

Lecture 13: Extrema

An important problem in multi-variable calculus is to **extremize** a function $f(x, y)$ of two variables. As in one dimensions, in order to look for maxima or minima, we consider points, where the "derivative" is zero.

A point (a, b) in the plane is called a **critical point** of a function $f(x, y)$ if $\nabla f(a, b) = \langle 0, 0 \rangle$.

Critical points are candidates for extrema because at critical points, all directional derivatives $D_{\vec{v}}f = \nabla f \cdot \vec{v}$ are zero. We can not increase the value of f by moving into a direction.

This definition does not include points, where f or its derivative is not defined. We usually assume that a function is arbitrarily often differentiable. Points where the function has no derivatives do not belong to the domain and need to be studied separately. For the function $f(x, y) = |xy|$ for example, we would have to look at the points on the coordinate axes separately.

- 1 Find the critical points of $f(x, y) = x^4 + y^4 - 4xy + 2$. The gradient is $\nabla f(x, y) = \langle 4(x^3 - y), 4(y^3 - x) \rangle$ with critical points $(0, 0), (1, 1), (-1, -1)$.
- 2 $f(x, y) = \sin(x^2 + y) + y$. The gradient is $\nabla f(x, y) = \langle 2x \cos(x^2 + y), \cos(x^2 + y) + 1 \rangle$. For a critical points, we must have $x = 0$ and $\cos(y) + 1 = 0$ which means $\pi + k2\pi$. The critical points are at $\dots(0, -\pi), (0, \pi), (0, 3\pi), \dots$
- 3 The graph of $f(x, y) = (x^2 + y^2)e^{-x^2 - y^2}$ looks like a volcano. The gradient $\nabla f = \langle 2x - 2x(x^2 + y^2), 2y - 2y(x^2 + y^2) \rangle e^{-x^2 - y^2}$ vanishes at $(0, 0)$ and on the circle $x^2 + y^2 = 1$. This function has infinitely many critical points.
- 4 The function $f(x, y) = y^2/2 - g \cos(x)$ is the energy of the pendulum. The variable g is a constant. We have $\nabla f = \langle y, -g \sin(x) \rangle = \langle 0, 0 \rangle$ for $(x, y) = \dots, (-\pi, 0), (0, 0), (\pi, 0), (2\pi, 0), \dots$ These points are equilibrium points, angles for which the pendulum is at rest.
- 5 The function $f(x, y) = a \log(y) - by + c \log(x) - dx$ is left invariant by the flow of the Volterra-Lotka differential equation $\dot{x} = ax - bxy, \dot{y} = -cy + dxy$. The point $(c/d, a/b)$ is a critical point. It is a place where the differential equation has stationary points.
- 6 The function $f(x, y) = |x| + |y|$ is smooth on the first quadrant. It does not have critical points there. The function has a minimum at $(0, 0)$ but it is not in the domain, where f and ∇f are defined.

In one dimension, we needed $f'(x) = 0, f''(x) > 0$ to have a local minimum, $f'(x) = 0, f''(x) < 0$ for a local maximum. If $f'(x) = 0, f''(x) = 0$, then the critical point was undetermined and could be a maximum like for $f(x) = -x^4$, or a minimum like for $f(x) = x^4$ or a flat inflection point like for $f(x) = x^3$.

Let now $f(x, y)$ be a function of two variables with a critical point (a, b) . Define $D = f_{xx}f_{yy} - f_{xy}^2$. It is called the **discriminant** of the critical point.

Remark: You might want to remember it better if you see that it is the determinant of the

Hessian matrix $H = \begin{bmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{bmatrix}$.

Second derivative test. Assume (a, b) is a critical point for $f(x, y)$.

If $D > 0$ and $f_{xx}(a, b) > 0$ then (a, b) is a local minimum.

If $D > 0$ and $f_{xx}(a, b) < 0$ then (a, b) is a local maximum.

If $D < 0$ then (a, b) is a saddle point.

In the case $D = 0$, we need higher derivatives to determine the nature of the critical point.

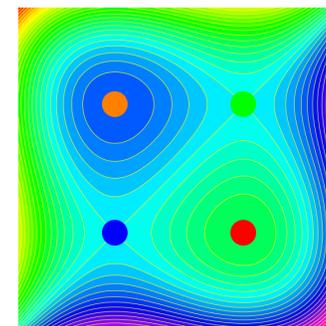
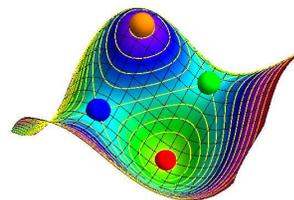
- 7 The function $f(x, y) = x^3/3 - x - (y^3/3 - y)$ has a graph which looks like a "napkin". It has the gradient $\nabla f(x, y) = \langle x^2 - 1, -y^2 + 1 \rangle$. There are 4 critical points $(1, 1), (-1, 1), (1, -1)$ and $(-1, -1)$. The Hessian matrix which includes all partial derivatives is $H = \begin{bmatrix} 2x & 0 \\ 0 & -2y \end{bmatrix}$.

For $(1, 1)$ we have $D = -4$ and so a saddle point,

For $(-1, 1)$ we have $D = 4, f_{xx} = -2$ and so a local maximum,

For $(1, -1)$ we have $D = 4, f_{xx} = 2$ and so a local minimum.

For $(-1, -1)$ we have $D = -4$ and so a saddle point. The function has a local maximum, a local minimum as well as 2 saddle points.



To determine the maximum or minimum of $f(x, y)$ on a domain, determine all critical points **in the interior the domain**, and compare their values with maxima or minima **at the boundary**. We will see next time how to get extrema on the boundary.

- 8 Find the maximum of $f(x, y) = 2x^2 - x^3 - y^2$ on $y \geq -1$. With $\nabla f(x, y) = \langle 4x - 3x^2, -2y \rangle$, the critical points are $(4/3, 0)$ and $(0, 0)$. The Hessian is $H(x, y) = \begin{bmatrix} 4 - 6x & 0 \\ 0 & -2 \end{bmatrix}$. At $(0, 0)$, the discriminant is -8 so that this is a saddle point. At $(4/3, 0)$, the discriminant is 8 and $H_{11} = 4/3$, so that $(4/3, 0)$ is a local maximum. We have now also to look at the boundary $y = -1$ where the function is $g(x) = f(x, -1) = 2x^2 - x^3 - 1$. Since $g'(x) = 0$ at $x = 0, 4/3$, where 0 is a local minimum, and $4/3$ is a local maximum on the line $y = -1$. Comparing $f(4/3, 0), f(4/3, -1)$ shows that $(4/3, 0)$ is the global maximum.

As in one dimensions, knowing the critical points helps to understand the function. Critical points are also physically relevant. Examples are configurations with lowest energy. Many physical laws are based on the principle that the equations are critical points. Newton equations in Classical mechanics are an example: a particle of mass m moving in a field V along a path $\gamma : t \mapsto \vec{r}(t)$ extremizes the integral $S(\gamma) = \int_a^b m\dot{r}'(t)^2/2 - V(r(t)) dt$ among all possible paths. Critical points γ satisfy the Newton equations $m\ddot{r}(t)/2 - \nabla V(r(t)) = 0$.

Why is the second derivative test true? Assume $f(x, y)$ has the critical point $(0, 0)$ and is a quadratic function satisfying $f(0, 0) = 0$. Then

$$ax^2 + 2bxy + cy^2 = a(x + \frac{b}{a}y)^2 + (c - \frac{b^2}{a})y^2 = a(A^2 + DB^2)$$

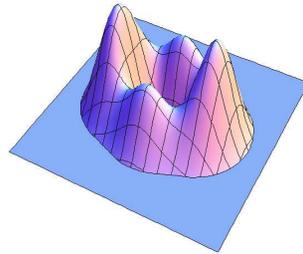
with $A = (x + \frac{b}{a}y)$, $B = b^2/a^2$ and discriminant D . You see that if $a = f_{xx} > 0$ and $D > 0$ then $c - b^2/a > 0$ and the function has positive values for all $(x, y) \neq (0, 0)$. The point $(0, 0)$ is a minimum. If $a = f_{xx} < 0$ and $D > 0$, then $c - b^2/a < 0$ and the function has negative values for all $(x, y) \neq (0, 0)$ and the point (x, y) is a local maximum. If $D < 0$, then the function can take both negative and positive values. A general smooth function can be approximated by a quadratic function near $(0, 0)$.

Sometimes, we want to find the overall maximum and not only the local ones.

A point (a, b) in the plane is called a **global maximum** of $f(x, y)$ if $f(x, y) \leq f(a, b)$ for all (x, y) . For example, the point $(0, 0)$ is a global maximum of the function $f(x, y) = 1 - x^2 - y^2$. Similarly, we call (a, b) a **global minimum**, if $f(x, y) \geq f(a, b)$ for all (x, y) .

- 9 Does the function $f(x, y) = x^4 + y^4 - 2x^2 - 2y^2$ have a global maximum or a global minimum? If yes, find them. **Solution:** the function has no global maximum. This can be seen by restricting the function to the x -axis, where $f(x, 0) = x^4 - 2x^2$ is a function without maximum. The function has four global minima however. They are located on the 4 points $(\pm 1, \pm 1)$. The best way to see this is to note that $f(x, y) = (x^2 - 1)^2 + (y - 1)^2 - 2$ which is minimal when $x^2 = 1, y^2 = 1$.

Here is a curious remark: let $f(x, y)$ be the height of an island. Assume there are only finitely many critical points on the island and all of them have nonzero determinant. Label each critical point with a $+1$ if it is a maximum or minimum, and with -1 if it is a saddle point. Sum up all these number and you will get 1, independent of the function. This theorem of Poincare-Hopf is an example of an "index theorem", a prototype for important theorems in physics and mathematics.



10

The following remarks can be skipped without problem:

- 1) for those of you who have taken linear algebra, you notice that the discriminant D is a determinant $\det(H)$ of the matrix $H = \begin{bmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{bmatrix}$. Besides the determinant, also the trace $f_{xx} + f_{yy}$ is

independent of the coordinate system. The determinant is the product $\lambda_1 \lambda_2$ of the eigenvalues of H and the trace is the sum of the eigenvalues. If the determinant D is positive, then λ_1, λ_2 have the same sign and this is also the sign of the trace. If the trace is positive then both eigenvalues are positive. This means that in the eigendirections, the graph is concave up. We have a minimum. On the other hand, if the determinant D is negative, then λ_1, λ_2 have different signs and the function is concave up in one eigendirection and concave down in an other. In any case, if D is not zero, we have an orthonormal eigenbasis of the symmetric matrix A . In that basis, the matrix H is diagonal.

2) The discriminant D can be considered also at points where we have no critical point. The number $K = D/(1 + |\nabla f|^2)^2$ is called the **Gaussian curvature** of the surface. It is remarkable quantity since it only depends on the intrinsic geometry of the surface and not on the way how the surface is embedded in space. This is the famous Theorema Egregia (=great theorem) of Gauss. Note that at a critical point $\nabla f(x) = \vec{0}$, the discriminant agrees with the curvature $D = K$ at that point. Since we mentioned one theorem for islands already: here is another one, which follows from the famous Gauss-Bonnet theorem: assume you measure the curvature K at each point of an island and assume there is a nice beach all around so that the land disappears flat into the water. In that case the average curvature over the entire island is zero.

3) You might wonder what happens in higher dimensions. The second derivative test then does not work well any more without linear algebra. In three dimensions for example, one can form the second derivative matrix H again and look at all the eigenvalues of H . If all eigenvalues are negative, we have a local maximum, if all eigenvalues are positive, we have a local minimum. If the eigenvalues have different signs, we have a saddle point situation where in some directions the function increases and other directions the function decreases.

Homework

- 1 Find all the extrema of the function $f(x, y) = 2x^3 + 4y^2 - 2y^4 - 6x$ and determine whether they are maxima, minima or saddle points.
 - 2 Where on the parametrized surface $\vec{r}(u, v) = \langle u^2, v^3, uv \rangle$ is the temperature $T(x, y, z) = 12x + y - 12z$ minimal? To find the minimum, look where the function $f(u, v) = T(\vec{r}(u, v))$ has an extremum. Find all local maxima, local minima or saddle points of f .
- Remark.** After you have found the function $f(u, v)$, you could replace the variables u, v again with x, y if you like and look at a function $f(x, y)$.
- 3 Find and classify all the extrema of the function $f(x, y) = e^{-x^2 - y^2}(x^2 + 2y^2)$.
 - 4 Find all extrema of the function $f(x, y) = x^3 + y^3 - 3x - 12y + 20$ on the plane and characterize them. Do you find a global maximum or global minimum among them?

- 5 The thickness of the region enclosed by the two graphs $f_1(x, y) = 10 - 2x^2 - 2y^2$ and $f_2(x, y) = -x^4 - y^4 - 2$ is denoted by $f(x, y) = f_1(x, y) - f_2(x, y)$. Classify all critical points of f and find the global minimal thickness.

