YNU 幾何学セミナー 第 3 回

多目的最適化と特異点論 2：特異点，パレート臨界点，パレート最適点

濱田 直希
株式会社富士通研究所 人工知能研究所 人工知能基盤 PJ
hamada-naoki@jp.fujitsu.com

2017 年 6 月 30 日

概要

本稿では，多目的最適化と特異点論の関係を紹介する．はじめに 1 章では，扱うべき写像 $f : X \to \mathbb{R}^m$ を関数 $f : X \to \mathbb{R}$ に単純化したケースとして，単目的最適化とモース理論の関係を紹介する．1.1 節では単目的最適化を，1.2 節ではモース理論を紹介する．1.3 節では，X が境界のない滑らかな多様体で f がモース関数の場合には，単目的最適化において求めるべき局所最適点がモース理論における安定臨界点と一致することを示す．

次に 2 章では，1 章の議論を写像 $f : X \to \mathbb{R}^m$ へと一般化することにより，多目的最適化と特異点論の関係を示す．1 章で紹介したそれぞれの概念に対して，その写像バージョンが得られる．2.1 節では多目的最適化を，2.2 節では特異点論を紹介する．2.3 節では，X が境界のない滑らかな多様体で f がモース関数の写像版の場合には，多目的最適化において求めるべき局所パレート最適点が特異点論における安定パレート臨界点と一致し，それらが特異点集合に含まれることを示す．

本稿で紹介する定義や定理は新しいものではない．1 章の大部分は 1930 年代から 1960 年代にかけての成果であり，2 章の大部分は 1960 年代から 1980 年代にかけての成果である．したがって，それぞれの分野の専門家にとっては周知の内容であると思われる．しかし，単目的最適化とモース理論，多目的最適化と特異点論のつながりを横断的な視点で紹介した文献は少ないため，それらをまとめた資料を作成しておくことには一定の意義があると考える．また，理論そのものは比較的整備されているものの，そこに登場する概念を具体的に計算するためのアルゴリズムについては，今まさに盛り上がりつつある研究領域である．本稿では詳しく紹介できないが，1.3 節と 2.3 節に簡単なコメントをつけた．なお，本稿 2.2 節の大部分は，山本 卓宏 先生（東京学芸大学 准教授）のセミナー（2017 年 4 月 28 日 九州大学）を LaTeX 原稿に書き起こしたものである．この場を借りて感謝の意を表する．

1　単目的最適化とモース理論

単純化したケースとして，単目的最適化とモース理論の関係から出発する．次章で述べる多目的最適化と特異点論の関係は，本章の内容を多値に拡張したものである．本章で議論の枠組みを掴んでおくことにより，次章が理解しやすくなると思われる．

1.1　単目的最適化

単目的最適化問題とその解の概念，最適性条件を紹介する．証明は標準的な教科書や，最近では WEB の講義資料でも豊富に見つかるため省略する．

定義 1.1　**単目的最適化問題** *(single-objective optimization problem)* とは，**目的関数** *(objective function)* $f : \mathbb{R}^n \to \mathbb{R}$ と**制約関数** *(constraint functions)* $g_1, \ldots, g_k : \mathbb{R}^n \to \mathbb{R}$ が与えられたとき，**実行可能領域** *(feasible region)*

$$X := \{x \in \mathbb{R}^n \mid g_1(x) \leq 0, \ldots, g_k(x) \leq 0\}$$

上に制限した目的関数 $f|_X : X \to \mathbb{R}$ の最小点を求める問題であり，以下のように表記される．

$$\begin{aligned} &\text{minimize} \quad f(x) \\ &\text{subject to} \quad g_i(x) \leq 0 \quad \text{for all } i = 1, \ldots, k. \end{aligned}$$

問題の解，すなわち，求めるべき最小点は以下のように定義される．

定義 1.2　f の最小値を与える点の集合

$$\theta_{\mathrm{O}} := \{x \in X \mid \nexists y \in X \text{ s.t. } f(y) < f(x)\}$$

を f の**最小点集合** *(minimum point set)* という．

問題に特別な仮定をおくことなく，最小点を求めることは難しい．より求めやすい解の概念として，局所的な最小点がしばしば用いられる．

定義 1.3　f の極小値を与える点の集合

$$\theta_{\mathrm{L}} := \{x \in X \mid \exists U_x : \text{ a nbd. of } x, \nexists y \in U_x \text{ s.t. } f(y) < f(x)\}$$

を f の**局所最小点集合** *(local minimum point set)* という．

f が滑らかな関数ならば，与えられた点が局所最小点であるかどうかは微分を用いて判定できる．局所最小点の必要条件は 1 階微分を用いて与えられる．

定理 1.1　f の局所最小点 x は以下の *1* **次の条件** *(first-order condition)*[*1]を満たす．$\exists \lambda \geq 0$, $\exists \mu = (\mu_1, \ldots, \mu_k) \geq 0$ s.t. $(\lambda, \mu) \neq 0$ について以下が成り立つ：

$$\begin{aligned} \lambda \nabla f(x) + \sum_j \mu_j \nabla g_j(x) = 0, \\ \mu_j g_j(x) = 0 \quad \text{for all } j = 1, \ldots, k. \end{aligned}$$

以下の例が示すように，これは十分条件ではない．

例 1　$X = \mathbb{R}, f(x) = x^3$ とする．$x = 0$ は 1 次の条件を満たすが，局所最小点ではない．

局所最小点の十分条件は 2 階微分を用いて与えられる．

定理 1.2　以下の *2* **次の条件** *(second-order condition)* を満たす点 x は f の局所最小点である．$J(x) = \{j \in \mathbb{N} \mid g_j(x) = 0\}$, $J^+(x) = \{j \in J(x) \mid \mu_j > 0\}$ とおく．1 次の条件を満たす λ, μ と，すべての

$$w \in \left\{0 \neq w \in \mathbb{R}^n \mid {}^t(\nabla f(x))w \leq 0,\ {}^t(\nabla g_j(x))w \leq 0 \text{ for all } j \in J(x)\right\}$$

[*1] *Fritz John* 条件 *(Fritz John condition)* ともよばれる．これに加えて，制約関数に若干の仮定をおき，$\lambda > 0$ を要求したものは *KKT* 条件 *(Karush-Kuhn-Tucker condition)* とよばれる．

または

$$w \in \left\{0 \neq w \in \mathbb{R}^n \mid {}^t(\nabla g_j(x))w = 0 \text{ for all } j \in J^+(x),\ {}^t(\nabla g_j(x))w \leq 0 \text{ for all } j \in J(x) \setminus J^+(x)\right\}$$

のいずれかについて，

$${}^t w \left(\lambda \nabla^2 f(x) + \sum_j \mu_j \nabla^2 g_j(x) \right) w > 0.$$

以下の例が示すように，これは必要条件ではない．

例 2　$X = \mathbb{R}, f(x) = 0$ とする．$x = 0$ は 2 次の条件を満たさないが，局所最小点である．

大域的な最小点が求められるケースとして，以下の仮定がしばしば用いられる．

定義 1.4　X が**凸集合** *(convex set)*，すなわち，

$$tx + (1-t)y \in X \quad \text{for all } t \in (0,1), x \neq y \in X$$

であり，f が**凸関数** *(convex function)*，すなわち，

$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y) \quad \text{for all } t \in (0,1), x \neq y \in X$$

であるとき，問題は**凸問題** *(convex problem)* であるという．

このとき，例 1 のようなケースは排除され，1 次の条件は最小点の必要十分条件になる．

定理 1.3　凸問題において，1 次の条件は局所最小点の必要十分条件であり，局所最小点は最小点である．

凸性だけでは解の一意性は保証されない．これを保証するために，より強い凸性が仮定されることもある．

定義 1.5　X が凸集合で f が**狭義凸関数** *(strictly convex function)*，すなわち，

$$f(tx + (1-t)y) < tf(x) + (1-t)f(y) \quad \text{for all } t \in (0,1), x \neq y \in X$$

であるとき，問題は**狭義凸問題** *(strictly convex problem)* であるという．

このとき，例 2 のようなケースは排除され，解は一意になる．

定理 1.4　狭義凸問題において，局所最小点は（存在すれば）一意な最小点である．

1.2　モース理論

最適化と関係の深いモース理論の概念や定理を紹介する．証明は広く知られているので省略する．例えば Milnor [8] などを参照されたい．

本節をとおして，X は滑らかな n 次元多様体，$f : X \to \mathbb{R}$ は滑らかな関数とする．

定義 1.6　f の微分

$$\nabla f(x) := {}^t \left(\frac{\partial f}{\partial x_1}(x), \ldots, \frac{\partial f}{\partial x_n}(x) \right)$$

が消える点の集合

$$\theta := \{x \in X \mid \nabla f(x) = 0\}$$

を f の**臨界点集合** *(critical point set)* という．

定義 1.7　f の臨界点 x において，ヘッセ行列

$$Hf(x) := \begin{pmatrix} \dfrac{\partial^2 f}{\partial x_1 \partial x_1}(x) & \cdots & \dfrac{\partial^2 f}{\partial x_1 \partial x_n}(x) \\ \vdots & \ddots & \vdots \\ \dfrac{\partial^2 f}{\partial x_n \partial x_1}(x) & \cdots & \dfrac{\partial^2 f}{\partial x_n \partial x_n}(x) \end{pmatrix}$$

が**非特異** *(non-singular)*，すなわち

$$\operatorname{rank} Hf(x) = n$$

であるとき，臨界点 x は**非退化** *(non-degenerate)* という．すべての臨界点が非退化である関数を**モース関数** *(Morse function)* という．

滑らかな関数のなかには，モース関数でない関数も存在する．しかし，そのような関数も，わずかな摂動を加えるだけでモース関数にできる．言い換えれば，モース関数は任意の関数を近似できるだけ豊富にある．

定理 1.5　モース関数は**ジェネリック** *(generic)* である．すなわち，C^∞ 級関数の集合 $C^\infty(X, \mathbb{R})$ において，モース関数全体は C^∞ 位相のもとで稠密な開集合をなす．

抽象的な存在証明だけでなく，より具体的に与えることもできる．

定理 1.6　$X \subseteq \mathbb{R}^n$ とする．ほとんどすべての $p \in \mathbb{R}^n$ について，**距離二乗関数** *(distance squared function)*

$$d_p^2(x) := \|x - p\|^2$$

は X 上のモース関数である．

臨界点の性質を調べるうえで，次のモースの補題が重要である．

定理 1.7　モース関数 f の任意の臨界点 p において，ある局所座標 $x = (x_1, \dots, x_n)$ がとれて，

$$f(x) = -x_1^2 - \cdots - x_i^2 + x_{i+1}^2 + \cdots + x_n^2 + f(p)$$

と書ける．

定義 1.8　上記の表示におけるマイナスの数 i を**指数** *(index)* という．指数 0 の臨界点を**極小点** *(minimal point)* という．

定義 1.9　モース関数 $f : X \to \mathbb{R}$ に対して，

$$f^{-1}(c) := \{x \in X \mid f(x) = c\}$$

を**レベルセット** *(level set)* といい，その連結成分を**等高線** *(contour)* という．商空間

$$X/\sim \quad (x \sim y \Leftrightarrow x, y \in X \text{ は } f \text{ の同じ等高線に属する})$$

を**レーブグラフ** *(Reeb graph)* という．

レーブグラフはグラフとして表示できる．ノードは臨界点からなる等高線に対応し，エッジは正則点からなる等高線に対応する．

レーブグラフは関数からレベルセットの変化に関する情報だけを抽出したものである．純粋に位相的な概念であるため，計量的な情報は完全に失われてしまう．ベクトル場を分析したいときには，もう少し計量的な情報を残したまま関数を要約したいこともある．そのようなときに有用なのがモース＝スメール複体である．

定義 1.10　曲線 $\varphi : \mathbb{R} \to X$ で

$$\frac{\partial}{\partial t}\varphi(t) = -\nabla f \circ \varphi(t) \quad \text{for all } t \in \mathbb{R}$$

を満たすものを**積分曲線** *(integral curve)* という．

定義 1.11　臨界点 x が**安定** *(stable)* であるとは，$x \in X$ の任意の近傍 V_x に対して，$x \in V_x$ のある近傍 U_x が存在して，積分曲線 $\varphi : \mathbb{R} \to X$ のうち

$$\varphi(0) \in U_x$$

であるものすべてが

$$\varphi(t) \in V_x \quad \text{for all } t \geq 0$$

を満たすことをいう．特に，$\lim_{t\to+\infty} \varphi(t) = x$ ならば**漸近的に安定** *(asymptotically stable)* という．安定でなければ，**不安定** *(unstable)* という．

$$\theta_{\mathrm{S}} := \{x \in \theta \mid x : \mathrm{stable}\}$$

を**安定臨界点集合** *(stable critical point set)* という．

臨界点 $x \in \theta$ について，

$$S(x) := \{x\} \cup \left\{x_0 \in X \mid x_0 \in \operatorname{Im}\varphi, \lim_{t\to+\infty} \varphi(t) = x\right\}$$

を**安定多様体** *(stable manifold)* といい，

$$U(x) := \{x\} \cup \left\{x_0 \in X \mid x_0 \in \operatorname{Im}\varphi, \lim_{t\to-\infty} \varphi(t) = x\right\}$$

を**不安定多様体** *(unstable manifold)* という．

定義 1.12　任意の $x, y \in \theta$ に対して，安定多様体 $S(x)$ と不安定多様体 $U(y)$ は横断的に交わるとする．

$$C := \{S(x) \cap U(y) \mid x, y \in \theta\}$$

を**モース＝スメール複体** *(Morse-Smale complex)* という．

1.3　両者の関係

1.3.1　最小点が内点のケース

定理 1.8（包含関係）　モース関数について，以下の関係が成り立つ．

$$\theta_{\mathrm{O}} \subseteq \theta_{\mathrm{L}} = \theta_{\mathrm{S}} \subseteq \theta$$

証明 1　定義より $\theta_{\mathrm{O}} \subseteq \theta_{\mathrm{L}}$ と $\theta_{\mathrm{S}} \subseteq \theta$ は自明．$\theta_{\mathrm{L}} = \theta_{\mathrm{S}}$ を背理法で示す．

もし $\theta_{\mathrm{L}} \not\subseteq \theta_{\mathrm{S}}$ ならば，ある点 $x \in \theta_{\mathrm{L}} \setminus \theta_{\mathrm{S}}$ が存在して，x は退化している．これはモース関数の仮定と矛盾する．

もし $\theta_{\mathrm{L}} \not\supseteq \theta_{\mathrm{S}}$ ならば，ある点 $x \in \theta_{\mathrm{S}} \setminus \theta_{\mathrm{L}}$ が存在して，x のどの近傍にも f の値を減少させる点が存在することになる．これは，安定臨界点の指数が 0，すなわち，その近傍で f が凸関数であることと矛盾する．　□

単目的最適化では，（局所）最小点を具体的に求めることが問題となる．そのために様々な解法が研究されている．以下では，その代表的な解法とモース理論がどのような関係にあるかを述べる．

例 3（勾配法）　単目的最適化問題の解法としては，勾配法が最も基本的である．勾配法の基本的な枠組みは以下のようになり，α_t の決め方によって様々な亜種がある [9]．

1. 初期点 $x_0 \in X$ を適当に選ぶ．
2. 点列 $x_1, x_2, \ldots$ を以下のように決める．

$$x_{t+1} = x_t - \alpha_t \nabla f(x_t). \tag{1}$$

勾配法の収束先は臨界点であるから，必ずしも大域的な最小点が得られるとは限らない．そこで，異なる初期点から再度探索を行うリスタートが行われる．

式 (1) からわかるように，系列 $x_0, x_1, \ldots$ は積分曲線の近似である．積分曲線がどの臨界点に流れ込むかを類別したものがモース＝スメール複体であった．近年，サンプルからモース＝スメール複体を近似計算するアルゴリズムが提案されている [2]．これを応用すれば，効率的なリスタート戦略が開発できる可能性がある．

例 4（遺伝的アルゴリズム）　勾配法のリスタートのように，1 点 1 点順番に収束するまで探索するのは効率が悪いこともある．そこで，探索途中の情報を共有して複数の点を同時に探索する方法が考えられる．そのようなものの 1 つとして，遺伝的アルゴリズムがある．遺伝的アルゴリズムの基本的な枠組みは以下のようになり，ステップ 2 と 3 の作り方によって様々な亜種がある [11]．

1. X の点をランダムに n 個生成し，それらの集合を X_0 とおく．$t = 0$ とする．
2. X_t の周辺に多数の点をランダムに生成し，Y_t とおく．
3. Y_t から n 点を f の値にしたがって選び，X_{t+1} とする．
4. $t := t + 1$ としてステップ 2 に戻る．

この方法は，レベルセット

$$f^{-1}(c) = \{x \in X \mid f(x) = c\}$$

をサンプル X_t で近似し，$c \to -\infty$ へと追跡しているといえる．したがって，遺伝的アルゴリズムが大域的最小点を見つけることができるかどうかは，サンプルの系列 $X_0, X_1, \ldots$ が関数のレーブグラフをうまく追跡できるかどうかにかかっている．

1.3.2　最小点が X の境界点であるケース

実際の最適化問題では，最小点が X の境界にあるケースがしばしば見られる．ジェネリックに，実行可能領域 X は**滑走分割集合** *(stratified set)* になる．滑層分割集合上でモース理論を考えることになるため，**滑走分割モース理論** *(stratified Morse theory)* [4] が必要になる．この理論では，臨界点の近傍において接線方向

の位相変化だけでなく法線方向の位相変化も考えることにより，多様体の張り合わせ構造を記述する．数学的にはすでに一通りの道具は揃っているが，計算アルゴリズムは未開拓であり，純粋な理論としても最適化に応用した研究はまだないと思われる．

2　多目的最適化と特異点論

本章では，前章の議論を関数から写像へと拡張することにより，多目的最適化と特異点論の関係について述べる．本章で示す概念は，すべて前章の概念の拡張になっている．おおむね同じ番号同士が対応するように記述してあるので，前章の対応する個所を見ることで理解の助けにしてほしい．

2.1　多目的最適化

多目的最適化問題とその解の概念，最適性条件を紹介する．証明は Miettinen [7] を参照されたい．

定義 2.1（[7, p. 37]）　**多目的最適化問題** *(multi-objective optimization problem)* とは，目的関数 $f_1, \dots, f_m : \mathbb{R}^n \to \mathbb{R}$ と制約関数 $g_1, \dots, g_k : \mathbb{R}^n \to \mathbb{R}$ が与えられたとき，**評価写像** *(evaluation mapping)* $f := (f_1, \dots, f_m) : \mathbb{R}^n \to \mathbb{R}^m$ を実行可能領域

$$X := \{x \in \mathbb{R}^n \mid g_1(x) \le 0, \dots, g_k(x) \le 0\}$$

上に制限した写像 $f|_X : X \to \mathbb{R}^m$ のパレート点を求める問題であり，以下のように表記される．

$$\begin{array}{ll} \text{minimize} & f_1(x), \dots, f_m(x) \\ \text{subject to} & g_i(x) \le 0 \quad \text{for all } i = 1, \dots, k. \end{array}$$

問題の解，すなわち，求めるべきパレート点は以下のように定義される．

定義 2.2（[6, Definition 1]）　どの f_i の値を減少させるためにも，いずれかの f_j の値を増加させなければならないような点の集合

$$\theta_{\mathrm{O}} := \{x \in X \mid \nexists y \in X \text{ s.t. } f(y) \le f(x), f(y) \ne f(x)\}$$

を f の**パレート点集合** *(Pareto point set)* という．

問題に特別な仮定をおくことなく，パレート点を求めることは難しい．より求めやすい解の概念として，局所的なパレート点がしばしば用いられる．

定義 2.3（[6, Definition 1]）　ある近傍においてパレートであるような点の集合

$$\theta_{\mathrm{L}} := \{x \in X \mid \exists U_x : \text{ a nbd. of } x, \nexists y \in U_x \text{ s.t. } f(y) \le f(x), f(y) \ne f(x)\}$$

を f の**局所パレート点集合** *(local Pareto point set)* という．

f が滑らかな関数ならば，与えられた点が局所パレート点であるかどうかは微分を用いて判定できる．局所パレート点の必要条件は 1 階微分を用いて与えられる．

定理 2.1（[7, Theorem 3.1.1]） f の局所パレート点 x は以下の *1* **次の条件** *(first-order condition)* を満たす．$\exists\lambda = (\lambda_1, \ldots, \lambda_m) \geq 0, \exists\mu = (\mu_1, \ldots, \mu_k) \geq 0 \text{ s.t. } (\lambda, \mu) \neq (0, 0)$ について以下が成り立つ：

$$\sum_i \lambda_i \nabla f_i(x) + \sum_j \mu_j \nabla g_j(x) = 0,$$
$$\mu_j g_j(x) = 0 \quad \text{for all } j = 1, \ldots, k.$$

以下の例が示すように，これは十分条件ではない．

例 5 $X = \mathbb{R}, f_1(x) = f_2(x) = x^3$ とする．$x = 0$ は 1 次の条件を満たすが，局所パレート点ではない．

局所パレート点の十分条件は 2 階微分を用いて与えられる．

定理 2.2（[7, Theorem 3.1.12]） 以下の *2* **次の条件** *(second-order condition)* を満たす点 x は f の局所パレート点である．$J(x) = \{j \in \mathbb{N} \mid g_j(x) = 0\}$, $J^+(x) = \{j \in J(x) \mid \mu_j > 0\}$ とおく．1 次の条件を満たす λ, μ と，すべての

$$w \in \left\{0 \neq w \in \mathbb{R}^n \mid {}^t(\nabla f_i(x))w \leq 0 \ \text{ for all } i,\ {}^t(\nabla g_j(x))w \leq 0 \ \text{ for all } j \in J(x)\right\}$$

または

$$w \in \left\{0 \neq w \in \mathbb{R}^n \mid {}^t(\nabla g_j(x))w = 0 \ \text{ for all } j \in J^+(x),\ {}^t(\nabla g_j(x))w \leq 0 \ \text{ for all } j \in J(x) \setminus J^+(x)\right\}$$

のいずれかについて，

$${}^t w \left(\sum_i \lambda_i \nabla^2 f_i(x) + \sum_j \mu_j \nabla^2 g_j(x) \right) w > 0.$$

以下の例が示すように，これは必要条件ではない．

例 6 $X = \mathbb{R}, f_1(x) = f_2(x) = 0$ とする．$x = 0$ は 2 次の条件を満たさないが，局所パレート点である．

大域的なパレート点が求められるケースとして，以下の仮定がしばしば用いられる．

定義 2.4（[7, Definition 2.1.3]） X が凸集合で $f_1, \ldots, f_m$ がすべて凸関数であるとき，問題は**凸問題** *(convex problem)* であるという．

このとき，例 5 のようなケースは排除され，1 次の条件はパレート点の必要十分条件になる．

定理 2.3（[7, Theorem 3.1.8]） 凸問題において，1 次の条件は局所パレート点の必要十分条件であり，局所パレート点はパレート点である．

凸性だけでは同じ値をもつパレート点が複数存在することがある．これを防ぐために，より強い凸性が仮定されることもある．

定義 2.5（[7, Definition 2.1.6]） X が凸集合で $f_1, \ldots, f_m$ がすべて狭義凸関数であるとき，問題は**狭義凸問題** *(strictly convex problem)* であるという．

このとき，例 6 のようなケースは排除され，すべてのパレート点は異なる値をもつようになる．

定理 2.4（[7, Theorem 3.1.12] の系） 狭義凸問題において，局所パレート点は（存在すれば）一意な値をもつパレート点である．

2.2　特異点論

本節では，X は滑らかな n 次元多様体とし，$f : X \to \mathbb{R}^m$ は滑らかな写像とする[*2].

定義 2.6（[6]） f のヤコビ行列

$$Df(x) = {}^t\left(Df_1(x), \ldots, Df_m(x)\right) = \begin{pmatrix} \dfrac{\partial f_1}{\partial x_1}(x) & \cdots & \dfrac{\partial f_1}{\partial x_n}(x) \\ \vdots & \ddots & \vdots \\ \dfrac{\partial f_m}{\partial x_1}(x) & \cdots & \dfrac{\partial f_m}{\partial x_n}(x) \end{pmatrix}$$

が特異な点の集合

$$\Sigma := \{x \in X \mid \operatorname{rank} Df(x) < \min(m, n)\}$$

を**特異点集合** *(singular point set)* という．

写像 $f = (f_1, \ldots, f_m)$ のパレート点は，必ずしも個々の目的関数 f_i の極小点ではない．したがって，臨界点の拡張概念としては，「$f_1, \ldots, f_m$ の勾配が消える点」ではなく「$f_1, \ldots, f_m$ の勾配が対立する点」を考えたい．

定義 2.7（[6, Definition 2]） $\mathbb{R}^m$ における**正開錐** *(positive open cone)* を

$$\mathrm{Pos} := \{y \in \mathbb{R}^m \mid y_1 > 0, \ldots, y_m > 0\}$$

とおく．集合

$$\theta := \{x \in X \mid \operatorname{Im} Df(x) \cap \mathrm{Pos} = \emptyset\}$$

を**パレート臨界点集合** *(Pareto critical point set)* という．

定理 2.5（[6, Proposition 3]） パレート臨界点 $x \in \theta$ について，以下が成り立つ（両者は同値である）．

(a)　$Df_1(x), \ldots, Df_m(x)$ は余接空間 T_x^*X の共通の開半空間に属さない．
(b)　ある $\lambda := (\lambda_1, \ldots, \lambda_m) \geq 0$ s.t. $\lambda \neq 0$ で $\sum_i \lambda_i Df_i(x) = 0$ を満たすものが存在する．

距離二乗関数の写像版もジェネリックである．

定理 2.6（[5, Proposition 2.1]） $2 \leq m \leq n$ とする．一般の位置にある $p_1, \ldots, p_m \in \mathbb{R}^n$ について，**距離二乗写像** *(distance squared mapping)*

$$d^2_{(p_1, \ldots, p_m)}(x) := (d^2_{p_1}(x), \ldots, d^2_{p_m}(x))$$

は正定値折り目写像と $\mathcal{A}$-同値である．ここで，$d_p^2(x)$ は距離二乗関数である．

レーブグラフの拡張を考えることができる．

[*2] つまり，前節の多目的最適化問題の X を $\operatorname{int} X$ に置き換えたケースに対応する．言い換えれば，最小点が実行可能領域の内点であるようなケースを扱っている．

定義 2.8（[1, p. 245]） 連続写像 $f : X \to \mathbb{R}^m$ に対して，

$$f^{-1}(c) := \{x \in X \mid f(x) = c\}$$

を**ファイバー** *(fiber)* といい，その連結成分を**断面** *(section)* という．商空間

$$X/\!\sim \quad (x \sim y \Leftrightarrow x, y \in X \text{ は } f \text{ の同じ断面に属する})$$

を**レーブ空間** *(Reeb space)* という．

モース＝スメール複体の多値拡張はまだないと思われるが，ベクトル場の分析は行われている．

定義 2.9（[6, Definition 6]） 写像 $f = (f_1, \dots, f_m) : X \to \mathbb{R}^m$ に対して，曲線 $\varphi : \mathbb{R} \to X$ で

$$\frac{\partial}{\partial t} f_i(\varphi(t)) < 0 \quad \text{for all } t \in \mathbb{R}, i = 1, \dots, m$$

を満たすものを**許容曲線** *(addmissible curve)* という．

明らかに，パレート臨界点を通過する許容曲線は存在しない．

定義 2.10（[6, Definition 7]） パレート臨界点 x が**安定** *(stable)* であるとは，$x \in X$ の任意の近傍 V_x に対して，$x \in V_x$ のある近傍 U_x が存在して，許容曲線 $\varphi : \mathbb{R} \to X$ のうち

$$\varphi(0) \in U_x$$

であるものすべてが

$$\varphi(t) \in V_x \quad \text{for all } t \geq 0$$

を満たすことをいう．安定パレート臨界点集合を

$$\theta_{\mathrm{S}} := \{x \in \theta \mid x : \mathrm{stable}\}$$

で表す．

パレート臨界点の安定性は目的関数の 2 階微分を調べればわかる．多目的では，2 階微分は不変的に定義することはできない．すなわち，

$$(D^2 f(x))(v, w) := \begin{pmatrix} {}^t v H f_1(x) w \\ \vdots \\ {}^t v H f_m(x) w \end{pmatrix}$$

で定義される写像 $D^2 f(x) : (T_x X) \times (T_x X) \to T_{f(x)}\mathbb{R}^m$ は局所座標の取り方に依存している．しかし，これを導関数の核 $\ker Df(x)$ に制限して，余核 $\mathrm{coker} Df(x) = T_{f(x)}\mathbb{R}^m / \operatorname{Im} Df(x)$ に値をとる対称双線形形式だと思えば，この形式は不変的に定義できる．

定義 2.11（[6, p. 468]） 余核への射影を

$$\pi : T_{f(x)}\mathbb{R}^m \to T_{f(x)}\mathbb{R}^m / \operatorname{Im} Df(x) = \mathrm{coker} Df(x)$$

とおき，

$$\pi \circ D^2 f(x)|_{\ker Df(x)} : \ker Df(x) \times \ker Df(x) \to \mathrm{coker} Df(x)$$

とする．

$$H_x(w) := \pi \circ D^2 f(x)|_{\ker Df(x)}(w, w)$$

を ***2* 階内在的導関数** *(second intrinsic derivative)* という．

$D^2 f(x)$ を $\ker Df(x)$ へと制限することは，以下のような意味をもつ．臨界点の近傍において許容曲線が引き込まれるか離れるかを調べるとき，パレート臨界点集合に平行な方向に関して何が起こっているかは問題ではなく，その直交空間こそが臨界点の安定性を決定づけている．

以下では，$\mathrm{corank} Df(x) = 1$，すなわち，$\mathrm{rank}\, Df(x) = m - 1$ を仮定する[*3]．このとき，H_x は 1 次元ベクトル空間に値をとる．$x \in \theta$ では $\mathrm{Im}\, Df(x) \cap \mathrm{Pos} = \emptyset$ なので，$\mathrm{coker} Df(x)$ は 1 次元であり，標準的な向きをもつ．このとき，H_x を**一般化ヘシアン** *(generalized Hessian)* とよび，その正定値性や指数を定義することができる．

定理 2.7（[6, Proposition 8]）

$$\partial\theta = \{x \in \theta \mid \mathrm{Im}(Df(x)) \cap \{\mathrm{Cl}(\mathrm{Pos}) \setminus \{0\}\} \neq \emptyset\}$$

とし，$x \in \theta, x \notin \partial\theta, \mathrm{corank} Df(x) = 1$ とおく．このとき，以下が成り立つ．

(a) もし一般化ヘシアン H_x が正定値ならば，$x \in \theta_{\mathrm{S}}$ である．

(b) $\lambda := (\lambda_1, \ldots, \lambda_m)$ を 1 次の条件のものとすると，正のスカラー倍を除いて

$$H_x = \sum_i \lambda_i D^2 f_i(x) \quad \text{on} \ \ker Df(x)$$

が成り立つ．

2.3 両者の関係

2.3.1 パレート点が X の内部にあるケース

定理 2.8（包含関係） パレート臨界点において非退化な一般化ヘシアンをもつ写像 $f : X \to \mathbb{R}^m$ $(\dim X \geq m)$ について，以下が成り立つ．

$$\theta_{\mathrm{O}} \subseteq \theta_{\mathrm{L}} = \theta_{\mathrm{S}} \subseteq \theta \subseteq \Sigma$$

証明 2 定義より $\theta_{\mathrm{O}} \subseteq \theta_{\mathrm{L}}$ と $\theta_{\mathrm{S}} \subseteq \theta \subseteq \Sigma$ は自明．背理法により $\theta_{\mathrm{L}} = \theta_{\mathrm{S}}$ を示す．

もし $\theta_{\mathrm{L}} \not\subseteq \theta_{\mathrm{S}}$ ならば，ある点 $x \in \theta_{\mathrm{L}} \setminus \theta_{\mathrm{S}}$ が存在して，x は退化している．これは非退化な一般化ヘシアンをもつと仮定したことと矛盾する．

もし $\theta_{\mathrm{L}} \not\supseteq \theta_{\mathrm{S}}$ ならば，ある点 $x \in \theta_{\mathrm{S}} \setminus \theta_{\mathrm{L}}$ が存在して，x のどの近傍にも f を改善する点が存在することになる．これは，安定パレート臨界点の指数が 0 であることと矛盾する． □

多目的最適化では，（局所）パレート点を具体的に求めることが問題となる．そのために様々な解法が研究されている．以下では，その代表的な解法と特異点論がどのような関係にあるかを述べる．

例 7（多目的勾配法） 多目的最適化問題に対する勾配法は，多目的勾配法とよばれる．多目的勾配法の基本的な枠組みは以下のようになり，$\alpha_t = (\alpha_{1t}, \ldots, \alpha_{mt}) \in \mathrm{Pos}$ の決め方によって様々な亜種がある [3].

1. 初期点 $x_0 \in X$ を適当に選ぶ.
2. 点列 $x_1, x_2, \ldots$ を以下のように決める.

$$x_{t+1} = x_t - \sum_i \alpha_{it} \nabla f_i(x_t). \tag{2}$$

[*3] この仮定は rank assumption とよばれ，ジェネリックな性質である.

多目的勾配法の収束先はパレート臨界点であるから，必ずしも大域的な最小点が得られるとは限らない．そこで，異なる初期点から再度探索を行うリスタートが行われる．

式 (2) からわかるように，系列 $x_0, x_1, \ldots$ は許容曲線の近似である．許容曲線がどのパレート臨界点に流れ込むかを類別する方法はまだ開発されていないと思われる．単目的最適化のケースから類推すれば，モース＝スメール複体の写像バージョンが必要になるはずである．この理論と計算アルゴリズムを確立すれば，効率的なリスタート戦略が開発できる可能性がある．

例 8（多目的遺伝的アルゴリズム） 多目的勾配法のリスタートのように，1 点 1 点順番に収束するまで探索するのは効率が悪いこともある．そこで，探索途中の情報を共有して複数の点を同時に探索する方法が考えられる．そのようなものの 1 つとして，多目的遺伝的アルゴリズムがある．多目的遺伝的アルゴリズムの基本的な枠組みは以下のようになり，ステップ 2 と 3 の作り方によって様々な亜種がある [10].

1. X の点をランダムに n 個生成し，それらの集合を X_0 とおく．$t = 0$ とする．
2. X_t の周辺に多数の点をランダムに生成し，Y_t とおく．
3. Y_t から n 点を f の値にしたがって選び，X_{t+1} とする．
4. $t := t + 1$ としてステップ 2 に戻る．

この方法は，**反鎖** *(antichain)*

$$C = \{c \in f(X) \mid d \in C \setminus \{c\} : c \not\leq d, c \not\geq d\}$$

に関するファイバーの合併

$$\bigcup_{c \in C} f^{-1}(c)$$

をサンプル X_t で近似し，$C \to f(\theta_{\mathrm{O}})$ へと追跡しているといえる．したがって，多目的遺伝的アルゴリズムがパレート点集合を見つけることができるかどうかは，サンプルの系列 $X_0, X_1, \ldots$ が写像のレーブ空間をうまく追跡できるかどうかにかかっている．

2.3.2 パレート点が X の境界にあるケース

実際の多目的最適化問題においては，パレート点が実行可能領域の境界にあるケースがしばしば見られる．このようなケースを特異点論的な視点から扱うためには，単目的最適化のケースから類推すれば，滑走分割モース理論の写像バージョンが必要になるものと思われる．しかし，これは理論的にも未開拓であると思われる．

参考文献

[1] H. Edelsbrunner, J. Harer, and A. K. Patel. Reeb spaces of piecewise linear mappings. In *Proceedings of the Twenty-fourth Annual Symposium on Computational Geometry*, SCG ’08, pages 242–250, New York, NY, USA, 2008. ACM.

[2] H. Edelsbrunner, J. Harer, and A. Zomorodian. Hierarchical Morse–Smale complexes for piecewise linear 2-manifolds. *Discrete & Computational Geometry*, 30(1):87–107, 2003.

[3] E. H. Fukuda and L. M. G. A. Drummond. A survey on multiobjective descent methods. *Pesquisa Operacional*, 34:585–620, 12 2014.

[4] M. Goresky and R. MacPherson. *Stratified Morse Theory*, volume 14 of *Ergebnisse der Mathematik und ihrer Grenzgebiete*. Springer Berlin Heidelberg, 1988.

[5] S. Ichiki and T. Nishimura. Distance-squared mappings. *Topology and its Applications*, 160(8):1005–1016, 2013.

[6] A. Lovison. Singular continuation: Generating piecewise linear approximations to Pareto sets via global analysis. *SIAM Journal on Optimization*, 21(2):463–490, 2011.

[7] K. M. Miettinen. *Nonlinear Multiobjective Optimization*, volume 12 of *International Series in Operations Research & Management Science*. Springer-Verlag, GmbH, 1999.

[8] J. Milnor. *Morse Theory*. Annals of Mathematics Studies. Princeton University Press, 1963.

[9] J. Nocedal and S. Wright. *Numerical Optimization*. Springer Series in Operations Research and Financial Engineering. Springer, 2nd edition, Jul 2006.

[10] A. Zhou, B.-Y. Qu, H. Li, S.-Z. Zhao, P. N. Suganthan, and Q. Zhang. Multiobjective evolutionary algorithms: A survey of the state of the art. *Swarm and Evolutionary Computation*, 1(1):32–49, 2011.

[11] 小林 重信. 実数値 GA のフロンティア. **人工知能学会論文誌**, 24(1):128–143, 2009.