# 9　非線形計画法 (I)：制約のない問題の最適化法／補遺

今，例題として次の 2 変数関数を考える．

$$f(x,y) = 3x^2 + 3y^2 + 2xy - 8x - 8y + 8$$

## 数学的準備

これの勾配ベクトル $\nabla f(x,y)$ は次のようになる．

$$\nabla f(x,y) = \begin{pmatrix} \dfrac{\partial f}{\partial x} \\ \\ \dfrac{\partial f}{\partial y} \end{pmatrix} = \begin{pmatrix} 6x + 2y - 8 \\ 6y + 2x - 8 \end{pmatrix}$$

同様に Hesse 行列 $\nabla^2 f(x,y)$ を求めると，

$$\nabla^2 f(x,y) = \begin{pmatrix} \dfrac{\partial^2 f}{\partial x^2} & \dfrac{\partial^2 f}{\partial x \partial y} \\ \\ \dfrac{\partial^2 f}{\partial y \partial x} & \dfrac{\partial^2 f}{\partial y^2} \end{pmatrix} = \begin{pmatrix} 6 & 2 \\ 2 & 6 \end{pmatrix}$$

この Hesse 行列は，0 ベクトルでない任意のベクトル $(x,y)^T$ に対して $(x+y)^2 + 4x^2 + 4y^2 > 0$ が成り立ち，正定値行列である．

## 制約なし最適化問題

次に，最適解の条件について確認してみる．上記の $f(x,y)$ の最小値は解析的に $(1,1)$ であることが分かっている．この解における勾配ベクトルの値は

$$\nabla f(1,1) = \begin{pmatrix} 6 \cdot 1 + 2 \cdot 1 - 8 \\ 6 \cdot 1 + 2 \cdot 1 - 8 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

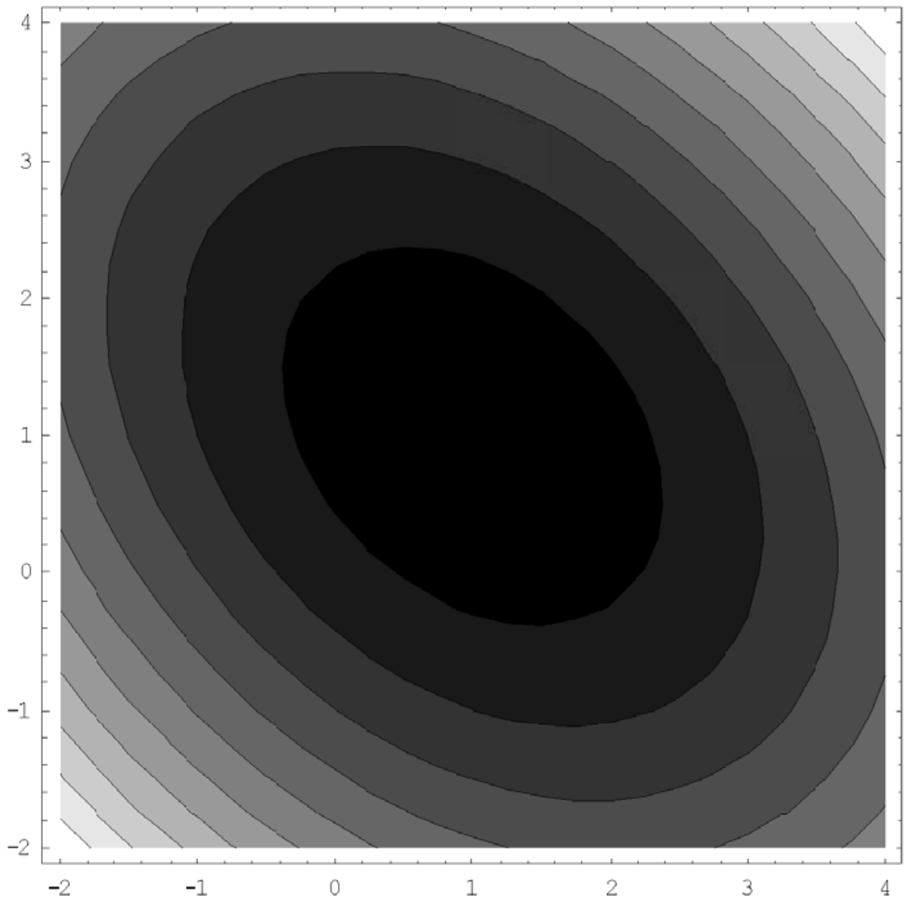
となることが確認できる．

## 最急降下法

$(1,3)$ を初期解とすると，最急降下の方向は $-\nabla f(1,3) = -(4,12)^T$ となる．この降下方向の方程式は $y = 3x$ となり，$f(x,y)$ に代入すると一変数関数の $f(x) = 36x^2 - 32x + 8$ となり，これの最小値は $f'(x) = 72x - 32 = 0$ の解で，$(x,y) = (4/9, 4/3)$．これを次の解として最急降下方向を求めると

$$-\nabla f(\frac{4}{9}, \frac{4}{3}) = -\begin{pmatrix} 6 \cdot \dfrac{4}{9} + 2 \cdot \dfrac{4}{3} - 8 \\ \\ 6 \cdot \dfrac{4}{3} + 2 \cdot \dfrac{4}{9} - 8 \end{pmatrix} = -\begin{pmatrix} -\dfrac{8}{3} \\ \\ \dfrac{8}{9} \end{pmatrix}$$

この降下方向は最初の降下方向と直交していることは，内積をとってみるとわかる．




## ニュートン法

初期解 $(1,3)$ における Hesse 行列を作成してみると，Hesse 行列は定数の行列であり，

$$\nabla^2 f(x,y) = \begin{pmatrix} 6 & 2 \\ 2 & 6 \end{pmatrix}$$

である．これを用いて降下方向のベクトル $(d_x, d_y)^T$ は次の方程式の解となる．

$$\begin{pmatrix} 6 & 2 \\ 2 & 6 \end{pmatrix} \begin{pmatrix} d_x \\ d_y \end{pmatrix} = - \begin{pmatrix} 4 \\ 12 \end{pmatrix}$$

この連立一次方程式を解くと $(d_x, d_y)^T = (0, -2)^T$ となり，次の解は $(x,y)^T = (1+0, 3-2)^T = (1,1)^T$ となる.（最適解）

## 非厳密直線探索

直線探索においては通常，厳密な解を求めることを行わず，以下の **Wolfe の条件**を満足する $\alpha$ を見つけた段階で反復を停止する．ただし，$0 < \sigma_1 < \sigma_2 < 1$ とする．

$$f(\boldsymbol{x}_k + \alpha \boldsymbol{d}_k) - f(\boldsymbol{x}_k) \leq \sigma_1 \alpha (\nabla f(\boldsymbol{x}_k))^T \boldsymbol{d}_k$$

$$(\nabla f(\boldsymbol{x}_k + \alpha \boldsymbol{d}_k))^T \geq \sigma_2 (\nabla f(\boldsymbol{x}_k))^T \boldsymbol{d}_k$$