2t)$, for $t \in \mathbb{R}$.

The above question becomes finding extrema of $f(\gamma(t)) = -3t^2 + 4t$.