最適化数学入門
—Karush–Kuhn–Tucker 条件の眺望—

奥野　貴之

受験数学で定番でもある関数の最小化問題もしくは最大化問題は，実は応用数学の基礎分野である最適化数学への入口とも言える．そして，最適化数学は高度情報化社会で意思決定をするためのツールとして欠かせないものとなっている．本稿では，お馴染みの変数が一つの関数の最小化問題から始めて最適化数学への簡単な入門を行う．とくに Karush–Kuhn–Tucker 条件と呼ばれる，最適解が満たすべき数学的性質の解説を行う．読者の知識としては，高校の数学Ⅱと数学 B における微分法とベクトルの知識があることを想定している．

キーワード：最適化，最適化問題，Karush–Kuhn–Tucker 条件

1. はじめに

今日複雑な世の中がうまく動いている背景には，さまざまな数学の問題が精錬されたテクニックによって絶え間なく解かれているという現実があります．その中の問題の一つに，高校生の皆さんが授業において既に中を覗き込んでいるものがあります．それが本稿で紹介する“最適化問題”であり，それらを取り扱う数学の分野を“最適化数学”と言います．コンピュータの能力の飛躍的向上と相まって現代社会で重要な意思決定の道具となっている分野です．本稿では，“最適化問題”を紹介するとともに，Karush–Kuhn–Tucker 条件という最適化問題において最も重要な数学的性質の一つを学んでいきたいと思います．

2. 最適化問題

最適化問題とは，制約となるいくつかの条件がある状況下で，ある目的にとって最適な選択肢を見つけ出す問題のことを言います．皆さんの身の回りにあったり，新聞・テレビなどで取り上げられる課題の多くが最適化問題として表現することが可能です．たとえば，出発地から目的地へ最短時間で到達する経路を見つける問題，数千種類ある株の中から見込まれる収益が最大となる株の組み合わせ（ポートフォリオと言います）を見つける問題，または限られた材料から製品を複数種類生産する際に，収益が最大となるように各製品の生産個数を決定する問題などが最適化問題として表すことができます．

また，皆さんにとって最適化問題は初めてのものではありません．実際，$y = x^2 + 2x \ (-1 \leqq x \leqq 3)$ の最小値，最大値を求める問題は高校で習う問題ですが，これも最適化問題です．最適化数学はこうした問題の延長線上にあるのです．

さて，こうした最適化問題は取りうる選択肢の集まりを S，目的の達成度を測る関数を f として下のように数式で表現することができます．この形の問題を“数理計画問題”とも呼び，集合 S を“**実行可能集合**”，関数 f を“**目的関数**”と呼びます．

$$\text{最小化} \quad f(x) \qquad \text{制約条件} \ x \in S$$

さて，これは実行可能集合 S に属する点（“実行可能解”と言います）の中で目的関数 $f(x)$ が最小となる“最適解” x を求めよと読みます[1]．先ほど述べた経路検索の問題を上の形として表すならば，S は出発地から目的地までの経路の集まり，$f(x)$ は経路 x に対する距離とおけばよいことになります．

ちなみに，実行可能集合 S は不等式や等式を用いて $S = \{x = (x_1, x_2, \ldots, x_n) \mid g(x) \leqq 0, h(x) = 0\}$ と表されることが多く[2]，その場合は集合の定義式である不等式や等式だけを用いて“最小化 $f(x)$ 制約条件 $h(x) = 0$, $g(x) \leq 0$”と書きます．本稿ではこの形を用います．

[1] $f(x)$ を最大化したい場合は $-f(x)$ を最小化することを考えてください．

[2] $x = (x_1, x_2, \ldots, x_n)$ とは n 個の実数の組で表される座標 x を意味します．$n = 2$ の場合はただの (x, y) 座標における点です．

おくの　たかゆき
東京理科大学工学部第一部経営工学科
〒 162–8601 東京都新宿区神楽坂 1–3

3. 具体例

具体的な問題を最適化問題として定式化してみましょう．

3.1 生産個数を決定する

ある工場で木材から鉛筆，机，椅子を生産する状況を考えましょう．各製品の1単位当たりの収益と重さは下の表に従うとします． このとき木材 1,000 kg から得られる収益を最大にするには鉛筆，机，椅子を何単位ずつ生産すべきでしょうか．鉛筆を x_1 ダース，机を x_2 台，椅子を x_3 脚生産するとしましょう．さて，x_1, x_2, x_3 が満たすべき条件としては

	収益	重さ
鉛筆（1 ダース）	a 円	0.05 kg
机（1 台）	b 円	50 kg
椅子（1 脚）	c 円	4 kg

(1) $0.05x_1 + 50x_2 + 4x_3 \leqq 1{,}000$（総重量の制約）
(2) x_1, x_2, x_3 は 0 以上の整数

で表現することができます．また目的としては，総収益 $ax_1+bx_2+cx_3$ の最大化，つまり $-ax_1-bx_2-cx_3$ の最小化です．そうすると解くべき最適化問題は次のように表現することができます．

$$
\begin{array}{ll}
\text{最小化} & -ax_1 - bx_2 - cx_3 \\
\text{制約条件} & (x_1, x_2, x_3) \in S; \\
 & S \text{ は (1), (2) を満たす} \\
 & (x_1, x_2, x_3) \text{ の集合}
\end{array}
$$

3.2 体積最大の直方体を求める

つぎに縦，横，高さの和が一定値 L をとる直方体の中で体積が最大のものを求める問題を考えましょう．直方体の 3 辺の長さを x_1, x_2, x_3 とします．このとき制約条件として，まず長さの和が一定より $x_1 + x_2 + x_3 = L$ が導かれます．加えて，長さは 0 以上ですから $x_1, x_2, x_3 \geqq 0$ です．最後に目的は体積 $x_1x_2x_3$ の最大化，すなわち $-x_1x_2x_3$ の最小化です．そうすると解くべき最適化問題は

$$
\begin{array}{ll}
\text{最小化} & -x_1x_2x_3 \\
\text{制約条件} & x_1 + x_2 + x_3 = L \\
 & x_1, x_2, x_3 \geqq 0
\end{array}
$$

となります．この問題の最適解は $(x_1, x_2, x_3) = (L/3, L/3, L/3)$ です．皆さん，理由を考えてみてください．

4. 最適解が満たすべき性質

前節の二つの具体例のように，制約条件と目的をどれほど達成できたかの目安を数学的に表現できたならば，多くの意思決定問題が最適化問題として表すことができます．このとき，得られた最適化問題の最適解が存在する[3]ならば，それをどのように見つけるのかは自然に生まれる疑問です．最適解を見つけるためには，候補となる点を実行可能集合の中から効率よく選び出す必要があります．そのためには最適解が満たすべき性質を利用するという手が考えられます．もちろん，最適解かどうかは，実行可能集合 S 上の他の点と目的関数の値を一つずつ比較すれば判定できます．ところが，3.2 節の直方体の例のように S の点の数が無限に考えられる場合はそんなことは不可能ですし，点の数が有限な場合でもその数が莫大な場合は現実的に難しいでしょう．したがって，闇雲に関数値を比較する方法とは異なるアプローチが必要です．それらは，考察対象となる最適化問題における目的関数や実行可能集合の特徴によって異なりますので，ここですべてを網羅することはできません．

そこで本稿では，最適化問題の中でも，制約条件としていくつかの 1 次不等式をもち，目的関数としてすべての点で微分可能な関数をもつ問題：

$$
\text{(P1)} \quad
\begin{array}{ll}
\text{最小化} & f(x) \\
\text{制約条件} & \displaystyle\sum_{j=1}^{n} a_{ij}x_j \leqq b_i \\
 & (i = 1, 2, \ldots, m)
\end{array}
$$

に限定します[4]．（a_{ij}，b_i $(i = 1, 2, \ldots, m,\ j = 1, 2, \ldots, n)$ は定数とします．）このタイプの問題は応用範囲が広く，とくに $f(x)$ が 1 次関数であるとき，この問題は線形計画問題と呼ばれ，非常に実用性が高いことで知られています．以降では，1 変数関数に関する最小化問題を皮切りに，この問題の最適解が満たすべき必要条件である Karush–Kuhn–Tucker 条件（KKT 条件）を解説していきます．

[3]「最小化 x 制約条件 $x \leq 0$」のように必ずしも最適解が存在するとは限りません．

[4] 1 次等式 $\sum_{j=1}^{n} c_jx_j = d$ が存在するときは，その式は 2 本の不等式 $\sum_{j=1}^{n} c_jx_j \geqq d$ と $\sum_{j=1}^{n} c_jx_j \leqq d$ とおき直せばよいことは明らかでしょう．

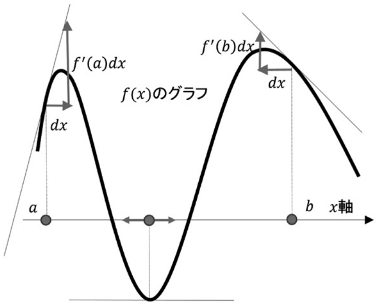

図 1 $f(x)$ のグラフと接線の様子

5. 微分可能な 1 変数関数の最小化問題

この節では，次のような微分可能な 1 変数関数（変数が一つの関数）の最小化問題を考えてみましょう．

$$\text{最小化} \quad f(x) \qquad \text{制約条件 } a \leqq x \leqq b$$

ここで a, b は $a \leqq b$ であるような定数とします．また，x^* を実行可能解とし，実行可能性を保持するような x^* からの変化量の集合，すなわち $a \leqq x^* + dx \leqq b$ となる dx の集合を $T(x^*)$ で表すとします．また $f'(x)$ を $f(x)$ の微分係数とします．

5.1 最適解であることの必要条件

x^* が最適解であることの必要条件を考えてみましょう．x^* が最適解ならば，$dx \in T(x^*)$ 方向に進むと $f(x)$ は増えなければいけません．一方，x^* における接線 $y = f'(x^*)(x - x^*) + f(x^*)$ の y の値は $f'(x^*)dx$ だけ変化します．ところが，x^* 付近では，接線の y の値の増減は $f(x)$ の増減を表しているのですから（図 1 参照）変化量 $f'(x^*)dx$ は 0 以上である必要があります．したがって，最適解であることの必要条件として以下が成り立ちます．

$$f'(x^*)dx \geqq 0 \quad (dx \in T(x^*)) \tag{1}$$

x^* の位置によって上の式を詳しく見てみましょう．

(i) $x^* = b$ の場合は，$dx \in T(x^*)$ として $dx \leqq 0$ のものしかとれません．よって (1) から $f'(x^*) \leqq 0$ です．

(ii) $a < x^* < b$ の場合は，dx として正負両方にとることができます．すると (1) から $f'(x^*) = 0$ です．

(iii) $x^* = a$ の場合は，(i) と同様にして $f'(x^*) \geqq 0$ が得られます．

以上から，最適解の候補として，区間中と区間の端で (i)，(ii)，(iii) が成り立っているような点を調べたらよいことがわかります．こうしたことは，教科書の内容を別の表現で書いただけのように見えますが，実は (1) は (P1) のような複雑な最適化問題に対しても考えることができるのです．詳しくは 6 節と 7 節で説明します．

5.2 性質 (1) の十分性について

図 1 で $x = a$ を見ればわかるように，(1) は x^* が最適解であることの十分条件ではありません．しかし，目的関数 f が下に凸な放物線である 2 次関数であれば，(1) は最適解であることの十分条件となることが知られています．より一般的に，目的関数 f が凸性[5]を満たせば (1) は十分条件となります．（証明は演習問題とします．）したがって，このような場合ならば (1) を満たす点は最適解であることが保証できます．

6. 2 変数 2 次関数の最適化問題

前節までは，1 変数関数の最適化問題で最適解が満たすべき性質を考えました．この節では変数を一つ増やして 2 変数の 2 次関数というものを考えます．x_1, x_2 を変数，a, b, c, d, e, f を定数とします．このとき下のような形をした関数を 2 変数 2 次関数と言います．

$$\frac{1}{2}ax_1^2 + bx_1x_2 + \frac{1}{2}cx_2^2 + dx_1 + ex_2 + f$$

では，具体的につぎのような 2 変数 2 次関数に関する最適化問題 (P2) を考えてみましょう．

$$\begin{array}{lll} & \text{最小化} & x_1^2 + x_2^2 - x_1x_2 - 5x_1 - 5x_2 \\ \text{(P2)} & \text{制約条件} & 2x_1 + 3x_2 \leqq 12, \\ & & x_1 \geqq 0,\ x_2 \geqq 0 \end{array}$$

この問題の最適解はどこにあるでしょうか．1 変数の場合と違い，一筋縄でいきそうにありませんね．イメージができるように $0 \leq x_1, x_2 \leq 10$ における目的関数のグラフを図 2 に，x_1x_2 座標における目的関数の等高線と実行可能集合を図 3 に示しておきます．ここで目的関数の等高線とは，地図における等高線と同じ概念で，関数値をその点における高さと見なして，同じ関数値をもつ点を結んだものです．図 2 と図 3 を見比べればわかるように，図 3 において楕円となっているのが $-20, -10, 0, 10, 20, 30, 40$ の関数値にそれぞれ相当する等高線です．全平面上で目的関数を最小化す

[5] $f(tx + (1-t)y) \leqq tf(x) + (1-t)f(y) \quad (0 \leqq t \leqq 1)$ がすべての x，y で成り立つという性質，直感的に言うとグラフ上の 2 点を結んだ線分よりもグラフが凹んでいる性質のことを言います．

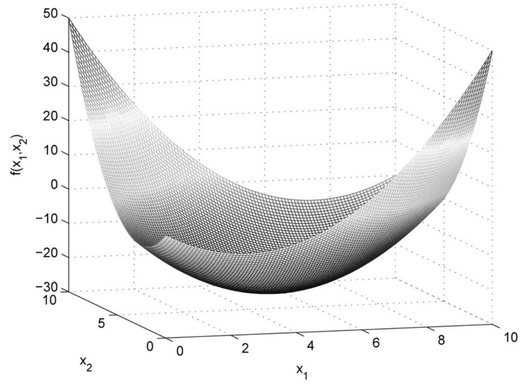

図 2 $f(x_1, x_2) = x_1^2 + x_2^2 - x_1x_2 - 5x_1 - 5x_2$ ($0 \leqq x_1, x_2 \leqq 10$) のグラフ

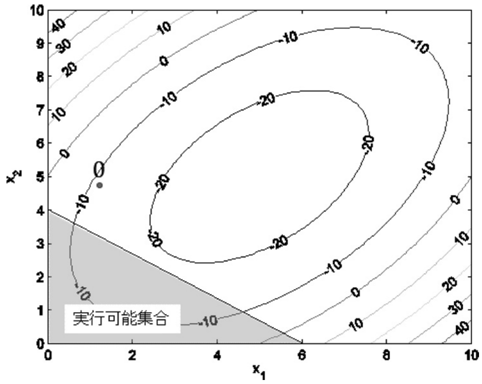

図 3 目的関数の等高線と実行可能集合 ($0 \leqq x_1, x_2 \leqq 10$)

る点は真ん中の楕円の中心部にあり，中心部に近づくにつれて目的関数値が低くなっていくのがわかります．そうすると図 3 から最適解は実行可能集合である直角三角形の斜辺の真ん中あたりにありそうです．

6.1 最適解であることの必要条件

最適解の位置や目的関数の増減のイメージはわきましたか？ では，$x^* = (x_1^*, x_2^*)$ を最適解として，x^* が満たすべき性質を導いてみましょう．そのために前節で考えた $T(x)$ と似たものを考えます．実行可能解 $x = (x_1, x_2)$ に対して，$(x_1 + dx_1, x_2 + dx_2)$ が実行可能解となるような 2 次元ベクトル $\vec{dx} = [dx_1, dx_2]$ の集合を考えます[6]．この集合を $T(x)$ と新たに表すことにします．加えて，この節では $T(x)$ から決まるつぎのベクトルの集合 $N(x)$ を定義します．

[6] 同じ成分 x_1, x_2 をもつベクトルと座標の違いを明らかにするために，ベクトルは [ ] で囲んで $[x_1, x_2]$ と表し，座標は ( ) で囲んで (x_1, x_2) と表すことにします．

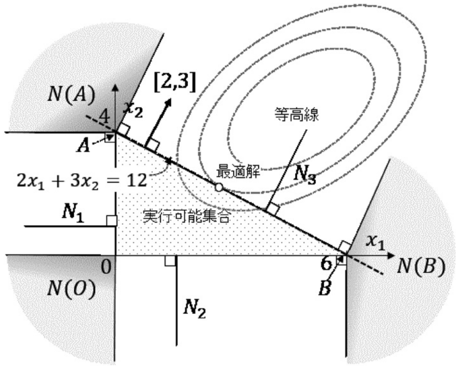

図 4 $N(x)$ の形状

表 1 各 x に対する $N(x)$

x の位置	$N(x)$
三角形内部	$\{[0, 0]\}$
頂点 O	$\{s[0, -1] + t[-1, 0] \mid s, t \geqq 0\}$
頂点 A	$\{s[2, 3] + t[-1, 0] \mid s, t \geqq 0\}$
頂点 B	$\{s[2, 3] + t[0, -1] \mid s, t \geqq 0\}$
辺 $OA \setminus \{O, A\}$	$N_1 = \{s[-1, 0] \mid s \geqq 0\}$
辺 $BO \setminus \{B, O\}$	$N_2 = \{s[0, -1] \mid s \geqq 0\}$
辺 $AB \setminus \{A, B\}$	$N_3 = \{s[2, 3] \mid s \geqq 0\}$

$$N(x) = \{\vec{e} \mid \text{すべての } \vec{d} \in T(x) \text{ について } \vec{d} \cdot \vec{e} \leqq 0 \text{ が成り立つ.}\}$$

ここで定義式中の“・”は内積記号です．したがって，$N(x)$ とは $T(x)$ に属するどのようなベクトル $\vec{d}$ とも，なす角が 90 度以上となるベクトル $\vec{e}$ の集合を意味しています．図 4 では，それぞれの実行可能解 x の位置における $N(x)$ の形を示しています．たとえば，実行可能集合である三角形の 3 頂点を $A(0, 4)$，$B(6, 0)$，$O(0, 0)$ としましょう．すると $A(0, 4)$ において $T(0, 4) = \{s\vec{AO} + t\vec{AB} \mid 0 \leqq s, t \leqq 1\}$ ですから，$N(0, 4)$（図 4 では $N(A)$ です）は $\{s[2, 3] + t[-1, 0] \mid s, t \geqq 0\}$ となります．ここで $[2, 3]$ と $[-1, 0]$ は，直線 $2x_1 + 3x_2 = 12$ と $x_1 = 0$ に対する法線ベクトルの中でも実行可能集合の外側を向いたベクトルです．その他の実行可能解 x に対する $N(x)$ を表 1 にまとめておきます．

さて，最適解であるための必要条件として，前節の (1) の拡張でもある下の補題が成り立ちます．

補題 6.1. $x^* = (x_1^*, x_2^*)$ を問題 (P2) の最適解とする．このとき

$$-[2x_1^* - x_2^* - 5, 2x_2^* - x_1^* - 5] \in N(x^*) \qquad (2)$$

が成り立つ．

証明. $f(x_1,x_2)=x_1^2+x_2^2-x_1x_2-5x_1-5x_2$ とおき，$\vec{dx}=[dx_1,dx_2]\in T(x^*)$ を一つとる．このとき，容易に確かめられるようにすべての自然数 $n\geqq 1$ に対して $\vec{dx}/n=[dx_1/n,dx_2/n]\in T(x^*)$ が成り立つ．x^* が最適解なので

$$\begin{aligned}0&\geqq f(x^*)-f\left(x^*+\left(\frac{dx_1}{n},\frac{dx_2}{n}\right)\right)\\&=-\frac{1}{n}\left((2x_1^*-x_2^*-5)\,dx_1+(2x_2^*-x_1^*-5)\,dx_2\right)\\&\quad-\frac{1}{n^2}\left((dx_1)^2+(dx_2)^2-dx_1dx_2\right)\\&=-\frac{1}{n}[2x_1^*-x_2^*-5,2x_2^*-x_1^*-5]\cdot\vec{dx}\\&\quad-\frac{1}{n^2}((dx_1)^2+(dx_2)^2-dx_1dx_2)\end{aligned}$$

である．ここで右辺の最後の式の第 2 項を左辺に移項して両辺に n を掛けると

$$\begin{aligned}&-[2x_1^*-x_2^*-5,2x_2^*-x_1^*-5]\cdot\vec{dx}\\&\qquad\leqq\frac{1}{n}\left((dx_1)^2+(dx_2)^2-dx_1dx_2\right)\\&\qquad=\frac{1}{n}\left(\left(dx_1-\frac{1}{2}dx_2\right)^2+\frac{3}{4}\,(dx_2)^2\right)\end{aligned}$$

を得る．すると，$\vec{dx}\in T(x^*)$ は任意に選んだので，左辺（$=\alpha$ とおく）が 0 以下であることを証明すれば題意は示される．$(dx_1-dx_2/2)^2+3(dx_2)^2/4=0$ となるベクトル $\vec{dx}$ に関しては $\alpha\leqq 0$ が成り立つので，$(dx_1-dx_2/2)^2+3(dx_2)^2/4>0$ となる $\vec{dx}$ のみを考える．さて，今から $\alpha\leqq 0$ を背理法で示す．$\alpha>0$ と仮定すると上の式から $0<\alpha\leqq((dx_1-dx_2/2)^2+3(dx_2)^2/4)\,/n$ が任意の自然数 n について成り立つ．ところが，n を十分大きくとると $(dx_1^2+dx_2^2-dx_1dx_2)/n<\alpha$ となり，矛盾が生じる．よって $\alpha\leqq 0$ である．以上から題意が示された． □

では，補題 6.1 の意味を考えてみましょう．まずベクトル $-[2x_1^*-x_2^*-5,2x_2^*-x_1^*-5]$ は唐突に出てきたように見えますが，証明からわかるように 2 点間の目的関数値の差に由来する自然なものです．次節でも書きますが，このベクトルは目的関数の勾配ベクトルの逆向きのベクトルに相当するものであり，5 節での $-f'(x^*)$ に対応しています．その方向と内積が正であるような向きに進むと関数値が下がる性質があります．逆に，内積が負であるような方向に進めば関数値は上がってしまいます．図 5 はいくつかの点におけるベクトル $-[2x_1-x_2-5,2x_2-x_1-5]$ を等高線図に

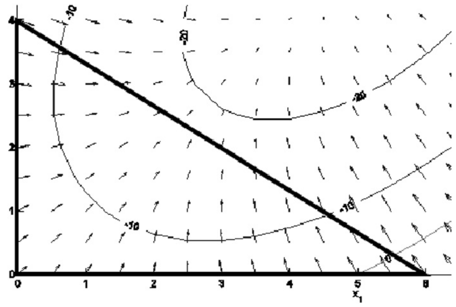

図 5　各点におけるベクトル $-[2x_1-x_2-5,2x_2-x_1-5]$

プロットしたものです．ある点から矢印の方向に進めば，その点の楕円の内部に入っていく様子から，目的関数値が下がる方向だということが観察できます．したがって，このようなベクトルが $N(x^*)$ に含まれているということは，x^* から実行可能集合のどの点に進んでも関数値が増大することを意味します．このことから (2) は x^* が最適解ならば自然に成り立つ性質であることがわかります．

6.2　条件 (2) を満たす x^* を求める

では，(2) を満たす x^* を実際に求めてみましょう．その前に，つぎのように関数と集合を定義しておきます．

$$\begin{aligned}&\vec{g}_*=[2x_1^*-x_2^*-5,2x_2^*-x_1^*-5],\\&g_1(x)=-x_1,\ g_2(x)=-x_2,\\&g_3(x)=2x_1+3x_2-12\\&\vec{v}_1=[-1,0],\ \vec{v}_2=[0,-1],\ \vec{v}_3=[2,3]\\&I(x)=\{i\mid g_i(x)=0,\ i\in\{1,2,3\}\}\end{aligned}$$

ここで $\vec{v}_i$ は直線 $g_i(x)=0$ の法線ベクトルの中で実行可能集合の外側に向いたベクトルであることに注意してください．さて，x^* の位置によって (2) をもう少し詳しく見ることができます．たとえば x^* がもし頂点 A にいるならば $I(x^*)=\{1,3\}$ であり，$g_1(x^*)=g_3(x^*)=0$，$g_2(x^*)<0$ です．また，表 1 から (2) の式は，ある $\lambda_1,\lambda_3\geqq 0$ が存在して $-\vec{g}_*=\lambda_1\vec{v}_1+0\vec{v}_2+\lambda_3\vec{v}_3$ と表せることを意味します．他の x^* の位置についても同様です．これらはまとめて下のようにスッキリと書くことができます．

x^* が最適解ならば，ある $\lambda_1,\lambda_2,\lambda_3$ が存在して

$$\begin{aligned}&\vec{g}_*+\sum_{i=1}^{3}\lambda_i\vec{v}_i=0,\\&g_i(x^*)=0,\ \lambda_i\geqq 0\ (i\in I(x^*)),\\&g_i(x^*)<0,\ \lambda_i=0\ (i\notin I(x^*))\end{aligned}$$

が成り立つ．

上の式を問題 (P2) に対する“Karush–Kuhn–Tucker 条件”あるいは“KKT 条件”と言い，λ_i を制約 $g_i(x) \leqq 0$ に関する“Lagrange 乗数”と言います．

KKT 条件は最適解を見つけ出すうえでとても有効な条件として知られており，最適化問題を解くための手法の多くがこの条件に基づいて開発されています．これを利用して，x^* を求めてみましょう．ずるいようですが，図 3 から $I(x^*) = \{3\}$ のようです．これを利用して上の KKT 条件を書き直すと $[2x_1^* - x_2^* - 5, 2x_2^* - x_1^* - 5] + \lambda_3[2,3] = 0, 2x_1^* + 3x_2^* - 12 = 0, \lambda_1 = \lambda_2 = 0, \lambda_3 \geqq 0, x_1^* \geqq 0, x_2^* \geqq 0$ となります．これらを解くと，$x_1^* = 99/38, x_2^* = 43/19, \lambda_1 = \lambda_2 = 0, \lambda_3 = 39/38$ が得られます．

6.3　KKT 条件の十分性について

最適解であることの必要条件として KKT 条件を導きましたが，1 変数の場合と同様，目的関数が凸性を満たすならば KKT 条件を満たす点は最適解であることが知られています．問題 (P2) の目的関数に関しても凸性が成り立っており，上で求めた解は問題 (P2) の最適解ということが保証できます．

7.　多変数関数に関する最適化問題

この節では，4 節における最適化問題 (P1) に対する KKT 条件を述べます．そのために

$$
\begin{aligned}
&\nabla f(x) = \left[\frac{\partial f(x)}{\partial x_1}, \frac{\partial f(x)}{\partial x_2}, \ldots, \frac{\partial f(x)}{\partial x_n}\right], \\
&I(x) = \left\{ i \;\middle|\; \sum_{j=1}^{n} a_{ij}x_j = b_i, i \in \{1, 2, \ldots, m\} \right\}, \\
&\vec{a}_i = [a_{i1}, a_{i2}, \ldots, a_{in}] \;\; (i = 1, 2, \ldots, m)
\end{aligned}
$$

とおきます．ここで第 1 行目のベクトルは関数 $f(x)$ の勾配ベクトルと呼ばれます．$\partial f(x)/\partial x_i$ は x_i に関する偏微分と言い，$f(x)$ を x_i のみを変数とする関数とみなしたときの微分係数を表します．$n = 1$ のときは $\nabla f(x) = f'(x)$ となります．また $n = 2$ として前節の目的関数 $f(x) = x_1^2 + x_2^2 - x_1x_2 - 5x_1 - 5x_2$ を考えると $\nabla f(x) = [2x_1 - x_2 - 5, 2x_2 - x_1 - 5]$ となります．

さて，最適化問題 (P1) に対する KKT 条件として以下が成り立ちます．

定理 7.1． $x^* = (x_1^*, x_2^*, \ldots, x_n^*)$ を問題 (P1) の最適解とする．このとき，

$$
\begin{aligned}
&\nabla f(x^*) + \sum_{i=1}^{m} \lambda_i \vec{a}_i = 0, \\
&\sum_{j=1}^{n} a_{ij}x_j^* - b_i = 0,\ \lambda_i \geqq 0\ (i \in I(x^*)), \\
&\sum_{j=1}^{n} a_{ij}x_j^* - b_i < 0,\ \lambda_i = 0\ (i \notin I(x^*))
\end{aligned}
$$

となる $\lambda_1, \lambda_2, \ldots, \lambda_m$ が存在する．逆に，目的関数が凸性を満たすならば，上が成り立てば x^* は (P1) の最適解である．

8.　おわりに

本稿では最適化数学入門として，最適化問題において最適解が満たすべき KKT 条件を紹介しました．本稿では，問題 (P1) を中心に考えましたが，より複雑な形をした最適化問題まで拡げて KKT 条件を考えることができます．こうした KKT 条件を含む最適化理論の入門書として [1] を，本格的な専門書として [2] をあげておきます．

また 6 節の問題では，手作業で KKT 条件が成り立つ解と Lagrange 乗数を求めました．しかし，変数や制約条件の個数が増大するにつれて手作業で最適化問題を解くのは難しくなっていきます．現代社会で解かれている最適化問題の中にはそれらが数万単位のものも少なくありません．そうした問題をコンピュータで効率的に解くための強力な数値解法も多く開発されています．数値解法に関する専門書として [3] をあげておきます．

最後になりましたが，本稿をきっかけに一人でも最適化数学に興味をもっていただけたら幸いです．

参考文献

[1] 福島雅夫，『数理計画入門』，朝倉書店，1996.
[2] 福島雅夫，『非線形最適化の基礎』，朝倉書店，2005.
[3] 矢部博，『工学基礎 最適化とその応用』，数理工学社，2006.

