Week 6: Extrema

Definition of Local and Global Extrema (M&T p. 168 + 180)

Let $f : U \subset \mathbb{R}^n \to \mathbb{R}$ be a function. A point $\mathbf{x}_0$ is called a local minimum if: there is a neighbourhood $\mathbf{x}_0 \in V \subseteq U$ such that:

MINIMA

$$\mathbf{x} \in V \Longrightarrow f(\mathbf{x}_0) \leq f(\mathbf{x}) \qquad \text{“}f(\mathbf{x}_0) \text{ is smaller than all values } \underbrace{f(\mathbf{x}) \text{ near } \mathbf{x}_0}_{\text{topology}}\text{”}$$

A point $\mathbf{x}_0$ is called a global minimum if: or "absolute".

$$\mathbf{x} \in U \Longrightarrow f(\mathbf{x}_0) \leq f(\mathbf{x}) \qquad \text{“}f(\mathbf{x}_0) \text{ is smaller than ALL values } f(\mathbf{x})\text{”}$$

Activity: Think-Pair-Share

Write out the corresponding definitions for local maximum and global maximum.

LOCAL: There is a neighbourhood $V \subseteq U$ such that $x \in V$ and $x \in V \Rightarrow f(x_0) \geq f(x)$

GLOBAL: $x \in U \Rightarrow f(x_0) \geq f(x)$

Example: A Simple Example

Find the global maximum of $f(x, y) = -3x - x^2 - y^2$. ← Complete the square.

Complete the square.

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

$$= -\left(-\left(\tfrac{3}{2}\right)^2 + \left(\tfrac{3}{2}\right)^2 + 3x + x^2\right) - (y)^2$$

$$= \left(\tfrac{3}{2}\right)^2 - \left(\left(\tfrac{3}{2}\right)^2 + 3x + x^2\right) - (y)^2$$

$$= \left(\tfrac{3}{2}\right)^2 - \left(\tfrac{3}{2} + x\right)^2 - (y)^2$$

The absolute max is $f\left(-\tfrac{3}{2}, 0\right) = \left(\tfrac{3}{2}\right)^2 = \tfrac{9}{4}$

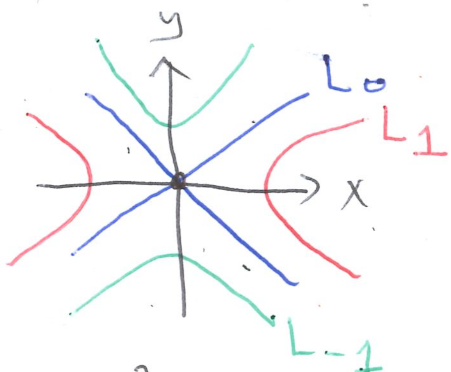
Critical Points and Extrema (M&T p. 168)

A point is a local or relative extrema if it is either a local minimum or maximum.
A point $\mathbf{x}_0$ is a critical point if EITHER: f is not differentiable at $\mathbf{x}_0$ OR $\mathbf{D}f(\mathbf{x}_0) = \mathbf{0}$.
A critical point that is NOT an extrema is a saddle point.

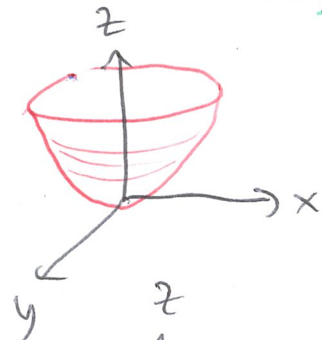
Confusingly, a saddle point does not need to look like a saddle.
We use this definition because it appears in M&T.

Some Examples:

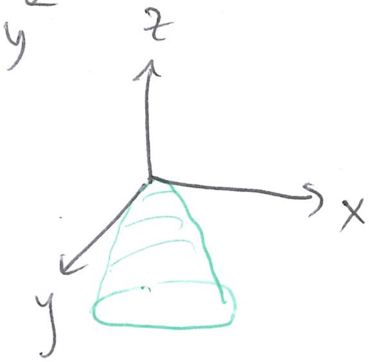
- saddle $f(x,y) = x^2 - y^2$
  $\nabla f = (2x, -2y)$



- minima $g(x,y) = x^2 + y^2$
  $\nabla g = (2x, 2y)$



- maximum $h(x,y) = -x^2 - y^2$
  $\nabla h = (-2x, -2y)$



All have a critical point at $(x,y) = (0,0)$.

$\exp(t) = e^t$

Example: A More Complex Example

Determine the critical points of $f(x, y) = xy\exp(-(x^2 + y^2)/2)$. $= xy\,e^{-(x^2+y^2)/2}$

This function is very nice. It is C^k for all k.

Thus, we check $Df = 0$.

Calculate gradient.

$$Df = \left[\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right] = \left[ye^{-(x^2+y^2)/2} - x^2ye^{-(x^2+y^2)/2}, \ldots\right]$$

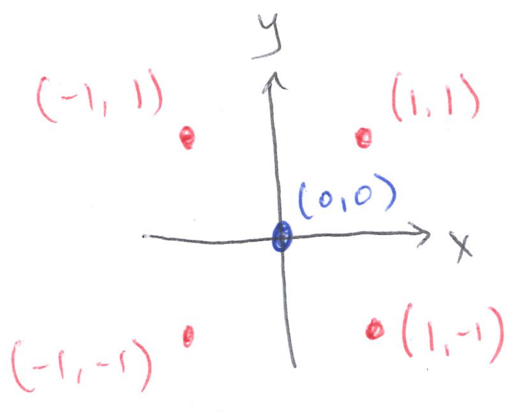
$$= \left[y(1-x^2)e^{-\frac{(x^2+y^2)}{2}}, \underbrace{x(1-y^2)e^{-\frac{(x^2+y^2)}{2}}}_{\text{Symmetry!}}\right] = [0,0]$$

Notice: The exp term will never equal zero.

Set $y(1-x^2) = x(1-y^2) = 0$.

We get: $y(1-x^2) = x(1-y^2) = 0$

Thus $(x,y) = (1,1), (0,0), (-1,1), (1,-1), (-1,-1)$



First Derivative Test for Local Extrema (M&T p. 169)

Suppose $U \subset \mathbb{R}^n$ is an open set and $f : U \subset \mathbb{R}^n \to \mathbb{R}$ is differentiable. If $\mathbf{x}_0 \in U$ is a local extrema, then $\mathbf{D}f(\mathbf{x}_0) = \mathbf{0}$.

Fermat's Theorem

Alternatively, "local extrema and differentiable $\implies$ critical point".

The Big Idea: Reduce to the single-variable case. If f has a local minimum at $\mathbf{x}_0$ then the single variable function $g(t) = f(\mathbf{x}_0 + t\mathbf{h})$ has a local minimum at $t = 0$ for all vectors $\mathbf{h}$. Thus, all the directional derivatives are zero at $\mathbf{x}_0$.

We need to show $\frac{\partial f}{\partial x_k} = 0$ for all k.

We will show the much stronger claim:

$Df\,h = 0$ for all unit vectors h.

(we have: $Df\,e_k = \frac{\partial f}{\partial x_k}$ for e_k being the k^{th} basis vector.)

To prove the stronger claim: pick h to be a unit vector.

Set $\boxed{g(t) = f(x_0 + th)}$ ← This is a single variable funct.

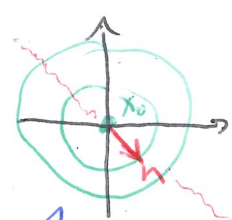
We know $f(x_0) = f(x_0 + 0h) = g(0)$ is a minima.

Thus, $g'(0) = Df\,h = 0$. by single variable calc.

It follows $Df\,h = 0$ for all unit vectors h.

Picking $h = e_1, e_2, \ldots, e_n$ we get:

$$Df = \left[\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots, \frac{\partial f}{\partial x_n}\right] = 0.$$



The function $g(t)$ is $f(x)$ along $x = x_0 + th$.

Example: Checking A Critical Point

Find the critical point of $f(x,y) = x^2y + xy^2$ and check whether it is a maxima, minima, or neither.
Hint: $f(t,t) = 2t^3$ ← How do you come up with this? "x^3" + "y^3"

Find where $Df = 0$.

$$\nabla f = (2xy + y^2, x^2 + 2xy) = (0,0)$$

We get: $\begin{cases} 2xy + y^2 = 0 \\ x^2 + 2xy = 0 \end{cases}$

Aaah! It's non-linear.

$$2xy + y^2 = x^2 + 2xy \Rightarrow x^2 = y^2$$
$$\Rightarrow x = \pm y$$ ← This gives two cases.

Case ①: $x = y$

$$0 = 2xy + y^2 = 2y \cdot y + y^2 = 3y^2 \Rightarrow y = 0.$$

Thus, $(x,y) = (0,0)$.

Case ②: $x = -y$

$$0 = 2xy + y^2 = -2yy + y^2 = -y^2 \Rightarrow y = 0.$$

Thus, $(x,y) = (0,0)$

Along the curve $c(t) = (t,t)$ we get $f(c(t)) = 2t^3$
Thus, $(x,y) = (0,0)$ is neither a min nor a max.
It follows $(0,0)$ is a saddle point.

The Hessian Function (M&T p. 172)

Suppose that $f : U \subset \mathbb{R}^n \to \mathbb{R}$ is C^2. The Hessian of f at $\mathbf{x}_0$ is the quadratic function:

$$Hf(\mathbf{x}_0)(\mathbf{h}) = \frac{1}{2}\sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x}_0) h_i h_j$$

$f(x) \approx \sum_{n=0} \frac{1}{n!} f^{(n)}(0)\, x^n$ Single Var.

$$= \frac{1}{2}[h_1, h_2, \ldots, h_n] \begin{bmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1}(\mathbf{x}_0) & \frac{\partial^2 f}{\partial x_1 \partial x_2}(\mathbf{x}_0) & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(\mathbf{x}_0) \\ \frac{\partial^2 f}{\partial x_2 \partial x_1}(\mathbf{x}_0) & \frac{\partial^2 f}{\partial x_2 \partial x_2}(\mathbf{x}_0) & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n}(\mathbf{x}_0) \\ \vdots & & \ddots & \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(\mathbf{x}_0) & \frac{\partial^2 f}{\partial x_n \partial x_2}(\mathbf{x}_0) & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_n}(\mathbf{x}_0) \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \\ \vdots \\ h_n \end{bmatrix}$$

cancels out deriv.s like x^2 or y^2.

Notice: The Hessian is analogous to the second-derivative of f.

Hessian at a Critical Point (M&T p. 173)

According to our formula for the Second-Order Taylor Series of $f(\mathbf{x}_0 + \mathbf{h})$ we have:

$$f(\mathbf{x}_0 + \mathbf{h}) = \underbrace{f(\mathbf{x}_0)}_{\text{constant}} + \underbrace{[\mathbf{D}f(\mathbf{x}_0)]\mathbf{h}}_{\text{first deriv}} + \underbrace{Hf(\mathbf{x}_0)(\mathbf{h})}_{\text{second deriv}} + R_2(\mathbf{x}_0, \mathbf{h})$$

At a critical point, $\mathbf{D}f(\mathbf{x}_0) = \mathbf{0}$ thus, we get:

$$f(\mathbf{x}_0 + \mathbf{h}) \approx f(\mathbf{x}_0) + \cancel{[\mathbf{D}f(\mathbf{x}_0)]\mathbf{h}} + Hf(\mathbf{x}_0)(\mathbf{h})$$

Notice that $Hf(\mathbf{x}_0)(\mathbf{h})$ is now the first non-constant term.

Example: Locally Quadratic

The function $f(x, y) = x^2 + 2x + 2y + y^2$ has a critical point at $(-1, -1)$.
Write the Second-Order Taylor expansion of $f(x, y)$ at $\mathbf{x}_0 = (-1, -1)$.
Use the Taylor expansion to determine whether $\mathbf{x}_0$ a max or min.

Apply second order Taylor formula.

$f(x_0 + h) \approx f(x_0) + Hf(x_0)(h)$

$= f(x_0) + \frac{1}{2}[h_1\ h_2]\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}\begin{bmatrix} h_1 \\ h_2 \end{bmatrix} = f(x_0) + \frac{1}{2}(2h_1^2 + 2h_2^2)$

$= f(x_0) + h_1^2 + h_2^2 \geq f(x_0)$

we have $f(x_0 + h) \geq f(x_0)$ for all $h \Longrightarrow f(x_0)$ is a min.

Positive Definite Matrices **(M&T p. 175)**

Suppose that $n \times n$ matrix M has quadratic function:

$$Q_M(\mathbf{h}) = \mathbf{h}^T M \mathbf{h}$$

scalar

We say that M is positive definite if: $Q_M(\mathbf{h}) > 0$ for all $\mathbf{h} \neq \mathbf{0}$ and $Q_M(\mathbf{h}) = 0$ is equivalent to $\mathbf{h} = \mathbf{0}$. There is a corresponding notion of negative definite matrices where $Q_M(\mathbf{h}) < 0$ for all $\mathbf{h} \neq \mathbf{0}$ and $Q_M(\mathbf{h}) = 0$ is equivalent to $\mathbf{h} = \mathbf{0}$.

scalar

Activity: Make Some Examples

Find some examples of postive definite and negative definite 2×2 matrices.

Positive $M = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$

$Q_M(h) = 2h_1^2 + 2h_2^2$

Negative $M = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$

$Q_M(h) = -h_1^2 - h_2^2$

Any examples $M \neq \lambda I$?

$M = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$

$Q_M(h) = 2h_1^2 + 2h_1h_2 + 2h_2^2$
$= 2(h_1^2 + h_1h_2 + h_2^2)$

$\det(A_1) = 2 > 0$
$\det(A_2) = 2^2 - 1^2 = 3 > 0$
Thus, positive definite.

Neither $M = \begin{bmatrix} -1 & -1 \\ -1 & -1 \end{bmatrix}$

$Q_M(h) = -x^2 - 2xy - y^2$

$(x,y) = (-1,1) \Rightarrow Q_M(h) = 0$

$\det(M_2) = 0$
Neither positive nor negative definite.

A Simple Test for Definiteness (M&T p. 175)

Give an $n \times n$ matrix $A = [a_{ij}]$ consider the submatrices A_k where $i, j \leq k$.
A is positive definite if all the determinants $\det(A_k)$ are positive.
A is negative definite if the determinants $\det(A_k)$ alternate sign with:

$$A = \begin{bmatrix} [A_k]_{k \times k} & \\ & \end{bmatrix}_{n \times n}$$

Exactly like this: $\det(A_1) < 0 \quad \det(A_2) > 0 \quad \det(A_3) < 0 \quad \det(A_4) > 0 \quad$ etc.

Example: Three Important Examples

I is positive definite and $-I$ is negative definite. The matrix $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ is neither! ← $\det(A_1) = 1$, $\det(A_2) = -1$

2x2

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$A_1 = [1] \Rightarrow \det(A_1) > 0$

$A_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \Rightarrow \det(A_2) > 0$

Positive definite!

$$A = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$$

$A_1 = [-1] \Rightarrow \det(A_1) < 0$

$A_2 = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} \Rightarrow \det(A_2) > 0$

Negative definite!

3x3

$$A = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

$\det(A_1) = -1 < 0$

$\det(A_2) = 1 > 0$

$\det(A_3) = -1 < 0$

Negative definite!

Second Derivative Test for Local Extrema (M&T p. 173)

Suppose that $f : U \subset \mathbb{R}^n \to \mathbb{R}$ is of class C^2 and $\mathbf{x}_0 \in U$ is a critical point.

- $Hf(\mathbf{x}_0)$ is positive definite $\Longrightarrow$ $\mathbf{x}_0$ is a relative minimum.
- $Hf(\mathbf{x}_0)$ is negative definite $\Longrightarrow$ $\mathbf{x}_0$ is a relative maximum.

Neither $\Rightarrow$ No conclusion.

Example: A Complicated Example

Consider the function $f(x, y) = xy \exp(-(x^2+y^2)/2)$. It has critical points $\mathbf{x} = (0, 0)$, $\cancel{(1, \pm 1), (\pm 1, 1)}$. $(\pm 1, \pm 1)$

1. Classify the critical points of $f(x, y)$.
2. Use software to sketch a contour plot of $f(x, y)$ and highlight the critical points.

Find the Hessian.

Recall $\nabla f = \left[y(1-x^2)e^{-\frac{(x^2+y^2)}{2}},\ x(1-y^2)e^{-\frac{(x^2+y^2)}{2}} \right]$

$\cancel{Hf = \begin{bmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{bmatrix} = \Big[-2xye^{-\frac{(x^2+y^2)}{2}} + y(1-x^2)(-x)e^{-\frac{(x^2+y^2)}{2}} \Big)}$

$$f_{xx} = -2xye^{-\frac{(x^2+y^2)}{2}} + y(1-x^2)(-x)e^{-\frac{(x^2+y^2)}{2}}$$

$$= \left(-2xy - xy(1-x^2)\right)e^{-\frac{(x^2+y^2)}{2}}$$

$$= xy\left(-2-(1-x^2)\right)e^{-\frac{(x^2+y^2)}{2}}$$

$$= xy(x^2-3)e^{-\frac{(x^2+y^2)}{2}}$$

$$f_{yy} = xy(y^2-3)e^{-\frac{(x^2+y^2)}{2}}$$ ← Symmetry!

$$f_{xy} = \frac{\partial}{\partial y}\left[y(1-x^2)e^{-\frac{(x^2+y^2)}{2}}\right]$$

$$= (1-x^2)e^{-\frac{(x^2+y^2)}{2}} + (1-x^2)(-y)ye^{-\frac{(x^2+y^2)}{2}}$$

$$= (1-x^2)(1-y^2)e^{-\frac{(x^2+y^2)}{2}}$$

Evaluate at critical points.

$$Hf(0,0) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$ → Saddle

$$Hf(1,1) = \begin{bmatrix} -2e^{-1} & 0 \\ 0 & -2e^{-1} \end{bmatrix}$$ → Maximum

Theorem: Minimize Distance Squared

Minimizing the square distance is equivalent to minimizing distance.
For non-negative d_1 and d_2:

$$d_1 \leq d_2 \iff d_1^2 \leq d_2^2$$

$d = \sqrt{x^2+y^2}$

$d^2 = x^2+y^2$

Notice: The distance function d is not differentiable at the origin, but d^2 is differentiable!

Example: A Simple Minimization

Find the point on the plane $x + y + z = 1$ closest to the origin $(0, 0, 0)$.

\# Minimize distance squared.

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

\# Reduce number of variables.

$f(x,y) = x^2+y^2+(1-x-y)^2$ because $z = 1-x-y$.

\# Find critical points

$$\nabla f = 0 \iff \nabla f = \big(2x-2(1-x-y),\ 2y-2(1-x-y)\big)$$
$$= (4x+2y-2,\ 4y+2x-2)$$

We get a linear system:

$$\begin{cases} 4x+2y = 2 \\ 2x+4y = 2 \end{cases} \iff (x,y) = \left(\tfrac{1}{3}, \tfrac{1}{3}\right)$$

This gives $z = 1-x-y = 1-\frac{1}{3}-\frac{1}{3} = \frac{1}{3}$

This point $(x,y,z) = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$ is a minimum.

Exercise: Check this using 2nd deriv test.

Bounded and Closed (Topology Sneak-Peek) (M&T p. 180)

A set $U \subset \mathbb{R}^n$ is bounded if there is some large constant M such that: $\mathbf{x} \in U \implies \|\mathbf{x}\| \le M$.
Recall, the disk $D_\epsilon(\mathbf{p})$ around $\mathbf{p}$ of radius $\epsilon > 0$ is the set of points at distance at most ϵ from $\mathbf{p}$.

$$D_\epsilon(\mathbf{p}) = \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x} - \mathbf{p}\| < \epsilon\}$$

A set is bounded if: $U \subset D_M(0)$

A boundary point of $U \subset \mathbb{R}^n$ is $\mathbf{p}$ such that $D_\epsilon(\mathbf{p})$ intersects U and $\mathbb{R}^n \setminus U$ for all $\epsilon > 0$.
A set $C \subset \mathbb{R}^n$ is closed if it contains all its boundary points.

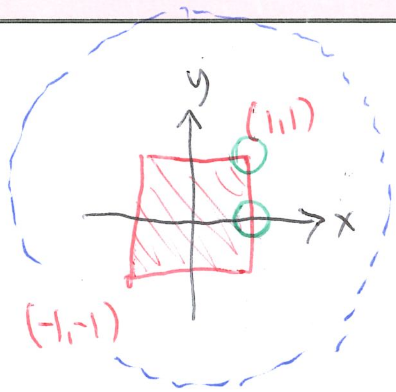
Example: Closed and Bounded Sets

Some examples of closed and bounded sets:

- The square $[-1, 1]^2$
- The circle $S^1 = \{(x, y) : x^2 + y^2 = 1\}$

$0 < x$

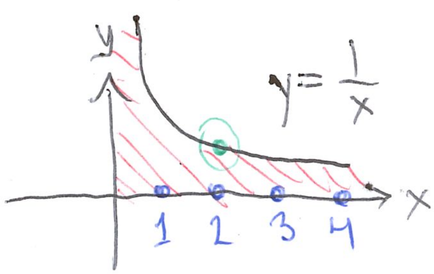
And a non-example: $\left\{(x, y) : 0 \le y < \frac{1}{x}\right\}$.



The square is bounded because $[-1,1]^2 \subset D_{1000}(0)$

Alternatively, $[-1,1]^2 \subset D_{\sqrt{2}}(0)$

The square is closed because it contains its boundary.



It cannot be bounded because it contains $(n, 0)$ for all $n > 0$. No finite radius ball contains all these points.

It is NOT closed because $(2, \frac{1}{2})$ is a boundary point but it is not contained in the set.

☆ Extreme Value Theorem (M&T p. 180)

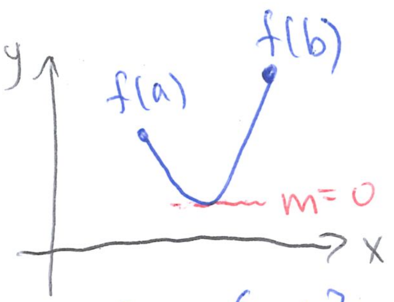
Suppose that $C \subset \mathbb{R}^n$ is closed and bounded. If $f : C \subset \mathbb{R}^n \to \mathbb{R}$ is contimuous then there are points $\mathbf{x}_0$ and $\mathbf{x}_1$ in C where f attains its global maximum and minimum values.

Minima and Maxima for Sets with Boundary (M&T p. 181)

Goal: To find all local/global extrema of f on U with boundary ∂U = "boundary of U"

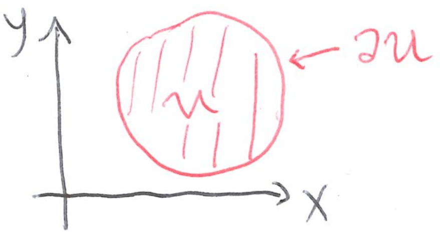
1. Locate all the critical points of f in U.
2. Find all the critical points of the function f on ∂U.
3. Compute f at all the critical points found above.
4. Compare the values of f and select the largest / smallest values.

In one-dimension:



$\partial U = \{a, b\}$

In two-dimensions:



$\partial U =$

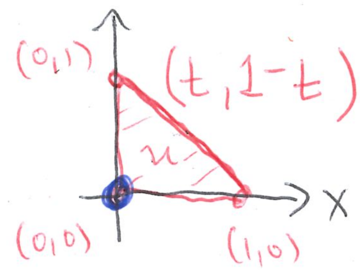
Example: Maximize on a box

Find the maximum of $f(x, y) = xy$ on the triangle with vertices $(0, 0)$, $(1, 0)$, and $(0, 1)$.

Find critical points.

$$\nabla f = (y, x) = (0,0)$$

$$\iff (x,y) = (0,0)$$



Classify critical point

$$\cancel{Hf = \begin{bmatrix} 0 & y \\ x & 0 \end{bmatrix} \implies Hf(0,0) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}}$$

$$Hf = \begin{bmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

$\det[0] = 0$

$\det\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = -1$

No conclusion from second deriv test.

We see $f(x,y) \geq 0$ for all $(x,y) \in \mathcal{U}$.

This critical point is the global min of f on $\mathcal{U}$.

Optimize f on $\partial \mathcal{U}$.

$f(t,0) = t \cdot 0 = 0$ for all t.

$f(0,t) = 0 \cdot t = 0$ for all t.

The hypotenuse of the triangle is $(t, 1-t)$

We get $f(t, 1-t) = t(1-t)$

This is maximized at $t = \frac{1}{2}$ by single var calc.

Thus, the global maximum is $f(\frac{1}{2}, \frac{1}{2}) = \frac{1}{4}$.