

5 制約付き最適化問題

制約付きの最適化問題を扱う. 今までと同様に最小化問題で主に説明するが最大化問題も同様である. また変数を一般に n 個として扱う.

C を数ベクトル空間 \mathbb{R}^n の部分集合とする. $x = (x_1, x_2, \dots, x_n)$ のように書く. 以下のような問題を**制約付き最小化問題**と呼ぶ.

$$(P) \quad \begin{aligned} & \text{最小化 } f(x) \\ & \text{制約 } x \in C \end{aligned}$$

最小化問題 (P) において, 上記のような集合 C を**実行可能領域**, C の点を**実行可能解**と呼ぶ. 制約なし最小化問題では, 最小解を任意の数ベクトルから探してきたが, 制約付き最小化問題では, C に入っている x の中で, 目的関数 $f(x)$ を最小にするものを探す.

例 12. 辺の長さが一定な長方形の中で, 面積が最大になるのはどのような場合か?

この問題は次のように定式化できる. 辺の長さを x と y とすると, 面積は xy になる. 辺の長さの合計を $a > 0$ とすると, 問題は

$$\begin{aligned} & \text{最大化 } xy \\ & \text{制約 } (x, y) \in C := \{(x, y) \mid x + y = a, x \geq 0, y \geq 0\} \end{aligned}$$

となる.

不等式を用いて最大解を求める. (x, y) を C 内の任意の点とする. すると, $x, y \geq 0$ なので, 相加相乗平均の不等式より

$$\sqrt{xy} \leq \frac{x+y}{2} = \frac{a}{2}$$

となり, $xy \leq a^2/4$ を満たす. 等号が成立するのは $x = y$ のときなので, 最大解は $(x, y) = (a/2, a/2)$ となり, これは正方形を表す.

補足. 相加相乗平均の不等式は一般に n 変数に対しても成り立つ. 例えば, 3 変数の場合を用いれば, 辺の和が一定の直方体の中で体積が最大のものも求まる.

例 13. 円 $x^2 + y^2 = 1$ 上の点で, 点 $(1, 2)$ に一番近い点はどこか?

一般に点 (x, y) と $(1, 2)$ の距離は $\sqrt{(x-1)^2 + (y-2)^2}$ と書ける. この問題は, 円上の点の中でこの距離を最小にする点を探すが, 距離を 2 乗した $(x-1)^2 + (y-2)^2$ を最小にするものを探しても一緒なので,

$$\begin{aligned} & \text{最小化 } f(x, y) := (x-1)^2 + (y-2)^2 \\ & \text{制約 } (x, y) \in C := \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\} \end{aligned}$$

という問題を解けばよい.

(x, y) を C の点とする. いま, C 上の点は $x^2 + y^2 = 1$ となるので, 目的関数を展開しこの関係式を代入すると, $f(x, y) = 1 - 2x - 4y + 3 = -2x - 4y + 4$ となる. ここで, 定数項を無視した部分にシュワルツの不等式を適用して

$$\begin{aligned} |-2x - 4y| &= | \langle (-2, -4), (x, y) \rangle | \leq \|(-2, -4)\| \| (x, y) \| \\ &= \sqrt{(-2)^2 + (-4)^2} \sqrt{x^2 + y^2} = \sqrt{20} \cdot 1 = 2\sqrt{5} \end{aligned}$$

を得る. これより特に, $f(x, y) = -2x - 4y + 4 \geq -2\sqrt{5} + 4$ を得る

シュワルツの不等式では, 等号は二つのベクトルの方向が同じ時に成り立つので, ある数 t が存在して, $(x, y) = t(-2, -4)$ が成り立つような (x, y) で, C 上の点になるものを探す. まず, t を消去すると, $y = 2x$ を得る. この関係を満たす点で, C 内の点になるのは, $x^2 + (2x)^2 = 1$ を満たすので, $(x, y) = (\pm 1/\sqrt{5}, \pm 2/\sqrt{5})$ を得る.

この二つの点での目的関数値を考えると, $f(1/\sqrt{5}, 2/\sqrt{5}) = -10/\sqrt{5} + 4 = -2\sqrt{5} + 4$ となるので, 最小解は $(x, y) = (1/\sqrt{5}, 2/\sqrt{5})$ である.

6 最適性条件

6.1 等式制約が一つの場合

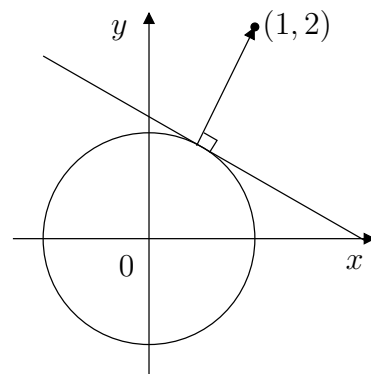
上記の例のように, 不等式を用いて綺麗に解ける問題もあるが, 制約なしの最適化問題で行ったように微分法を用いた一般的な解法を学ぶ.

例 13 で得た解を幾何的に解釈してみよう.

まず, テーラー展開を考えると, 点 (\bar{x}, \bar{y}) において, 方向 $-\nabla f(\bar{x}, \bar{y})$ は (ゼロベクトルでなければ) $f(\bar{x}, \bar{y})$ を減少させる方向である. 例 13 で, この「減少方向ベクトルが C のその点における接線と直交する方向を指す」ことが確認できる.

例 13 で, 点 $(1, 2)$ までの距離を最小にする円上の点は $(1/\sqrt{5}, 2/\sqrt{5})$ であることを示した. 実際, この点で減少方向ベクトルは $-\nabla f(1/\sqrt{5}, 2/\sqrt{5}) = (2(1/\sqrt{5} - 1), 2(2\sqrt{5} - 2)) = 2(1/\sqrt{5} - 1, 2\sqrt{5} - 2)$ となる.

これは $(1/\sqrt{5}, 2/\sqrt{5})$ と $(1, 2)$ を通る直線と同じ方向を指す. 距離が最短であることを考えると, C のこの点での接線と, この点と点 $(1, 2)$ を通る直線は直交しているはずである. よって上記の主張は正しそうだ.



定理 8 (ラグランジュの乗数法). 最小化問題

$$\text{最小化 } f(x)$$

$$\text{制約 } g(x) = 0$$

を考える (実行可能領域を集合ではなく集合を制約する式だけで書くときもある). \bar{x} を局所最小解とする. $\nabla g(\bar{x}) \neq (0, \dots, 0)$ ならば, ある数 λ が存在して,

$$\begin{aligned}\nabla f(\bar{x}) &= \lambda \nabla g(\bar{x}) \\ g(\bar{x}) &= 0\end{aligned}$$

が成り立つ.

Proof. 2変数の場合を証明する. $\bar{x} = (\bar{x}_1, \bar{x}_2)$ を局所最適解とする. $g(\bar{x}) = 0$ は制約を満たすことから成り立つ. いま, 関数 $u(t) = (x_1(t), x_2(t))$ で, $t = 0$ の近くで微分可能で, $u(0) = \bar{x}$, $g(u(t)) = 0$ を満たすものが存在するとしよう. すると定理の主張は簡単に示せる. 実際, \bar{x} は局所最小解で, $u(t)$ は制約を満たしているので, $f(u(t)) \geq f(\bar{x}) = f(u(0))$ が成り立つ. よって, 制約なしの最適化問題

$$\text{最小化 } f(u(t)), t \in \mathbb{R}$$

は $t = 0$ で局所最小値を取る. したがって, 合成関数 $f \circ g$ の $t = 0$ における微分 $\left. \frac{d}{dt} f(u(t)) \right|_{t=0} = \nabla f(\bar{x}) u'(0) = 0$ が成り立つ. また, $t = 0$ の近くで $g(u(t)) = 0$ なので, 微分も $\left. \frac{d}{dt} g(u(t)) \right|_{t=0} = \nabla g(\bar{x}) u'(0) = 0$ を満たす. したがって, $\nabla f(\bar{x}), \nabla g(\bar{x})$ 2次元ベクトルで一つのベクトル $u'(0)$ に直交しているので, ある定数 λ が存在して, $\nabla f(\bar{x}) = \lambda \nabla g(\bar{x})$ が成り立つ.

さて, 上記のような $u(t)$ の存在を示そう. これには陰関数定理を用いる. $\nabla g(\bar{x}) \neq 0$ より, 特に $g_{x_2}(\bar{x}) \neq 0$ とする. すると, 陰関数定理より関数 φ で, \bar{x}_1 の近くで微分可能で, $g(x_1, \varphi(x_1)) = 0$ (x_1 は \bar{x}_1 に近い任意の数), $\varphi(\bar{x}_1) = \bar{x}_2$ を満たすものが存在する. よって, $t = 0$ の近くで, $u(t) = (\bar{x}_1 + t, \varphi(\bar{x}_1 + t))$ と定義すると上記を満たす関数 u を得る. \square

補足. 3変数以上の場合には, 上の議論を工夫して $\langle \nabla g(\bar{x}), d \rangle = 0$ を満たす任意のベクトル d に対して, $u'(0) = d$ となるような $u(t)$ を見つける必要がある. すると, $\nabla g(\bar{x})$ に直交するすべてのベクトルに $\nabla f(\bar{x})$ が直交することが言えて, $\nabla f(\bar{x}) = \lambda \nabla g(\bar{x})$ が成り立つ.

補足. 証明を見れば分かるように最大化問題に対しても同様のことが言える.

定理の式を満たすような点 \bar{x} を **実行可能領域に関する停留点** と呼ぶ. 通常の停留点と同じように, 実行可能領域に関する停留点であっても最適解にならないものは存在する. 省略しても誤解がない場合は, 単に停留点と呼ぶ.

証明はしないが (実数の性質を使う) 最適解の存在定理を一つ挙げておく.

定理 9. 目的関数が連続で, 実行可能領域が有界閉集合ならば, 最適化問題は最小解と最大解を持つ.

制約付き最適化問題で特に等式制約を持つ問題は, 実行可能領域が有界閉集合になる場合が多い. 最適解の存在があらかじめ分かっていると議論が楽になり, 1階微分の情報 (基本最適性条件) のみで, 最適解を探せることも多い.

6.2 等式制約付き最小化問題の解き方

例 14.

$$\begin{aligned} \text{最小化 } f(x, y) &:= x - y \\ \text{制約 } g(x, y) &:= 2x^2 + 3y^2 - 1 = 0 \end{aligned}$$

基本最適性条件より, 実行可能領域に関する停留点を求める. 定理 8 (直後の補足にも注意) の適用を目指す. まず, 目的関数と制約式の勾配ベクトルは, それぞれ $\nabla f(x, y) = (1, -1)$, $\nabla g(x, y) = (4x, 6y)$ となる. ここで, $g(x, y) = 0$ を満たす点には $\nabla g(x, y)$ を零ベクトルにする点はないので, 定理 8 が適用できる. 停留点を見つけるには,

$$\begin{cases} \nabla f(x, y) = \lambda \nabla g(x, y), \\ g(x, y) = 0 \end{cases}$$

を解けばよい. つまり

$$\begin{cases} 1 = \lambda 4x \\ -1 = \lambda 6y \\ 2x^2 + 3y^2 - 1 = 0 \end{cases}$$

を満たす $(x, y), \lambda$ を見つければ, その (x, y) が求める点になる (λ は存在すれば良く, 必ずしも計算する必要はない).

第一式より $4x \neq 0$ となるので, $\lambda = 1/(4x)$ として第二式に代入すると, $y = -2x/3$ を得る. これを第三式に代入すると, $x = \pm\sqrt{3/10}$ が求まる. このとき $y = \mp\sqrt{2/15}$ となる. よって停留点は $(x, y) = (\pm\sqrt{3/10}, \mp\sqrt{2/15})$ になる.

さて, ここで実行可能領域は楕円を表すので有界閉集合になり, 定理 9 よりこの問題は最小解を持つ. 最小解が存在することが分かれば, それは停留点の内のどれかになる (講義ノートの包含関係を表す図を参照). よって, $(x, y) = (\pm\sqrt{3/10}, \mp\sqrt{2/15})$ での目的関数 J の値を調べれば, $(-\sqrt{3/10}, \sqrt{2/15})$ が最小解であることが分かる.

例 15 (2 次形式の最大値, 最小値). $A = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$ を実対称行列として以下の最小化問題を考える.

$$\begin{aligned} (P) \text{ 最小化 } f(x, y) &:= \begin{bmatrix} x & y \end{bmatrix} A \begin{bmatrix} x \\ y \end{bmatrix} = ax^2 + 2bxy + cy^2 \\ \text{制約 } g(x, y) &= x^2 + y^2 - 1 \end{aligned}$$

最小値は A の最小固有値, 最小解はその固有ベクトルになる.

実行可能領域に関する停留点を求める. $g(x, y) = 0$ を満たす点で $\nabla g(x, y)$ は零ベクトルにならない. よって, 定理 8 を適用できる. 停留点を見つけるには,

$$\begin{cases} \nabla f(x, y) = \lambda \nabla g(x, y), \\ g(x, y) = 0 \end{cases}$$

を (x, y) , λ について解けばよい. いま $\nabla f(x, y) = (2ax + 2by, 2bx + 2cy)$, $\nabla g(x, y) = (2x, 2y)$ より, 第一式は

$$\begin{bmatrix} 2ax + 2by \\ 2bx + 2cy \end{bmatrix} = \lambda \begin{bmatrix} 2x \\ 2y \end{bmatrix}$$

なので, 停留点の式は

$$\begin{cases} \begin{bmatrix} a & b \\ b & c \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \lambda \begin{bmatrix} x \\ y \end{bmatrix}, \\ g(x, y) = 0 \end{cases}$$

となる. これは λ が A の固有値で, (x, y) がその固有ベクトルあり, かつ大きさが 1 であるということを示す. よって, 停留点は A の固有ベクトルとなる.

ここで, (\bar{x}, \bar{y}) を A の固有ベクトルとすると, $f(\bar{x}, \bar{y}) = \lambda(\bar{x}^2 + \bar{y}^2) = \lambda$ となるので, 停留点の中で f を最小にするのは, (\bar{x}, \bar{y}) が A の最小固有値に対する固有ベクトルのときであり, そのとき f の値は A の最小固有値と等しくなる.

6.3 等式制約が複数の場合

制約は複数ある場合もある.

定理 10. 最小化問題

最小化 $f(x)$

$$\text{制約 } g_1(x) = 0, g_2(x) = 0, \dots, g_m(x) = 0$$

を考える. \bar{x} を局所最小解とする. $\nabla g_1(\bar{x}), \dots, \nabla g_m(\bar{x})$ が一次独立ならば, ある数 $\lambda_1, \dots, \lambda_m$ が存在して,

$$\begin{aligned} \nabla f(\bar{x}) &= \lambda_1 \nabla g_1(\bar{x}) + \dots + \lambda_m \nabla g_m(\bar{x}) \\ g_1(\bar{x}) &= 0, \dots, g_m(\bar{x}) = 0 \end{aligned}$$

が成り立つ.

証明には次のような陰関数定理を用いる. 簡単のため変数の数が 3 つ, 等式の数 が 2 つの場合を記述する.

定理 11. g_1, g_2 を 3 変数関数, $g_1(\bar{x}) = 0, g_2(\bar{x}) = 0$ とする.

$$\text{rank} \begin{bmatrix} \nabla g_1(\bar{x})^T \\ \nabla g_2(\bar{x})^T \end{bmatrix} = 2$$

ならば, ある関数 φ_1, φ_2 で,

$$\begin{bmatrix} g_1(x_1, \varphi_1(x), \varphi_2(x)) \\ g_2(x_1, \varphi_1(x), \varphi_2(x)) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

を満たすものが存在する.

解説. g_1, g_2 が線形だとすると, $g_1(\bar{x}) = 0, g_2(\bar{x}) = 0$ は, ある 2×3 行列を用いて,

$$\begin{bmatrix} g_1(\bar{x}) \\ g_2(\bar{x}) \end{bmatrix} = A\bar{x} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

と書ける. また, $\text{rank} \begin{bmatrix} \nabla g_1(\bar{x})^T \\ \nabla g_2(\bar{x})^T \end{bmatrix} = \text{rank } A$ である. $\text{rank } A = 2$ の場合, 方程式の解の自由度が 1 になるので 1 つのパラメータを用いて解を表せる. この定理が言っているのは, g_1, g_2 が非線形の場合にも, 微分の階数が 2 であれば, 方程式の解は 1 つのパラメータと非線形関数を用いて表せる, ということである. \square

練習問題 3. 1. 実行可能領域に関する停留点とその点での目的関数の値を求め, 最小解を探せ.

(1)

$$\begin{aligned} \text{最小化 } f(x, y) &:= 2x - 3y \\ \text{制約 } g(x, y) &:= 2x^2 + y^2 - 1 = 0 \end{aligned}$$

(2)

$$\begin{aligned} \text{最小化 } f(x, y, z) &:= x - y + 2z \\ \text{制約 } g(x, y, z) &:= x^2 + y^2 + 2z^2 - 1 = 0 \end{aligned}$$

(3)

$$\begin{aligned} \text{最小化 } f(x, y) &:= x^3 + y^3 \\ \text{制約 } g(x, y) &:= x^2 + y^2 - 1 = 0 \end{aligned}$$

2. x, y が原点を中心とする半径 1 の円周上の点であるとき, $2x^2 - 6xy + 2y^2$ の最小値と最大値を求めよ.