

1 Session 3. Unconstrained optimization

Let f be a function of n variables and let $1 \leq i \leq n$ and $1 \leq j \leq n$. If the partial derivatives f'_i and f'_j of f exist in an open set S containing (x_1, \dots, x_n) and both of these partial derivatives are differentiable at (x_1, \dots, x_n) then $(\text{Hess } f)_{ij} = (\text{Hess } f)_{ji}$.

- If the Hessian at a given point has all positive eigenvalues, it is said to be a positive-definite matrix. This is the multivariable equivalent of “concave up” (we called it ”convex” in this course).
- If all of the eigenvalues are negative, it is said to be a negative-definite matrix. This is like “concave down”. (we call it ”concave” in this course).
- If either eigenvalue is 0, then you will need more information (possibly a graph or table) to see what is going on.
- If the eigenvalues are mixed, you have a saddle point.

Exercise 1

Determine the concavity of

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

at the following points:

1. $(0, 0)$
2. $(3, 3)$
3. $(3, -3)$
4. $(-3, 3)$
5. $(-3, -3)$

Exercise 2

Study the critical points of the functions

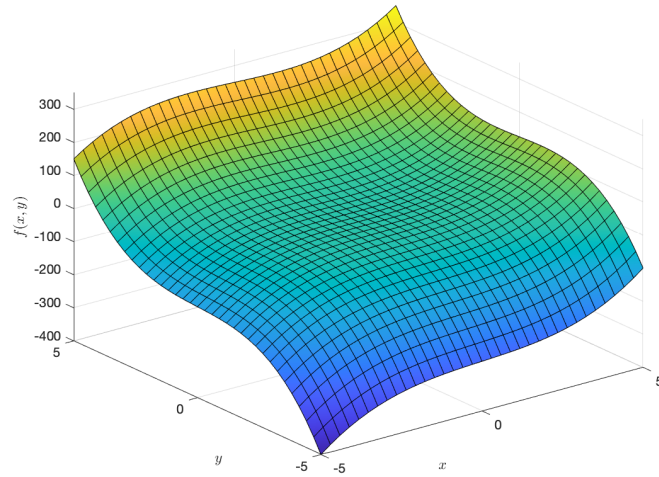
1. $f(x, y) = 4x + 2y - x^2 - 3y^2$
2. $f(x, y) = x^4 + y^2 - xy$

Exercise 3

Check the character of the stationary point in $(0, 0)$ for the functions:

1. $f(x, y) = x^2 + y^2$
2. $g(x, y) = -x^2 - y^2$
3. $h(x, y) = x^2 - y^2$

4. $k(x, y) = x^3 + 2y^3 - xy$ (is it concave or convex in $(3, 3)$? what about in $(-3, -3)$?)



For unconstrained problems with just one variable x , if first and second derivatives exist in x^* :

Necessary conditions If $\frac{df}{dx} = 0$ at $x = x^*$

- $\frac{d^2f}{dx^2} \geq 0$, for a local minimum at $x = x^*$
- $\frac{d^2f}{dx^2} \leq 0$, for a local maximum at $x = x^*$

Sufficient conditions If $\frac{df}{dx} = 0$ at $x = x^*$

- $\frac{d^2f}{dx^2} > 0$, for a local minimum at $x = x^*$
- $\frac{d^2f}{dx^2} < 0$, for a local maximum at $x = x^*$

For unconstrained problems with several variables $\mathbf{x} = (x_1, \dots, x_n)$:

Sufficient condition If $\frac{\partial f}{\partial x_i} = 0$ at $\mathbf{x} = \mathbf{x}^*$:

- $\nabla f = \mathbf{0}$, for a local minimum at $x = x^*$ if the function is convex;
- $\nabla f = \mathbf{0}$, for a local maximum at $x = x^*$ if the function is concave.

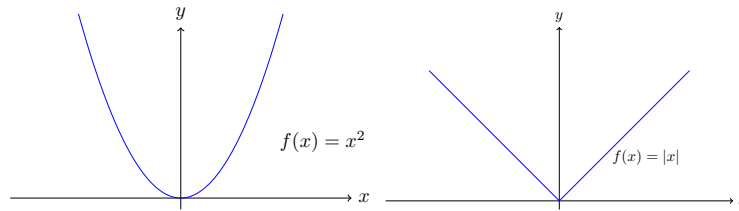
Decision conditions The function f is a convex function if H_f is positive definite or positive semidefinite for all \mathbf{x} ; and f is concave if H_f is negative definite or negative semidefinite for all \mathbf{x} . For $f : \mathbf{R}^2 \rightarrow \mathbf{R}$

$$\mathbf{H}_f(x, y) = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{pmatrix} = \begin{pmatrix} A & C \\ C & B \end{pmatrix}$$

Let f be a twice-differentiable function of many variables on the convex open set S and denote the Hessian of f at the point x by $H(x)$. Then

- f is concave if and only if $H(x)$ is negative semidefinite $\forall x \in S$
- if $H(x)$ is negative definite $\forall x \in S$ then f is strictly concave (see Eq. ??)
- f is convex if and only if $H(x)$ is positive semidefinite $\forall x \in S$
- if $H(x)$ is positive definite $\forall x \in S$ then f is strictly convex (see Eq. ??)

Things are not so simple if, for example, functions are not derivable in x^* :



Or, if the function is not convex (nor concave), the sufficient condition is lost and we may find local optimal points, instead of global.

For constrained optimization problems, the shape of the feasible region adds difficulty.

1.1 Iterative methods

1.1.1 Line search

Let us assume a concave function. First, establish error tolerance ϵ . Then:

1. Set up lower and upper bounds for x by finding x_l, x_u such that $\frac{df}{dx_l} \geq 0$ and $\frac{df}{dx_u} \leq 0$
2. Get new trial intermediate solution (eg, $x = \frac{x_l + x_u}{2}$)
3. If $x_u - x_l \leq \epsilon$, then terminate
4. If $\frac{df}{dx} \geq 0$, set $x_l = x$; if $\frac{df}{dx} \leq 0$, set $x_u = x$
5. Go to step 1

For an implementation, see this link For a continuous function it guarantees to find an extremum.

1.1.2 Golden search

Two quantities are in the golden ratio if their ratio is the same as the ratio of their sum to the larger of the two quantities:

$$\frac{a+b}{a} = \frac{a}{b} = \psi \Rightarrow \psi = \frac{1+\sqrt{5}}{2} = 1.616033988\dots$$

Procedure of the Golden search method, given a function to be minimized, $f(x)$, the interval to be searched as $x \in (x_1, x_4)$, and their functional values $f(x_1)$ and $f(x_2)$:

1. Calculate an interior point and its functional value $f(x_3)$. Ensure the position of x_3 follows the golden ratio between the distances $\frac{x_2-x_1}{x_3-x_1} = \frac{x_3-x_1}{x_2-x_3} = \psi$.
2. Using the triplet, determine if convergence criteria are fulfilled. If they are, estimate the x at the minimum from that triplet and terminate/return.
3. Select x_4 within the largest interval, following the same ratios as in Step 1.
4. The three points for the next iteration will be the one where $f(x)$ is a minimum, and the two points closest to it in x .
5. Go to step 2.

Exercise 4

Use the Golden Search method to find the minimum of function $f(x) = x^2 - 6x + 15$ in the interval $x \in (0, 10)$. Set up a spreadsheet to do so.

Programming exercise 1. *E2. Build a Python code that can run one-dimensional search algorithms for user defined functions, using or not first derivatives. Implement the one-dimensional line search and the golden search algorithms and find the optimal solution for the functions:*

- $f(x) = x^4 - 16x^3 + 45x^2 - 20x + 203$ within the range $x \in (2.5, 14)$
- $g(x) = x^5 - 2x^4 - 23x^3 - 12x^2 + 36x$ within the range $x \in (2, 3)$

Try both a direct implementation as well as the use of solvers in scipy.

Exercise 5

Draw the *Rosenbruck's function*, $f(x, y) = 100(y - x^2)^2 + (1 - x)^2$ and discuss on its convexity, the feasibility of its minimization and the most appropriate numerical method to use.

As a reference (not the central aim of this course) here is a list of some algorithms/methods to explore to perform unconstrained optimization in multiple dimensions:

- Newton's method,
- Quasi-Newton methods (e.g., BFGS),
- Trust-region methods,
- Conjugated gradient methods.

and their many variants.

References