第5章　非線形計画

5.1　概要

非線形計画問題とは，一般に次のように定式化される最適化問題である．

$$\begin{aligned} &\text{目的関数：} \quad f(\mathbf{x}) \rightarrow \text{ 最小化} \\ &\text{制約条件：} \quad \mathbf{x} \in S \end{aligned}$$

$\mathbf{x}$ は n 変数ベクトル $(x_1, \ldots, x_n) \in R^n$，$f : R^n \rightarrow R$ は実数値をとる関数，S は n 次元空間 R^n の部分集合（実行可能領域）．$S = R^n$ のとき，つまり，

$$\text{目的関数：} \quad f(\mathbf{x}) \rightarrow \text{ 最小化}$$

である問題を**制約なし問題**と呼ぶ．

大域的最適解: その問題に対する，本当の最適解．

局所的最適解: $\mathbf{x}$ の十分近くのどの実行可能解 $\mathbf{x}+\delta\mathbf{x}$ に対しても $f(\mathbf{x}) < f(\mathbf{x}+\delta\mathbf{x})$ であるような $\mathbf{x}$ を，局所的最適解という．局所的最適解はたくさんある可能性がある．大域的最適解は局所的最適解のうちの１つである．

凸関数: 関数 f が

$$\begin{gathered} \mathbf{x}_1, \mathbf{x}_2 \in R^n, 0 \leq \alpha \leq 1 \\ \Downarrow \\ f(\alpha\mathbf{x}_1 + (1-\alpha)\mathbf{x}_2) \leq \alpha f(\mathbf{x}_1) + (1-\alpha) f(\mathbf{x}_2) \end{gathered}$$

を満たすとき，f は**凸関数**であるという．（下に凸）

凸集合: ある集合 S が，

$$\mathbf{x}_1, \mathbf{x}_2 \in S, 0 \leq \alpha \leq 1 \;\Rightarrow\; \alpha\mathbf{x}_1 + (1-\alpha)\mathbf{x}_2 \in S$$

を満たすとき，S は**凸集合**であるという．

凸計画問題: f が凸関数，S が凸集合であるような非線形計画問題.

定理 5.1 凸計画問題には，局所的最適解は 1 つしかなく，それが大域的最適解となる.

したがって，凸計画問題では局所的最適解を求めれば十分である．しかし，一般に，凸計画問題でない非線形問題に対しては多くの局所的最適解が存在する．この講義では，局所的最適解を求めることを目標とする．また，f は 2 階微分可能な関数であるとする.

5.2　準備（勾配，ヘッセ行列）

関数の勾配: 点 $\mathbf{x} = (x_1, \ldots, x_n)$ における関数 $f(\mathbf{x})$ の勾配を次の $\nabla f(\mathbf{x})$ で表す.

$$\nabla f(\mathbf{x}) = \begin{pmatrix} \frac{\partial f(\mathbf{x})}{\partial x_1} \\ \vdots \\ \frac{\partial f(\mathbf{x})}{\partial x_n} \end{pmatrix}$$

$\nabla f(\mathbf{x})$ は，点 $\mathbf{x}$ 付近において，f が最も増加する方向と，その増加量を示す．詳しく言えば，$\nabla f(\mathbf{x}) \neq \mathbf{0}$ である場合，$\mathbf{x}$ から $\nabla f(\mathbf{x})$ 方向 ± 90 °の方向に（わずかに）進むと $f(\mathbf{x})$ の値は必ず増加し，逆に，$-\nabla f(\mathbf{x})$ 方向 ± 90 °の方向に（わずかに）進むと $f(\mathbf{x})$ の値は必ず減少する.

$f(\mathbf{x})$ の値が等しい $\mathbf{x}$ からなる線（面）を等高線（面）という．勾配を示すベクトルは，$f(\mathbf{x})$ の等高線と _____ になっている.

ヘッセ行列: $\nabla f(\mathbf{x})$ をさらに微分した行列:

$$\nabla^2 f(\mathbf{x}) = \begin{pmatrix} \frac{\partial^2 f(\mathbf{x})}{(\partial x_1)^2} & \frac{\partial^2 f(\mathbf{x})}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f(\mathbf{x})}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f(\mathbf{x})}{\partial x_2 \partial x_1} & \frac{\partial^2 f(\mathbf{x})}{(\partial x_2)^2} & \cdots & \frac{\partial^2 f(\mathbf{x})}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f(\mathbf{x})}{\partial x_n \partial x_1} & \frac{\partial^2 f(\mathbf{x})}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f(\mathbf{x})}{(\partial x_n)^2} \end{pmatrix}$$

を，**ヘッセ行列**と呼ぶ.

$\frac{\partial^2 f(\mathbf{x})}{\partial x_i \partial x_j} = \frac{\partial^2 f(\mathbf{x})}{\partial x_j \partial x_i}$ であるから，ヘッセ行列は _____ 行列である．ヘッセ行列は，点 $\mathbf{x}$ 付近で，f の勾配がどうなっているかを示している．ヘッセ行列と，関数の幾何学的な性質との関係について，もう少し調べてみよう．解析の道具は,

- 対称行列の固有値は全て実数となる.

- 任意のベクトル $\mathbf{x}$ に対して $\mathbf{x}^T\mathbf{A}\mathbf{x} \geq 0$ となる行列 $\mathbf{A}$ を**半正定値**と呼ぶ．行列 $\mathbf{A}$ が対称行列であるとき，$\mathbf{A}$ が半正定値であることと $\mathbf{A}$ の全ての固有値が非負であることは等価.

- 全てのベクトル $\mathbf{x} \neq \mathbf{0}$ に対して $\mathbf{x}^T\mathbf{A}\mathbf{x} > 0$ となる行列 $\mathbf{A}$ を**正定値**と呼ぶ．行列 $\mathbf{A}$ が対称行列であるとき，$\mathbf{A}$ が正定値であることと全ての固有値が正であることは等価.

まず，関数の例として，２次の 2 変数関数 $f(x,y) = x^2 + 2y^2$ を考える．$f(x,y)$ の等高線は，原点を中心とし，方向 $(1,0),(0,1)$ を短軸と長軸とした同心の _____ になる．一方，ヘッセ行列は

$$\nabla^2 f(x,y) = \begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix}$$

となり，２つの固有値は $2,4$，固有ベクトルはそれぞれ $(1,0)^T,(0,1)^T$ であり，等高線を表す _____ の軸方向に一致している．

では，この f を原点を中心に 45 度回転させた関数ではどうなるか見てみよう．関数は

$$f'(x,y) = \left(\frac{1}{\sqrt{2}}x + \frac{1}{\sqrt{2}}y\right)^2 + 2\left(-\frac{1}{\sqrt{2}}x + \frac{1}{\sqrt{2}}y\right)^2 = \frac{3}{2}x^2 + \frac{3}{2}y^2 - xy$$

である．$f'(x,y)$ の等高線は，原点を中心とし，方向 _____ を短軸と長軸とした同心の楕円になる．ヘッセ行列は

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

となり，２つの固有値は _____, 固有ベクトルはそれぞれ _____ であり，これもまた等高線の軸方向に一致している．

一般に，関数 $f(\mathbf{x})$ のヘッセ行列が正定値であるとき，f の等高線（面）は同心の楕円となり，固有ベクトルはその楕円の軸方向を示している．

5.3 最適性条件

大域的最適解であることを判定することは大変な作業であるので，局所的最適解（その近辺においては目的関数が最小となる解）を考える．この章では，与えれた解が局所的最適解であることを判定するための条件について説明する．

5.3.1 制約なし問題

制約条件のない非線形問題の場合，とにかく目的関数 f を小さくすればよいので比較的簡単である．

図 5.1: 局所最適解の十分条件と必要条件の関係

１次の必要条件： $\mathbf{x}^*$ が局所的最適解であるためには

$$\nabla f(\mathbf{x}^*) = \mathbf{0} \tag{5.1}$$

でなければならない.

２次の必要条件： $\mathbf{x}^*$ が局所的最適解であるためには式 (5.1) かつ

$$\nabla^2 f(\mathbf{x}^*) \text{ が半正定値} \tag{5.2}$$

でなけらばならない.

２次の十分条件： 式 (5.1) かつ

$$\nabla^2 f(\mathbf{x}^*) \text{ が正定値} \tag{5.3}$$

ならば，$\mathbf{x}^*$ は局所的最適解である.

式（5.1）は必要条件であるが十分条件ではないので，$\nabla f(\mathbf{x}^*) = \mathbf{0}$ であっても，$\mathbf{x}^*$ が局所的最適解であるとは限らない.（たとえば，_____ などの**停留点**.）同様に，式（5.2）も必要条件であって十分条件ではない．逆に，式 (5.3) は十分条件であって必要条件ではないので，局所的最適解であっても式 (5.3) を満たさないものもあるので，注意が必要である.

これらの条件の関係を図 5.1 に示す.

例として，$\mathbf{x} = (x_1, x_2)^T$，$f(\mathbf{x}) = x_1^2 + 2x_2^2$ を考えてみよう．最適解は $\mathbf{x}^* = (0, 0)$ であるが,

$$\nabla f(\mathbf{x}) = \begin{pmatrix} \text{—} \\ \text{—} \end{pmatrix}, \quad \nabla f(\mathbf{x}^*) = \begin{pmatrix} \text{—} \\ \text{—} \end{pmatrix}$$

$$\nabla^2 f(\mathbf{x}) = \begin{pmatrix} \text{—} & \text{—} \\ \text{—} & \text{—} \end{pmatrix} = \nabla^2 f(\mathbf{x}^*), \quad \text{固有値は} \underline{\qquad\qquad}$$

であるから，1 次の必要条件，2 次の必要条件，2 次の十分条件は_____
.