第13回　非線形計画法II
— 最適性条件2, 解法 —

13.1　制約つき問題における最適性条件

制約つき非線形計画問題は，一般に

$$\begin{aligned} &\text{目的関数：} && f(\mathbf{x}) \to \text{ 最小化} \\ &\text{制約条件：} && g_i(\mathbf{x}) = 0 \ (1 \le i \le l) \\ & && g_i(\mathbf{x}) \le 0 \ (l+1 \le i \le m) \end{aligned}$$

のように書くことができる．g_i $(1 \le i \le l)$ を等式制約関数，g_i $(l+1 \le i \le m)$ を不等式制約関数という．等式制約は，２つの不等式制約に書き換えることができる．

$$g_i(\mathbf{x}) = 0 \ \Leftrightarrow \ \begin{cases} g_i(\mathbf{x}) \le 0 \\ \underline{\qquad\qquad\qquad\qquad} \end{cases}$$

したがって，全ての制約条件が不等式制約であるとしても一般性を失わない．以後は等式制約のない問題：

$$\begin{aligned} &\text{目的関数：} && f(\mathbf{x}) \to \text{ 最小化} \\ &\text{制約条件：} && g_i(\mathbf{x}) \le 0 \ (1 \le i \le m) \end{aligned}$$

のみを扱う．ここでは，目的関数と全ての制約関数は微分可能であるとする．

制約つき問題における最適性は，最適解が境界上にあるかないかによって状況が変わってくる．最適解 $\mathbf{x}^*$ が境界上にない（内点である）場合，制約なし問題の場合と同様に

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

が１次の必要条件となる．しかし，最適解 $\mathbf{x}^*$ が境界上にある場合，

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

が成り立つとは限らない．では，最適解 $\mathbf{x}^*$ が境界上にある場合の必要条件はどのように表されるか，考えてみよう．

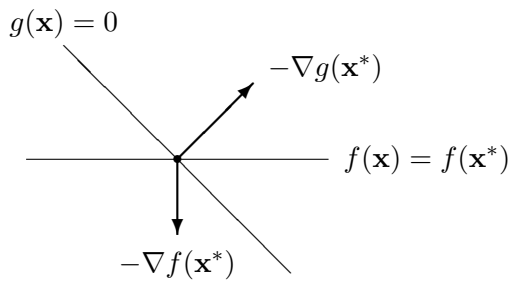

図 13.1: f の勾配と g の勾配との関係

たとえば，１つの不等式制約関数 $g(\mathbf{x}) \leq 0$ の境界上の点 $\mathbf{x}^*$ を考える（図 13.1 の中心にある点）．点 $\mathbf{x}^*$ から，点 $\mathbf{x}^*$ における g の勾配方向 ± 90 度以内の方向に進むと，$g(\mathbf{x})$ の値は _____ なり，その領域内の $\mathbf{x}$ は制約条件 $g(\mathbf{x}) \leq 0$ を満たさなくなる．逆に，点 $\mathbf{x}^*$ から，点 $\mathbf{x}^*$ における g の勾配と逆の方向 ± 90 度以内の方向に進むと，$g(\mathbf{x})$ の値は _____ なり，制約条件は満たされている（図 13.1 の縦斜線部分）．

一方，点 $\mathbf{x}^*$ における f の勾配と反対の方向（$-\nabla f(\mathbf{x}^*) \pm 90$ 度未満の方向）にある点では，$f(\mathbf{x})$ は $f(\mathbf{x}^*)$ より小さくなっている（図 13.1 の斜線部分）．

$-\nabla g(\mathbf{x}^*) \pm 90$ 度以内の領域と，$-\nabla f(\mathbf{x}^*) \pm 90$ 度未満の領域に，重なる部分（図の斜線の重なった部分）が存在する場合，その領域内にある点は実行可能解であって，$\mathbf{x}^*$ よりも目的関数を小さくできる．したがって，このときは $\mathbf{x}^*$ は明らかに最適でない．

つまり，$\mathbf{x}^*$ が局所的最適解であるためには，$\nabla g(\mathbf{x}^*) \pm 90$ 度以内の領域と $-\nabla f(\mathbf{x}^*) \pm 90$ 度未満の領域に共通部分が存在してはいけない．このようなことは，$-\nabla g(\mathbf{x}^*)$ と $-\nabla f(\mathbf{x}^*)$ が _____ を向いているときに起こる．すなわち，ある非負の数 u が存在して

$$u \nabla g(\mathbf{x}^*) + \nabla f(\mathbf{x}^*) = \mathbf{0}$$

が成り立つときである．したがってこれが１つの不等式制約関数の境界上の点に対する必要条件となる．

次に，2 つの不等号制約関数 $g_1(\mathbf{x}) \leq 0$ と $g_2(\mathbf{x}) \leq 0$ の境界線の交点にある点 $\mathbf{x}^*$ を考える．$-\nabla g_1(\mathbf{x}^*) \pm 90$ 度以内**かつ** $-\nabla g_2(\mathbf{x}^*) \pm 90$ 度以内の方向にある点は，制約条件を _____．したがって，そのような点が $-\nabla f(\mathbf{x}^*) \pm 90$ 度未満の方向に存在する場合，$\mathbf{x}^*$ は明らかに _____．

したがって，$\mathbf{x}^*$ が局所的最適解であるためには，非負の数 u_1, u_2 が存在して，

$$u_1 \nabla g_1(\mathbf{x}^*) + u_2 \nabla g_2(\mathbf{x}^*) + \nabla f(\mathbf{x}^*) = \mathbf{0} \tag{13.1}$$

となる必要がある．

例外として，$-\nabla g_1(\mathbf{x}^*)$ と $-\nabla g_2(\mathbf{x}^*)$ がちょうど逆方向を向いている（線形従属である）場合，$\mathbf{x}^*$ が局所的最適解であっても式 (13.1) を満たす u_1, u_2 が存在しない場合もある．

以上の考察を一般化すると，次の条件となる．

１次の必要条件（キューン・タッカー条件）： 解 $\mathbf{x}^*$ が問題

$$\begin{array}{ll} \text{目的関数：} & f(\mathbf{x}) \to \text{ 最小化} \\ \text{制約条件：} & g_i(\mathbf{x}) \le 0 \ (1 \le i \le m) \end{array}$$

の局所的最適解であり，すべての g_i に対して勾配ベクトル $\nabla g_i(\mathbf{x}^*)$ が１次独立でないならば，$u_1, \ldots, u_m$ が存在して，

$$\left\{ \begin{array}{l} \nabla f(\mathbf{x}^*) + u_1 \nabla g_1(\mathbf{x}^*) + \cdots + u_m \nabla g_m(\mathbf{x}^*) = \mathbf{0} \\ \left. \begin{array}{l} g_i(\mathbf{x}^*) \le 0,\ u_i \ge 0 \\ g_i(\mathbf{x}^*) < 0 \text{ ならば } u_i = 0 \end{array} \right\} (1 \le \forall i \le m) \end{array} \right.$$

が成り立つ．

キューン・タッカー条件は，制約なし問題の１次の必要条件式 (12.1) を拡張したものである．u_i を**ラグランジュ乗数**と呼ぶ．$g_i(\mathbf{x}^*) = 0$ である制約条件を**有効制約**と呼ぶ．

キューン・タッカー条件は必要条件であるから，もちろん，条件を満たしているからといって $\mathbf{x}^*$ が局所最適解になっているとは限らない．ただし，目的関数 f が凸関数，等式制約関数 g_i が１次関数，不等式制約関数が凸関数であるような，凸計画問題に対しては，キューン・タッカー条件を満たす $\mathbf{x}^*$ はその問題の大域的最適解であることがわかっている．次の問題を解いてみよう．

[**問題 13.1**]　2 変数の凸計画問題:

$$\begin{array}{lll} \text{目的関数} & : & f(\mathbf{x}) = (x_1 - 3)^2 + 2(x_2 - 3)^2 + 4(x_3 - 3)^2 \to \text{ 最小化} \\ \text{制約条件} & : & g(\mathbf{x}) = x_1 + 2x_2 + 3x_3 - 11 \le 0 \end{array}$$

の最適解（大域的最適解＝局所最適解）を求めよ．

まず，

$$\nabla f(\mathbf{x}) = \begin{pmatrix} \underline{\qquad\qquad} \\ \underline{\qquad\qquad} \\ \underline{\qquad\qquad} \end{pmatrix}$$

より，$\nabla f(\mathbf{x}) = \mathbf{0}$ となる $\mathbf{x}$ は _____ のみであり，これは制約条件を満たさないので，最適解は境界上にあることがわかる．$\mathbf{x}^* = (x_1^*, x_2^*, x_3^*)$ を境界上の最適解とすると，キューン・タッカー条件よりある非負の数 u が存在して

$$\nabla f(\mathbf{x}^*) + u \nabla g(\mathbf{x}^*) = \mathbf{0}$$

が成り立つ.

$$\nabla f(\mathbf{x}^*) = \begin{pmatrix} 2(x_1^* - 3) \\ 4(x_2^* - 3) \\ 8(x_3^* - 3) \end{pmatrix}, \nabla g(\mathbf{x}^*) = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

より，x_1^*, x_2^*, x_3^* をすべて u で表わすと

$$x_1^* = \underline{\qquad\qquad\qquad}$$
$$x_2^* = \underline{\qquad\qquad\qquad}$$
$$x_3^* = \underline{\qquad\qquad\qquad}$$

であり，境界にあるので $g(\mathbf{x}^*) = 0$ を満たしているので,

$$g(\mathbf{x}^*) = \underline{\qquad\qquad\qquad\qquad\qquad\qquad} = 0.$$

これを解くと，$u =$___，$x_1^* =$___，$x_2^* =$___, $x_3^* =$___ を得る．$u > 0$ であるから，$\mathbf{x}^* = ($___ , ___$)$ は最適解である．

[**課題 13.1**] 次の 2 変数の凸計画問題を解け.

$$\begin{aligned} \text{目的関数} &: \ f(\mathbf{x}) = e^{-x_1} + e^{-x_2} \rightarrow \text{ 最小化} \\ \text{制約条件} &: \ x_1 + 3x_2 \leq 4 \end{aligned}$$

13.2 解法 1：最急降下法

制約なし問題に対する 1 次の必要条件を用いた解法に，最急降下法がある．これは，ある適当な解 $\mathbf{x}^{(t)}$ に対して $\nabla f(\mathbf{x}^{(t)}) \neq \mathbf{0}$ ならば，$\mathbf{x}^{(t)}$ から $\mathbf{d}^{(t)} = -\nabla f(\mathbf{x}^{(t)})$ の方向に進むと，必ず目的関数の値が _____ ことを利用している．具体的な手順は以下のとおり.

最急降下法

(Step 1) 適当に近似解 $\mathbf{x}^{(0)}$ を選ぶ．$t = 0$ とする.

(Step 2) $\nabla f(\mathbf{x}^{(t)}) = \mathbf{0}$ であれば，$\mathbf{x}^{(t)}$ が停留点であるので終わり．そうでなければ，$\mathbf{d}^{(t)} = -\nabla f(\mathbf{x}^{(t)})$ とし，$f(\mathbf{x}^{(t)} + \alpha^{(t)}\mathbf{d}^{(t)})$ が最小となるような正の数 $\alpha^{(t)}$ を選ぶ.

(Step 3) $\mathbf{x}^{(t+1)} = \mathbf{x}^{(t)} + \alpha^{(t)}\mathbf{d}^{(t)}$ を次の近似解とし，$t \leftarrow t + 1$ として step 2 に戻る.

$\alpha^{(t)}$ を**ステップ幅**と呼ぶ．実際に解く場合は，有限回の繰り返しで $\nabla f(\mathbf{x}^{(t)}) = \mathbf{0}$ となる $\mathbf{x}^{(t)}$ を得ることは期待できないので，十分小さい正数 ϵ を定めておき，$||\nabla f(\mathbf{x}^{(t)})|| < \epsilon$ であれば終了する.

真の最適解を $\mathbf{x}^*$ としたとき，最急降下法ではある正の定数 $\beta < 1$ が存在して，

$$||\mathbf{x}^{(t+1)} - \mathbf{x}^*|| \leq \beta ||\mathbf{x}^{(t)} - \mathbf{x}^*||$$

となることが保証されている．このような性質を持つ解法は，_____ という．

最急降下法の性質

- どこを出発点 $\mathbf{x}^{(0)}$ として選んでも，必ず１つの停留点に収束していく．
  これを _____ という．
- 収束の速さが遅い．

最急降下法は，制約なし問題の解法であるが，少し変形することによって制約つき問題を解くのにも使うことができる．制約なし問題の場合，ステップ幅は $f(\mathbf{x}^{(t+1)})$ が最小となるように選ぶが，制約つき問題の場合，$f(\mathbf{x}^{(t+1)})$ が最小となる前に，$\mathbf{x}^{(t+1)}$ が実行可能領域を出てしまうこともありうる．したがって，上記の step 2 のステップ幅は，$f(\mathbf{x}^{(t+1)})$ が最小となる，または $\mathbf{x}^{(t+1)}$ が実行可能領域の境界になるように選ぶ．また，終了条件もキューン・タッカー条件を用いる．

13.3 解法２：ニュートン法

制約なし問題に対する２次の必要条件を用いた解法であるニュートン法について説明する．

目的関数を点 $\mathbf{x}^*$ の付近で近似すると，

$$f(\mathbf{x}^* + \mathbf{d}) \simeq f(\mathbf{x}^*) + \nabla f(\mathbf{x}^*)^T \mathbf{d} + \frac{1}{2}\mathbf{d}^T \nabla^2 f(\mathbf{x}^*)\mathbf{d}$$

となる．ここで $\mathbf{d} = (d_1, \ldots, d_n)^T$ は微小なベクトル．$\nabla^2 f(\mathbf{x}^*)$ が正定値であると仮定すれば，２次関数

$$q(\mathbf{d}) \triangleq f(\mathbf{x}^*) + \nabla f(\mathbf{x}^*)^T \mathbf{d} + \frac{1}{2}\mathbf{d}^T \nabla^2 f(\mathbf{x}^*)\mathbf{d}$$

は凸関数となる．（$f(\mathbf{x}^*)$ は定数項，$\nabla f(\mathbf{x}^*)^T \mathbf{d}$ は１次の項，$\frac{1}{2}\mathbf{d}^T \nabla^2 f(\mathbf{x}^*)\mathbf{d}$ は２次の項.）$q(\mathbf{d})$ の勾配を計算すると，

$$\nabla q(\mathbf{d}) = \begin{pmatrix} \frac{\partial q(\mathbf{d})}{\partial d_1} \\ \vdots \\ \frac{\partial q(\mathbf{d})}{\partial d_n} \end{pmatrix}$$

$$= \underline{\hspace{10cm}}$$

$$= \underline{\hspace{4cm}}$$

であるから，$\nabla f(\mathbf{x}^*) + \nabla^2 f(\mathbf{x}^*)\mathbf{d} = \mathbf{0}$ となる $\mathbf{d}$ で，$q(\mathbf{d})$ は極小値を取る．$q(\mathbf{d})$ は，$\mathbf{x}^*$ の近傍において目的関数を近似した関数であるから，$\mathbf{x}^* + \mathbf{d}$ を次の近似解として考えることができる．以上の考え方を利用した解法の具体的な手順は以下のとおり．

ニュートン法

(Step 1) 適当に初期近似解 $\mathbf{x}^{(0)}$ を定める．$t = 0$ とする．

(Step 2) $\nabla f(\mathbf{x}^{(t)}) = \mathbf{0}$ であれば $\mathbf{x}^{(t)}$ を出力して終了．

(Step 3) $\nabla^2 f(\mathbf{x}^{(t)})$ が正定値であれば

$$\nabla f(\mathbf{x}^{(t)}) + \nabla^2 f(\mathbf{x}^{(t)})\mathbf{d} = \mathbf{0}$$

を満たす $\mathbf{d}$ を求め，$\mathbf{x}^{(t+1)} = \mathbf{x}^{(t)} + \mathbf{d}$ とする．$t \leftarrow t+1$ として step 2 にもどる．

真の最適解を $\mathbf{x}^*$ としたとき，ニュートン法ではある正の定数 β が存在して，

$$||\mathbf{x}^{(t+1)} - \mathbf{x}^*||_G \leq \beta ||\mathbf{x}^{(t)} - \mathbf{x}^*||_G^2$$

となることが保証されている．（ただし，$\mathbf{G} = \nabla^2 f(\mathbf{x}^*), ||\mathbf{x}||_G = \mathbf{x}^T\mathbf{G}\mathbf{x}$.）このような性質を _____ という．

ニュートン法の性質

- 収束が速い．
- $\nabla^2 f(\mathbf{x}^{(t)})$ が正定値でなければいけない．$\mathbf{x}^{(t)}$ が十分局所的最適解に近ければ，$\nabla^2 f(\mathbf{x}^{(t)})$ は正定値となるが，任意の点について言えるわけではない．したがって，ニュートン法は大域的収束性はもっていない．（_____ ）

13.4 まとめ

最適性条件

- 制約なし問題

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

１次の必要条件：$\nabla f(\mathbf{x}^*) = 0$

２次の必要条件：$\nabla f(\mathbf{x}^*) = 0,$
$\nabla^2 f(\mathbf{x}^*)$ が半正定値

２次の十分条件：$\nabla f(\mathbf{x}^*) = 0,$
$\nabla^2 f(\mathbf{x}^*)$ が正定値

- 制約つき問題

$$\begin{aligned} &\text{目的関数：} \quad f(\mathbf{x}) \to \text{最小化} \\ &\text{制約条件：} \quad g_i(\mathbf{x}) \le 0 \ (1 \le i \le m) \end{aligned}$$

１次の必要条件（キューン・タッカー条件）：$\mathbf{x}^*$ が局所最適解であり，全ての $\nabla g_i(\mathbf{x}^*)$ が１次独立ならば，ラグランジュ乗数 $u_1, \ldots, u_m$ が存在して

$$\begin{cases} \nabla f(\mathbf{x}^*) + u_1 \nabla g_1(\mathbf{x}^*) + \cdots + u_m \nabla g_m(\mathbf{x}^*) = \mathbf{0} \\ g_i(\mathbf{x}^*) \le 0, \ u_i \ge 0 \\ g_i(\mathbf{x}^*) < 0 \text{ ならば } u_i = 0 \end{cases}$$

が成り立つ.

解法

最急降下法 ：１次の必要条件を利用．１次収束（遅い）．大域的収束性.

ニュートン法 ：２次の必要条件を利用．２次収束（速い）．局所的収束（初期近似解が局所的最適解に十分近くないと収束しない）.