Zaczynamy metody optymalizacji funkcji wielu zmiennych. W naszym przypadku ograniczymy się do funkcji dwóch zmiennych. W pierwszej kolejności omówimy sposób wyznaczania pochodnych numerycznych w dowolnym punkcie, niezbędne do uogólnienia metod na dowolny przykład bez przeliczania ponownie pochodnej. W związku z tym w pierwszej kolejności mają Państwo następujący wzór pierwszej pochodnej bazując na funkcji dwóch zmiennych po zmiennej x:

$$\frac{\partial f(x,y)}{\partial x} = \frac{f(x+h,y) - f(x,y)}{h}$$

Zakładamy, że dla x i y ustawiamy **h=0.00001**. Analogicznie po zmiennej y:

$$\frac{\partial f(x,y)}{\partial y} = \frac{f(x,y+h) - f(x,y)}{h}$$

Kolejne pochodne można wyznaczyć poprzez zagnieżdżenie funkcji licząc pochodną z pochodnej lub policzyć z wyznaczonego wzoru (kolejno mają Państwo drugie pochodne: podwójną po x, podwójną po y oraz podwójną po x i y):

$$\frac{\partial f(x,y)}{\partial^2 x} = \frac{f(x+2h,y) - 2f(x+h,y) + f(x,y)}{h^2}$$

$$\frac{\partial f(x,y)}{\partial^2 y} = \frac{f(x,y+2h) - 2f(x,y+h) + f(x,y)}{h^2}$$

$$\frac{\partial f(x,y)}{\partial x \partial y} = \frac{f(x+h,y+h) - f(x+h,y) - f(x,y+h) + f(x,y)}{h^2}$$

Proszę sprawdzić działanie pochodnych numerycznych na jakiejś prostej funkcji. Podstawić jakiś punkt pod jej pochodną i drugie pochodne i sprawdzić czy w programie wychodzi to samo. Np. x^2y^2, pochodna po *x* to $2xy^2$, po *y* to $2x^2y$, druga pochodna podwójna po *x* to $2y^2$, natomiast podwójna po y, to $2x^2$, druga pochodna mieszana to $4xy$. Teraz stosując w powyższych wzorach funkcję x^2y^2 i podstawiając np. x=2, y=1, powinniśmy otrzymać kolejno: 4, 8, 2, 8, 8.

METODA NEWTONA

DANE: $f(x,y),\ \boldsymbol{x}_0 = [x_0, y_0],\ \varepsilon$ (PRZYKŁAD I ZADANIA NA NASTĘPNEJ STRONIE)

Wzór ogólny: $$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k - \nabla^2 f(\boldsymbol{x}_k)^{-1}\, \nabla f(\boldsymbol{x}_k)$$

gdzie: $$\nabla f(\boldsymbol{x}_k) = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix}, \qquad \nabla^2 f(\boldsymbol{x}_k) = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix}$$

W związku z tym wzór można rozpisać następująco:

$$\begin{bmatrix} x_{k+1} \\ y_{k+1} \end{bmatrix} = \begin{bmatrix} x_k \\ y_k \end{bmatrix} - \begin{bmatrix} \frac{\partial^2 f(x_k,y_k)}{\partial x^2} & \frac{\partial^2 f(x_k,y_k)}{\partial x \partial y} \\ \frac{\partial^2 f(x_k,y_k)}{\partial y \partial x} & \frac{\partial^2 f(x_k,y_k)}{\partial y^2} \end{bmatrix}^{-1} \begin{bmatrix} \frac{\partial f(x_k,y_k)}{\partial x} \\ \frac{\partial f(x_k,y_k)}{\partial y} \end{bmatrix}$$

WARUNEK STOPU: $|\nabla f(x_{k+1}, y_{k+1})| \le \varepsilon \quad \text{lub} \quad |x_{k+1} - x_k| \le \varepsilon \quad oraz \quad |y_{k+1} - y_k| \le \varepsilon$

Dla przypomnienia wzór na macierz odwrotną:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

PRZYKŁAD. Dane: $f(x,y) = 10x^2 + 12xy + 10y^2, \quad x_0 = [10,12]$

W pierwszej kolejności policzymy gradient i macierz Hessego:

$$\nabla f(\boldsymbol{x}_k) = \begin{bmatrix} 20x + 12y \\ 12x + 20y \end{bmatrix}, \qquad \nabla^2 f(\boldsymbol{x}_k) = \begin{bmatrix} 20 & 12 \\ 12 & 20 \end{bmatrix}$$

Wtedy wzór można zapisać:

$$\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} = \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} - \begin{bmatrix} 20 & 12 \\ 12 & 20 \end{bmatrix}^{-1} \begin{bmatrix} 20x_0 + 12y_0 \\ 12x_0 + 20y_0 \end{bmatrix}$$

Podstawiając dane $x_0 = [10,12]$

$$\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} = \begin{bmatrix} 10 \\ 12 \end{bmatrix} - \frac{1}{20 \cdot 20 - 12 \cdot 12} \begin{bmatrix} 20 & -12 \\ -12 & 20 \end{bmatrix} \begin{bmatrix} 20 \cdot 10 + 12 \cdot 12 \\ 12 \cdot 10 + 20 \cdot 12 \end{bmatrix} = \begin{bmatrix} 10 \\ 12 \end{bmatrix} - \frac{1}{256} \begin{bmatrix} 20 & -12 \\ -12 & 20 \end{bmatrix} \begin{bmatrix} 344 \\ 360 \end{bmatrix} =$$

$$= \begin{bmatrix} 10 \\ 12 \end{bmatrix} - \frac{1}{256} \begin{bmatrix} 20 \cdot 344 - 12 \cdot 360 \\ -12 \cdot 344 + 20 \cdot 360 \end{bmatrix} = \begin{bmatrix} 10 \\ 12 \end{bmatrix} - \frac{1}{256} \begin{bmatrix} 2560 \\ 3072 \end{bmatrix} = \begin{bmatrix} 10 \\ 12 \end{bmatrix} - \begin{bmatrix} 10 \\ 12 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Przykładowa funkcja ekstremum osiąga w punkcie [0, 0].