最適化問題

（制約なしと制約付き）

2016/5/2

スタートアップゼミ＃３

B4　山本正太郎

はじめに

・僕自身の理解過程をスライドに入れたので冗長な部分があるかもしれません、ご容赦ください

・最適化問題について、制限なし、制限つき双方の場合について基礎的な手法の理解をすることを目指します

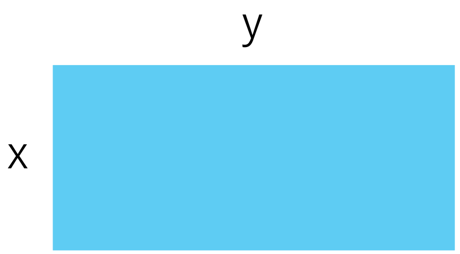
・問題と手法を結び付けてパターン化するだけの文系的な考えを脱し、より根本的な理解を目指します

・とはいえ、理解が甘い部分がございますのでバンバンご指摘ください

序：最適化問題とは？

例）

縦横の辺の長さの和が４となる長方形の中で、面積が最大になるのはどのような長方形か？



定式化すると

最大化　$f(x, y) = \ xy$

制約　$x + y = 4, x \geq 0, y \geq 0$

序：最適化問題とは？

最適化問題とは、関数を最小化、または最大化する問題である。

変数を選ぶ範囲になんらかの制約があるものを「制約付き」、変数の範囲に制約がないものを「制約なし」とよぶ。

まずは、「制約なし最適化問題」から導入する

1. 制約なし最適化問題

例）

最小化：　$f(x) = x^2 + 2x + 3$

制約条件：なし

$$f(x) = x^2 + 2x + 3$$
$$= (x+1)^2 + 2$$

より、すべてのxについて

$f(x) \geq f(-1) = 2$　（<u>大域最小解</u>）

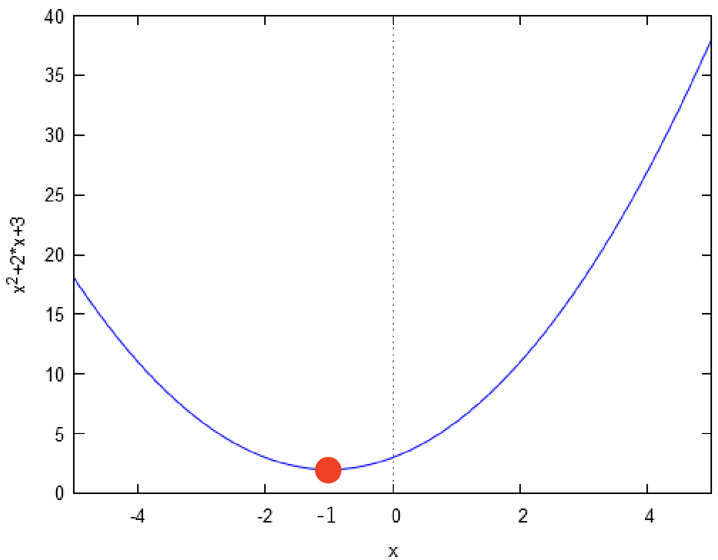


図１：f(x)のグラフ

1. 制約なし最適化問題

例)

最小化：$f(x) = x^3 - x$

図2より大域最小解は存在しない。

ただし、1/√3に十分近いxについて

$f(x) \geq f\left(\frac{1}{\sqrt{3}}\right) = \frac{-2}{3\sqrt{3}}$となる。

このように、局所的に最小になっている点を

局所最小解という。

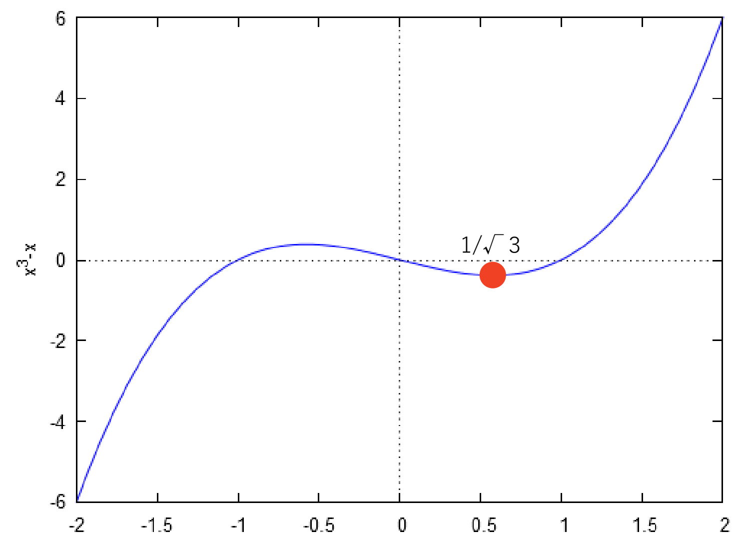
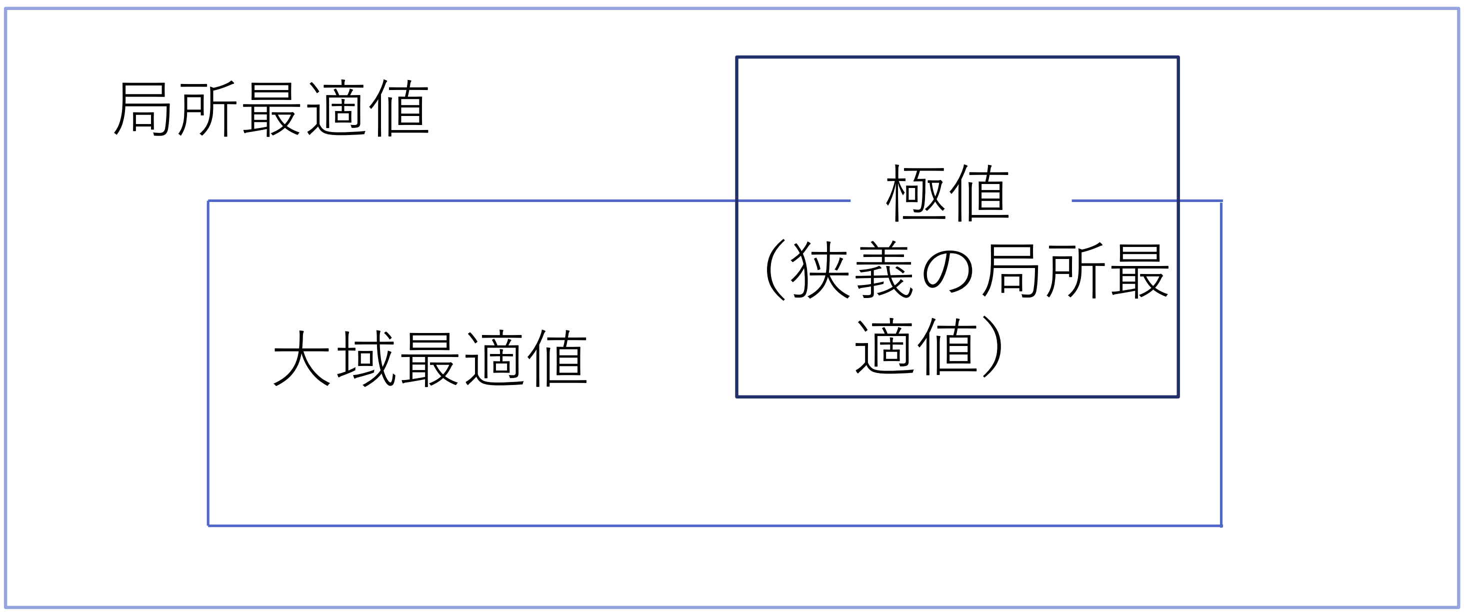


図２：f(x)のグラフ

1. 制約なし最適化問題

（極値と最適値の関係）



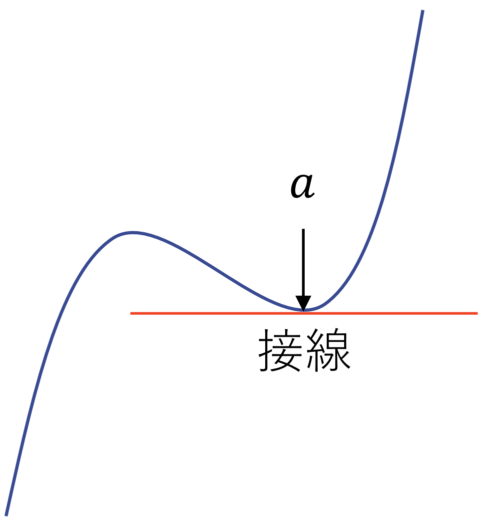
とりあえず、局所最適解を求める方法を考える

1-1 1次の最適性条件

（1変数関数について）

定理：1変数関数$f(x)$に対して、点aが局所最適解ならば、

$f'(a) = 0$となる。



1-1 １次の最適性条件

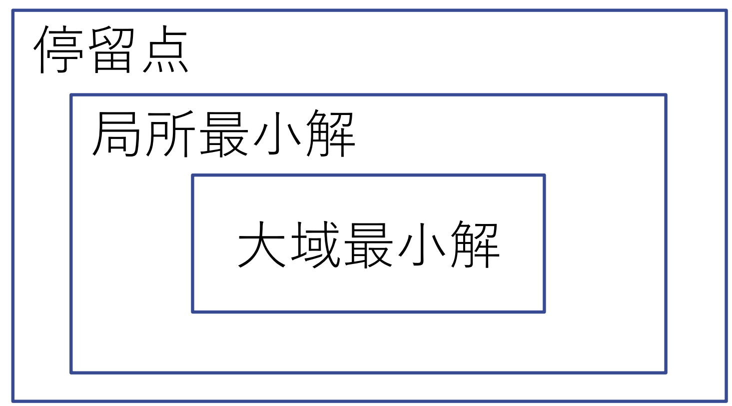
（多変数関数について）

定理１：点$\bar{x}$が局所最適解ならば、$\nabla f(\bar{x}) = \mathbf{0}$（零ベクトル）

定理２：点pが$\nabla f(p) = \mathbf{0}$を満たすとき、pをfの停留点と呼ぶ。

つまり、局所最適解は常に停留点である。
よって停留点をすべて見つければ最適解はその中にある。

停留点と最適解の関係は右図の通り。



1-1 １次の最適性条件

（停留点の導出）

最小化問題　$f(x,y)=x^3-3xy+y^3$

目的関数の勾配ベクトルは

$$\nabla f(x,y)=\begin{bmatrix}3x^2-3y\\3y^2-3x\end{bmatrix}$$

$\nabla f(x,y)=0$を解くと

$(x,y)=(0,0),(1,1)$

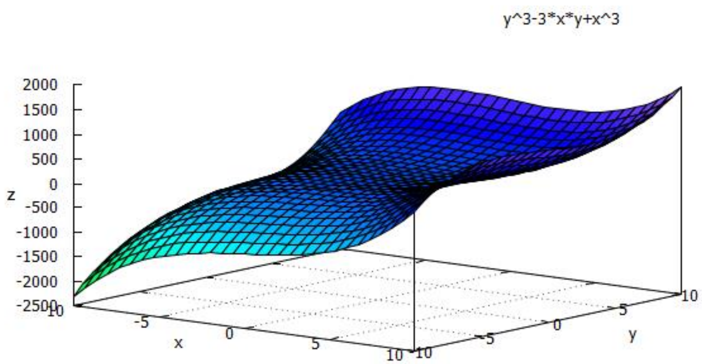


図３：f(x,y)のグラフ

問題は、「停留点であっても局所最適解とは限らない」こと。

1-2 ２次の最適性条件

停留点を求めた後に、どのようにして局所最適解を見つけるか

参考書によると…

やり方１．グラフをコンピュータで描画し、推測する

やり方２．ヘッセ行列を用いて判定する

疑問：ヘッセ行列ってなんだろう？

1-2 2次の最適性条件

<ヘッセ行列とは>

定義：２変数関数fと(x, y)に対して、

$\nabla^2 f(x, y) = \begin{bmatrix} f_{xx}(x, y) & f_{xy}(x, y) \\ f_{yx}(x, y) & f_{yy}(x, y) \end{bmatrix}$ をヘッセ行列と呼ぶ。

一般化すると

n 変数関数 $f(x_1, x_2, \dots, x_n)$ に対して,

ij成分が $\frac{\partial f}{\partial x_i \partial x_j}$であるような$n \times n$ 行列をヘッセ行列と言う。

例）$f(x, y) = x^3 - 3xy + y^3$について、ヘッセ行列は

$\nabla^2 f(x, y) = \begin{bmatrix} 6x & 2 \\ 2 & 6 \end{bmatrix}$ となる。

1-2 2次の最適性条件

定理１．（必要条件）

$\bar{x}$が局所最小解である ⇒ $\nabla f(\bar{x}) = 0$かつ$\nabla^2 f(\bar{x})$が半正定値

定理２．（十分条件）

$\nabla f(\bar{x}) = 0$かつ$\nabla^2 f(\bar{x})$が正定値⇒$\bar{x}$は局所最小解であり、そこで極小値をとる。

定理３．（否定）

$\nabla^2 f(\bar{x})$が不定値のとき、$\bar{x}$は局所最適解ではない。

1-2 ２次の最適性条件

＜行列の正値性について＞

Aを対称行列とすると、以下の主張が成り立つ

定理１.

Aが半正定値⇔Aの固有値がすべて0以上

定理２.

Aが正定値⇔Aの固有値が全て正

定理３.

Aが不定⇔Aが正と負の固有値を持つ

1-2 2次の最適性条件

任意の(x, y)についてヘッセ行列$\nabla^2 f(x, y)$が正定値⇒ fが狭義凸関数

（狭義凸関数とは、グラフ上の任意の2点を線分で結んだとき、その線分がグラフよりも上部にある関数のこと）

(x, y)に十分近い点(a, b)を考える。

$\nabla^2 f(a, b)$が正定値⇒(a, b)に近い(x, y)で$\nabla^2 f(x, y)$が正定値

⇒ fが「(a, b)の近くで狭義凸関数」

1-2 2次の最適性条件

１次の最適化条件と合わせると、

$\nabla f(a,b) = 0$かつ$\nabla^2 f(a,b)$が正定値

⇒ (a,b)はfの「局所的に狭義凸」な部分の底にある

したがって、(a,b)の近くで狭義凸であることから、点(a,b)の近くで極小値をとり、(a,b)は局所最小解である。

また、ヘッセ行列$\nabla^2 f(a,b)$が不定の場合、fは(a,b)の周りで凸関数でも凹関数でもない（捻じれている）。

1-2 2次の最適性条件

（２次の最適化条件の幾何学的理解）

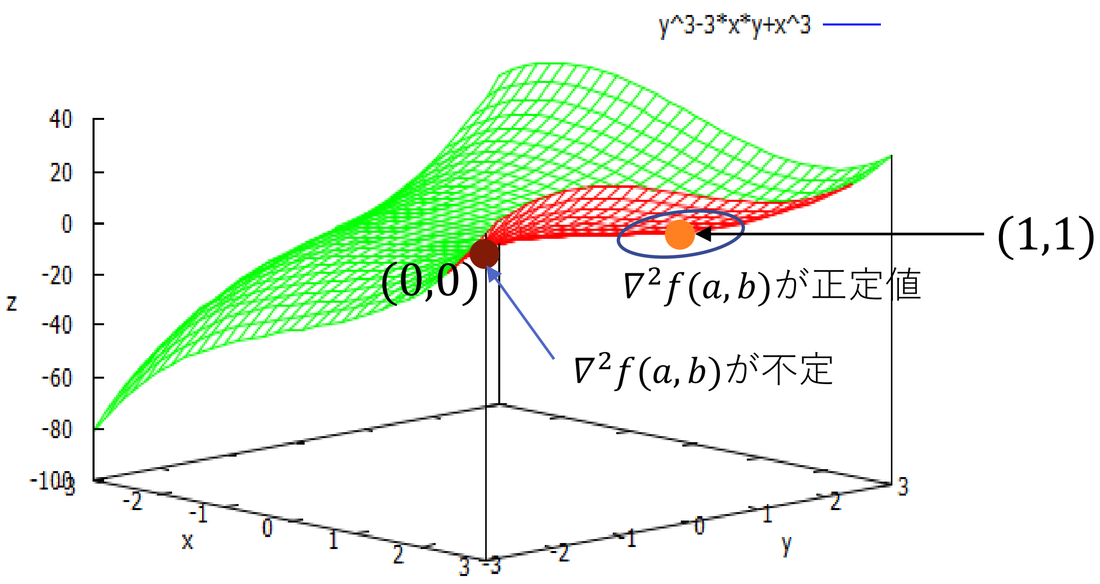
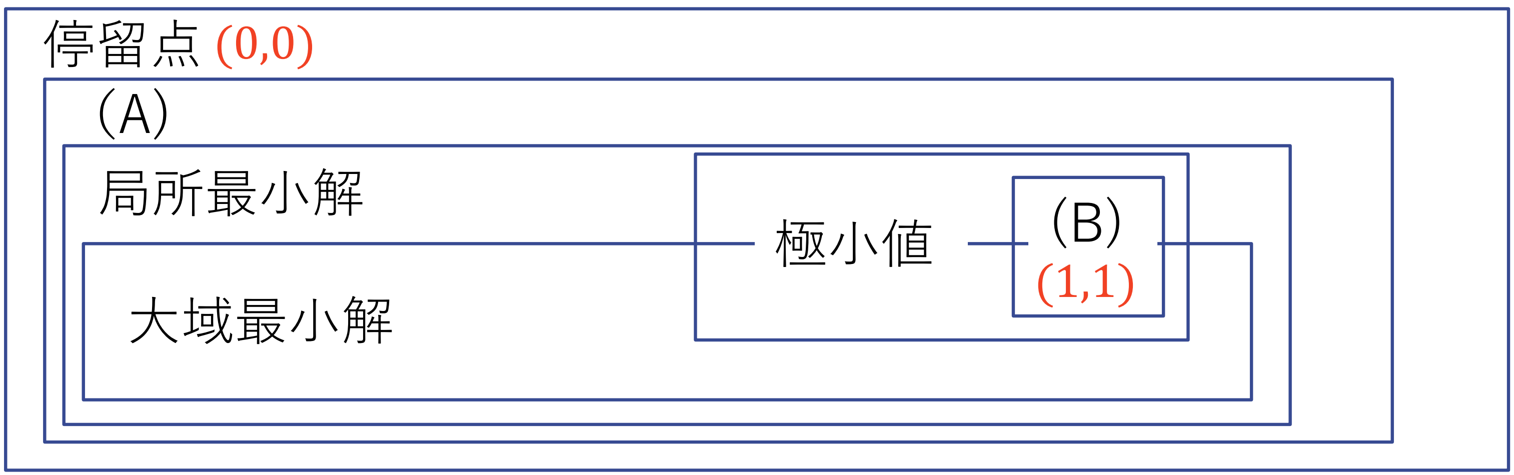


図5：$f(x,y) = x^3 - 3xy + y^3$のグラフ

1. 制約なし最適化問題

したがって、$f(x,y)=x^3-3xy+y^3$の停留点（0,0）と(1,1)の分類は以下の図のようになる。



$(A)\ \nabla f(\bar{x})=0$かつ$\nabla^2 f(\bar{x})$が半正定値

$(B)\ \nabla f(\bar{x})=0$かつ$\nabla^2 f(\bar{x})$が正定値

1. 制約なし最適化問題

これまでの流れを踏まえ、制約なし最適化問題の解き方を手続き化

１．停留点を求める

勾配ベクトルについて、$\nabla f(x, y) = 0$の解(a, b)を求める

２．ヘッセ行列$\nabla^2 f(a, b)$の正値性を調べる（狭義凸関数か調べる）

$\nabla^2 f(a, b)$が正定値または負定値のとき、(a, b)は局所最適解

$\nabla^2 f(a, b)$が不定のとき、(a, b)は局所最適解でない（鞍点）。

2. 制約付き最適化問題

制約付き最適化問題とは？

（制約付き最小化問題）

最小化:　$f(x)$

制約:　$x \in C$

ここで、集合Cを実行可能領域、 Cの点を実行可能解と呼ぶ。

例）

最大化:　$f(x, y) = xy$

制約:　$x + y = 4, x \geq 0, y \geq 0$

(x, y)	(1,3)		(2,2)		(3,1)
$f(x, y)$	3	↗	4	↘	3

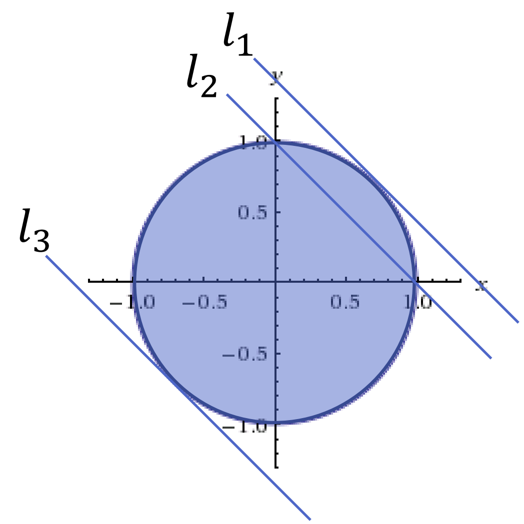
局所最大解

2-1 等式制約が一つの場合

・1次関数を円周上で最適化

最小化または最大化:　$f(x, y) = x + y$

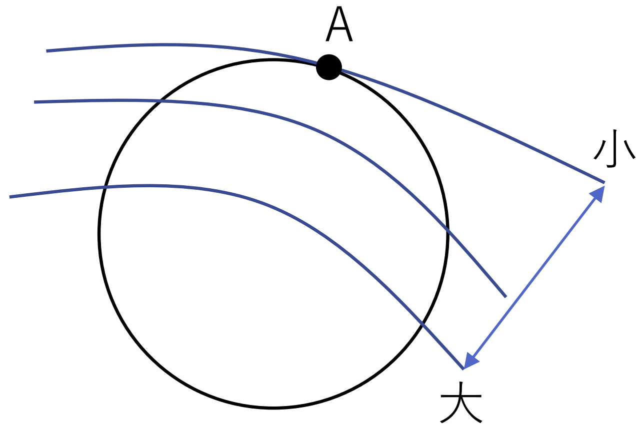
制約:　$x^2 + y^2 = 1$



このとき、l_1と円の接点が局所最大解、

l_3と円の接点が局所最小解になっている。

2-1 等式制約が一つの場合



・一般の関数を円周上で最適化
目的関数の等高線が右のようになっている
と仮定すると、点Aは局所最小解になる。

↓

定理「局所最適解で目的関数の等高線と実行可能領域は接する」

2-1 等式制約が一つの場合

（ラグランジュ乗数法）

最適化問題

最小化または最大化:　$f(x)$

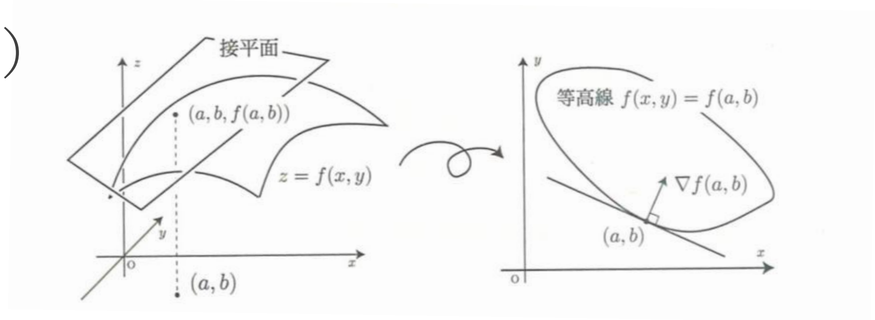
制約:　$g(x)=0$　を考え、$\bar{x}$を局所最適解とする。

$\nabla g(\bar{x}) \neq 0$ならば、ある数 λ が存在して、

$\nabla f(\bar{x}) = \lambda \nabla g(\bar{x}),\ g(\bar{x}) = 0$　が成り立つ。

2-1 等式制約が一つの場合

（ラグランジュ乗数法の幾何的理解）



図のような関数fのグラフ

$z = f(x, y)$と(a, b)における接平面

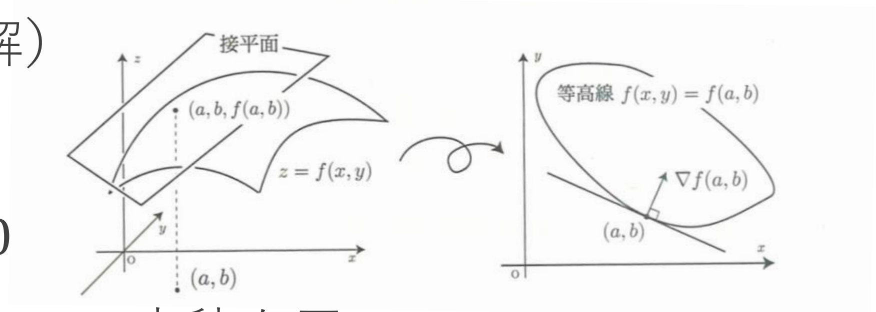
$z = f(a, b) + f_x(a, b)(x - a) + f_y(a, b)(y - b)$

を考える。点$(\mathrm{a}, \mathrm{b}, f(a, b))$を通り$xy$平面に平行な平面

$z = f(a, b)$で、fのグラフと接平面を切ったときの断面図は図右のようになり、グラフの断面は等高線、接平面の断面は等高線の$(x, y) = (a, b)$における接線になっている。

2-1 等式制約が一つの場合

（ラグランジュ乗数法の幾何的理解）



図右の接線の式は

$$f_x(a,b)(x-a)+f_y(a,b)(y-b)=0$$

という式で与えられるから、ベクトルの内積を用いて

$$\begin{bmatrix} f_x(a,b) \\ f_y(a,b) \end{bmatrix} \cdot \begin{bmatrix} x-a \\ y-b \end{bmatrix} = 0$$ と変形できる。

この式から、勾配ベクトル$\nabla f(a,b)$と等高線の$(x-y)=(a,b)$における接線が直交していることを表す。したがって、勾配ベクトルは等高線と直交している。

2-1 等式制約が一つの場合

（ラグランジュ乗数法の幾何的理解）

さらに、接平面上で点(a, b)を$\nabla f(a, b)$方向へ移動した点

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} a \\ b \end{bmatrix} + \nabla f(a, b) = \begin{bmatrix} a + f_x(a, b) \\ b + f_y(a, b) \end{bmatrix}$$

におけるz座標を調べると

$$z = f(a, b) + f_x(a, b)^2 + f_y(a, b)^2 > f(a, b)$$

となるので、$\nabla f(a, b)$ は接平面のz座標が増加する方向を向いている。

よって、勾配ベクトルは関数値が増える方向を向いている。

2-1 等式制約が一つの場合

（ラグランジュ乗数法の幾何的理解）

以上の準備をもとにまとめに入る。

$\bar{x}$を局所最適解とすると、局所最適解$\bar{x}$においてfの等高線と実行可能領域は接している。

ところで、制約は$g(x)=0$ なので、実行可能領域は値0に対する関数gの等高線である。よって、点$\bar{x}$において実行可能領域と$\nabla g(\bar{x})$は直交している。

したがって以下の関係が成り立つ。

$\nabla f(\bar{x}) \perp \{f$の等高線$\}$…①

$\nabla g(\bar{x}) \perp \{$実行可能領域$\}$…②

また、局所最適解$\bar{x}$においてfの等高線と実行可能領域は接するので、$\nabla f(\bar{x})$と$\nabla g(\bar{x})$は共通の接線に直交する。したがって$\nabla f(\bar{x})$と$\nabla g(\bar{x})$は平行であり、

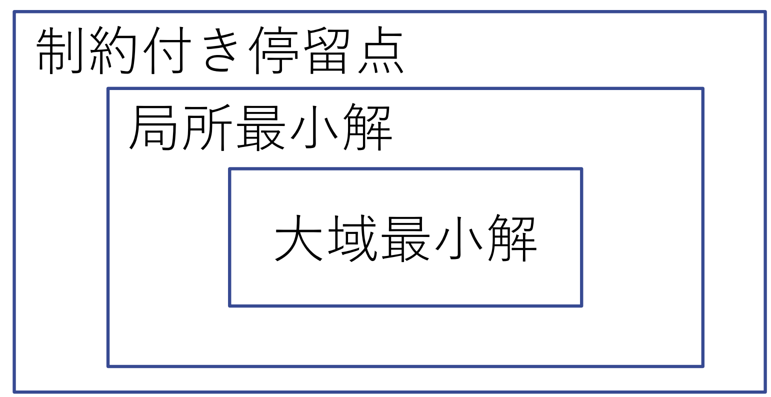
$\nabla f(\bar{x}) = \lambda \nabla g(\bar{x})$と書ける。

2-1 等式制約が一つの場合

（ラグランジュ乗数法）

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

を満たす点$\bar{x}$を制約つき停留点と呼ぶとすると、最適解との包含関係は右のようになる。



2-1 等式制約が一つの場合

ラグランジュ乗数法はあくまで**局所最適解の候補**を求める方法であるので、以下の定理を利用して局所最適解かどうか判定する。

定理：

$F(x,y,\lambda)=f(x,y)-\lambda g(x,y)$と定義する。

$(x,y,\lambda)=(\mathrm{a},\mathrm{b},\lambda_0)$が$F(x,y,\lambda)$の停留点だったとする。

このとき$(x,y,\lambda)=(\mathrm{a},\mathrm{b},\lambda_0)$における3変数関数$F(x,y,\lambda)$のヘッシアンが正ならば極大値、負ならば極大値をもつ。このヘッシアンは縁付きヘッシアン(Bordered Hessian)とよばれ、具体的には下の行列式に等しい。

$$\begin{vmatrix} 0 & g_x(a,b) & g_y(a,b) \\ g_x(a,b) & F_{xx}(a,b,\lambda_0) & F_{xy}(a,b,\lambda_0) \\ g_y(a,b) & F_{yx}(a,b,\lambda_0) & F_{yy}(a,b,\lambda_0) \end{vmatrix}$$

2-1 等式制約が一つの場合

以上のように、等式制約が一つの場合はラグランジュ乗数法と縁付きヘッシアンによる判定によって局所最適解を求めることができる。

さらに、「**目的関数が連続で、実行可能領域が有界閉集合ならば、最適化問題は大域最小解と大域最大解をもつ**」（**ワイヤシュトラウスの定理**）

という定理を用いれば、上記の条件を満たす実行可能領域に関してラグランジュ乗数法で導出した停留点の値を比較することによって大域最適解を得ることができる。

2-2 等式制約が複数の場合

制約数が任意の個数の場合は以下のように解く。

最小化問題

最小化：$f(x)$

制約：$g_1(x) = 0, g_2(x) = 0, \ldots, g_n(x) = 0$

を考え、 xを局所最小解とする。$\nabla g_1(\bar{x}), \ldots, \nabla g_n(\bar{x})$が一次独立ならば、ある数$\lambda_1, \ldots, \lambda_n$が存在して、

$\nabla f(\bar{x}) = \lambda_1 \nabla g_1(\bar{x}) + \cdots + \lambda_n \nabla g_m(\bar{x})$

$g_1(\bar{x}) = 0,\ g_2(\bar{x}) = 0, \ldots, g_n(\bar{x}) = 0$

が成り立つ。

2-2 等式制約が複数の場合

結局、ラグランジュ乗数法とは？

「**制約付き最適化問題**」を、
目的関数と制約条件を未定係数 λ によって線形結合させることで、
「**制約なし停留点探索問題**」に変換する方法のこと。

2-3-1 不等式制約が一つの場合

等式の場合と同じように未定乗数 λ を導入して解く。

最小化問題

最小化：$f(x)$

制約：$g(x) \leq 0$

に対して、$\bar{x}$を局所最小解として、$\nabla g(\bar{x}) \neq 0$ならば、ある数 λ が存在して、以下が成り立つ。

$$\begin{cases} \nabla f(\bar{x}) = \lambda \nabla g(\bar{x}) \\ \boxed{\begin{array}{c} \lambda g(\bar{x}) = 0, \quad \lambda \leq 0 \\ g(\bar{x}) \leq 0 \end{array}} \end{cases}$$

2-3-1 不等式制約が一つの場合

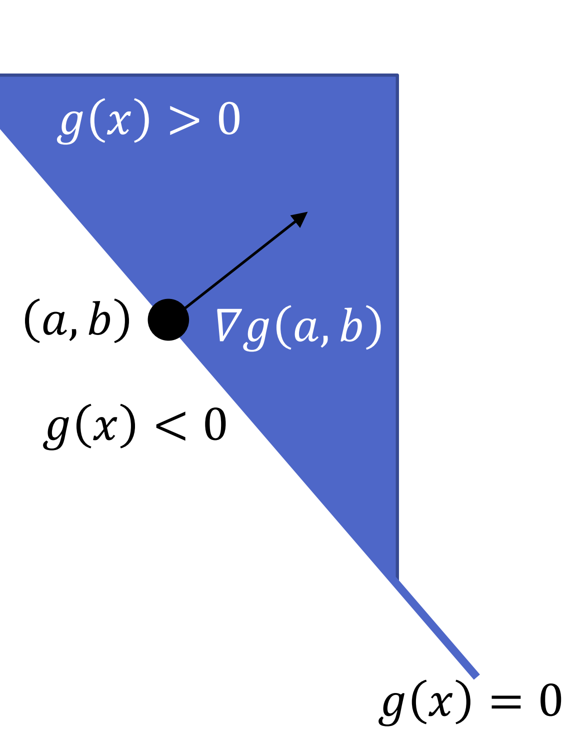
等式制約の場合と違うところはどこか？

1）　停留点が制約条件を十分に満たしている場合

（ $g(\bar{x}) < 0$のとき）制約条件は最適化問題に関係がない。

2）　局所最適解が境界線上にある場合（ $g(\bar{x}) = 0$のとき)

制約条件をλの正負によって効かさなければならない。



2-3-1 不等式制約が一つの場合

局所最適解が境界線上に位置する場合について、

最大化or最小化、∇fと∇gの向きが同じor反対の4つの場合について考えると

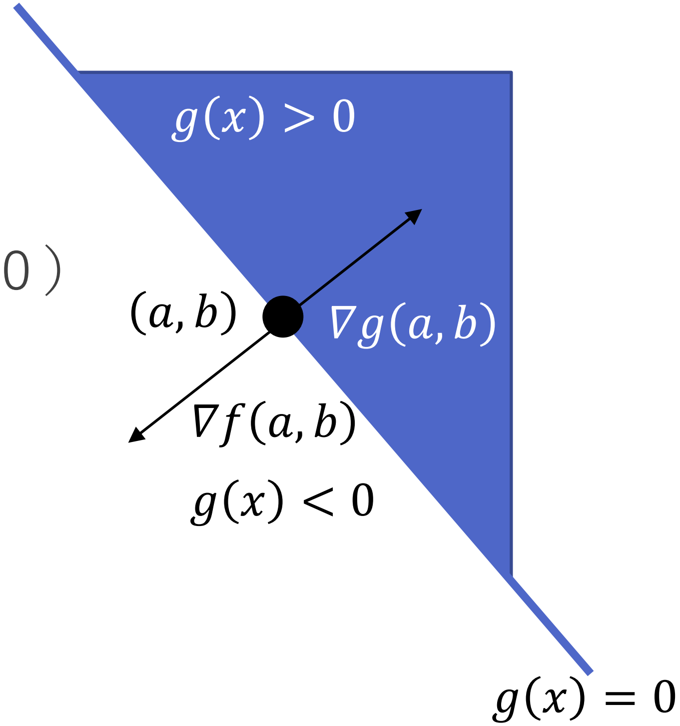
１）最大化かつベクトルの向きが同じ

この場合、制約条件を満たす上で問題は

生じない。よって制約条件は無効。（$\lambda = 0$）

２）最大化かつベクトルの向きが反対

$\nabla f + \lambda \nabla g = 0$において、必ず$\lambda > 0$である。



2-3-1 不等式制約が一つの場合

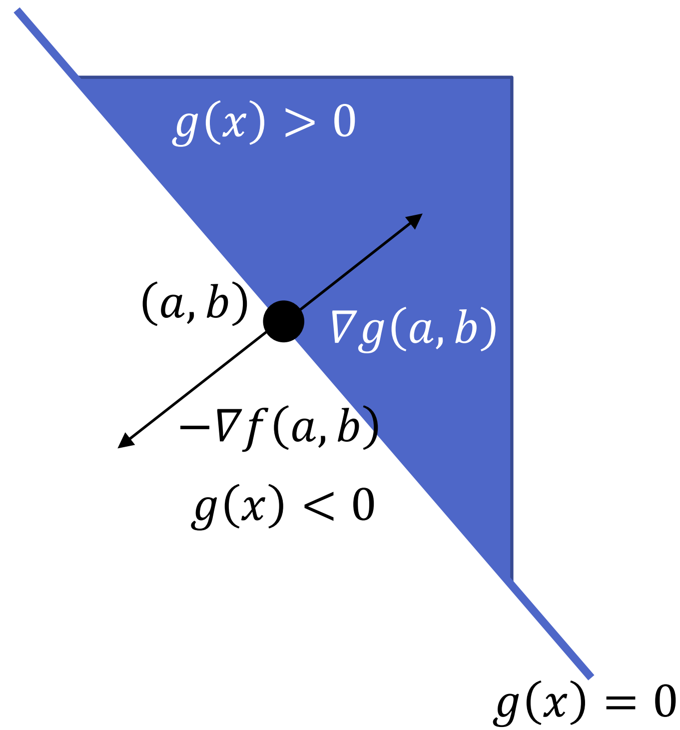
最小化の場合も、最大化の場合と同じようにして

３）最小化かつベクトルの向きが同じ

$\nabla f+\lambda\nabla g=0$において、必ず$\lambda<0$である。

４）最小化かつベクトルの向きが反対

最大化の時と同様に、制約条件は無効。



2-3-1 不等式制約が一つの場合

以上をまとめて書くと、不等式制約のもとでの最適化問題について

最小化：$f(x)$

制約：$g(x) \leq 0$　ならば

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

が成り立つが、λ の正負は最適化の方向と制約条件の不等号の向きによって変わるので留意しておく必要がある。

2-3-2 不等式制約が複数の場合

複数の制約がある場合も、一つの場合の解き方を拡張できる。

最適化問題

最小化：$f(x)$

制約：$g_1(x) \leq 0, g_2(x) \leq 0$

に対して、$\bar{x}$を局所最小解であり、$\nabla g_1(\bar{x})$と$\nabla g_2(\bar{x})$が一次独立ならば、ある数λ_1, λ_2が存在して、以下が成り立つ。

$$
\begin{cases}
\nabla f(\bar{x}) = \lambda_1 \nabla g_1(\bar{x}) + \lambda_2 \nabla g_2(\bar{x}) \\
\qquad \lambda_i g(\bar{x}) = 0, \quad \lambda_i \geq 0 \\
\qquad g_i(\bar{x}) \leq 0 \ \ (i = 1,2)
\end{cases}
$$

2-3-3 制約に不等式と等式がある場合

最適化問題

最小化：$f(x)$

制約：$g_1(x) \leq 0, g_2(x) \leq 0, h(x) = 0$

に対して、$\bar{x}$を局所最小解であり、$\nabla g_1(\bar{x}), \nabla g_2(\bar{x}), h(x)$が一次独立であるとする。すると、ある数$\lambda_1, \lambda_2, \mu$ が存在して、

$$\begin{cases} \nabla f(\bar{x}) = \lambda_1 \nabla g_1(\bar{x}) + \lambda_2 \nabla g_2(\bar{x}) + \mu \nabla h(\bar{x}) \\ \qquad \lambda_i g_i(\bar{x}) = 0, \quad \lambda_i \geq 0, \ \ g_i(\bar{x}) \leq 0 \\ \qquad\qquad \mu: \text{任意}, \ \ h(\bar{x}) = 0 \end{cases}$$

が成り立つ。

2-3-3 制約に不等式と等式がある場合

$$
\begin{cases}
\nabla f(\bar{x}) = \lambda_1 \nabla g_1(\bar{x}) + \lambda_2 \nabla g_2(\bar{x}) + \mu \nabla h(\bar{x}) \\
\lambda_i g_i(\bar{x}) = 0, \quad \lambda_i \leq 0,\ g_i(\bar{x}) \geq 0 \\
\mu\text{: 任意},\ h(\bar{x}) = 0
\end{cases}
$$

この条件式をKarush-Kuhn-Tucker条件、略してKKT条件と呼ぶ。

制約を増やすと以下のようになる。（最小化問題）

$$
\begin{cases}
\nabla f(\bar{x}) = \lambda_1 \nabla g_1(\bar{x}) + \cdots + \lambda_m \nabla g_m(\bar{x}) + \mu_1 \nabla h_1(\bar{x}) + \cdots + \mu_l \nabla h_l(\bar{x}) \\
\lambda_i g_i(\bar{x}) = 0, \quad \lambda_i \leq 0,\ g_i(\bar{x}) \geq 0 \\
\mu\text{: 任意},\ h_j(\bar{x}) = 0
\end{cases}
$$

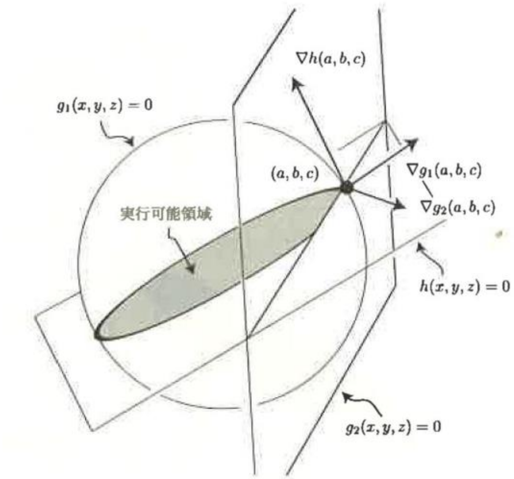
2-3-3 制約に不等式と等式がある場合

参考書掲載の具体例を用いてKKT条件の幾何的意味を考える。

例）$\alpha, \beta, \gamma, \delta$を定数として、以下の1次関数を最小化する

最小化　$f(x, y, z) = \alpha x + \beta y + \gamma z + \delta$

制約　$\begin{cases} g_1(x, y, z) = x^2 + y^2 + z^2 - 1 \leq 0 \\ g_2(x, y, z) = -x - 1/2 \leq 0 \\ h(x, y, z) = x + y + z - 1 = 0 \end{cases}$



実行可能領域は右のようになる。

2-3-3 制約に不等式と等式がある場合

1)点(a, b, c)が空間内の円の内側にある場合

g_1とg_2の制約が外れるので、

この最小化問題は

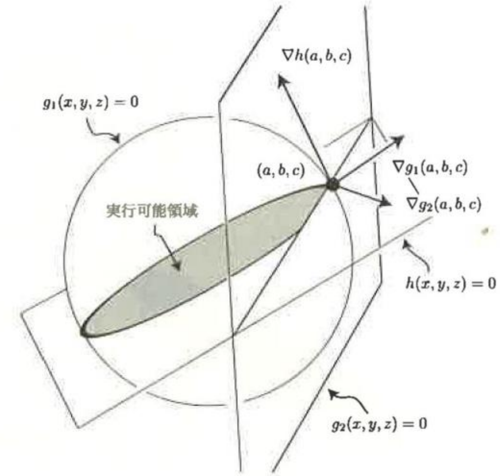
最小化：$f(x, y, z)$

制約：$h(x, y, z) = 0$

に帰着する。したがって、あるμが存在して

$\nabla f(a, b, c) = \mu \nabla h(a, b, c)$が成り立つ。

よってKKT条件式で$\lambda_1 = \lambda_2 = 0$と置けばよい。



2-3-3 制約に不等式と等式がある場合

2)点(a,b,c)が空間内の円の境界にある場合

g_2の制約がないので、

最小化：$f(x,y,z)$

制約：$\begin{cases} g_1(x,y,z)=0 \\ h(x,y,z)=0 \end{cases}$

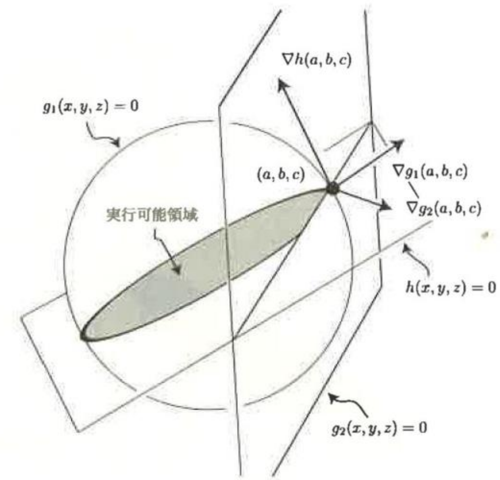
の局所最小解を求めればよい。

したがってラグランジュ乗数λ_1を導入して

$\nabla f(a,b,c)=\lambda_1 \nabla g_1(x,y,z)+\mu \nabla h(a,b,c)$を解く。

ここで、以前の議論から$\lambda_1 \leq 0$である。

これより、KKT条件式で$\lambda_2=0$と置けばよい。



2-3-3 制約に不等式と等式がある場合

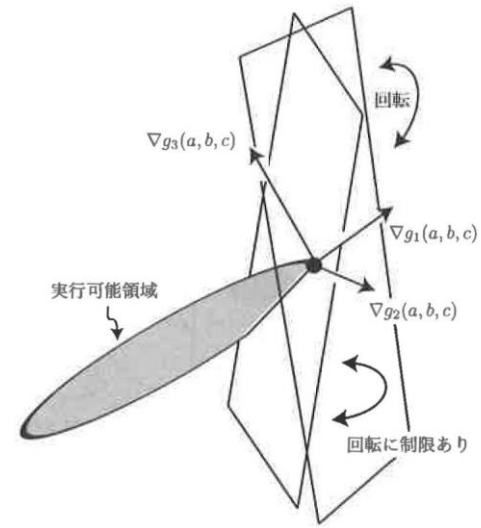
3)点(a,b,c)が実行可能領域の角張った部分にある場合

今、点(a,b,c)で実行可能領域に接する平面は

右図のように回転させても実行可能領域に接し

続けるので、実行可能領域に接する平面の法線

ベクトルは、回転させた平面のものも含めて、



$\lambda_1 \nabla g_1(a,b,c) + \lambda_2 \nabla g_2(a,b,c) + \boxed{\mu \nabla h(a,b,c)}$,

$\lambda_1 \leq 0, \lambda_2 \leq 0, \mu$: 任意　となる。

いま、目的関数の(a,b,c)を通る等高面の法線ベクトルは$\nabla f(a,b,c)$なので、以下のように書ける。

$\nabla f(a,b,c) = \lambda_1 \nabla g_1(a,b,c) + \lambda_2 \nabla g_2(a,b,c) + \mu \nabla h(a,b,c)$,

$\lambda_1 \leq 0, \lambda_2 \leq 0, \mu$: 任意と書ける。

2-3-3 制約に不等式と等式がある場合

改めてKKT条件式をみると、

$$\begin{cases} \nabla f(\bar{x}) = \lambda_1 \nabla g_1(\bar{x}) + \cdots + \lambda_m \nabla g_m(\bar{x}) + \mu_1 \nabla h_1(\bar{x}) + \cdots + \mu_l \nabla h_l(\bar{x}) \\ \lambda_i g_i(\bar{x}) = 0,\quad \lambda_i \leq 0,\ g_i(\bar{x}) \geq 0 \\ \mu\text{: 任意},\ h_j(\bar{x}) = 0 \end{cases}$$

ここで、等式制約の場合には現れなかった、λ_iに関する条件
$\lambda_i g_i(\bar{x}) = 0,\quad \lambda_i \geq 0,\ g_i(\bar{x}) \leq 0$は相補性条件とよばれ、有効でない制約条件についてはλ = 0となる。（$g_i(\bar{x}) = 0$でないということは$\bar{x}$はg_iの境界線上にないので、g_iについて考慮する必要がない）

KKT条件式は凸計画問題と線形計画問題においては最適性の必要十分条件である。

2-4 凸計画問題

最後に、凸計画問題について軽く触れておく。

目的関数も制約式もすべて凸関数である問題を凸計画と呼ぶ。

このとき、KKT条件式を満たす$\bar{x}$は大域最適解になっている。

目的関数や制約式が全て線形である線形計画問題も凸計画問題であり、KKT条件式のみによって大域最適解を見つけることができる。

2-5 制約つき最適化問題まとめ

（等式制約の場合）

最小化：$f(x,y,z)$

制約：$g_1(x,y,z)=0$

$g_2(x,y,z)=0$

→

（等式制約の場合）

$\nabla f(\bar{x})=\lambda_1\nabla g_1(\bar{x})+\lambda_2\nabla g_2(\bar{x})$

を解く。

（不等式制約の場合）

最小化：$f(x,y)$

制約：$g_1(x,y)\leq 0$

$g_2(x,y)\leq 0$

→

（不等式制約の場合）

$$\begin{cases}\nabla f(\bar{x})=\lambda_1\nabla g_1(\bar{x})+\lambda_2\nabla g_2(\bar{x})\\ \lambda_1\leq 0,\ \lambda_2\leq 0\end{cases}$$

を解く。

2-5 制約つき最適化問題まとめ

（等式制約と不等式制約が混在している場合）

KKT条件式

$$\begin{cases} \nabla f(\bar{x}) = \lambda_1 \nabla g_1(\bar{x}) + \cdots + \lambda_m \nabla g_m(\bar{x}) + \mu_1 \nabla h_1(\bar{x}) + \cdots + \mu_l \nabla h_l(\bar{x}) \\ \lambda_i g_i(\bar{x}) = 0, \quad \lambda_i \leq 0,\ g_i(\bar{x}) \geq 0 \\ \mu: \text{任意},\ h_j(\bar{x}) = 0 \end{cases}$$

を解く。

凸計画問題においてはKKT条件式は必要十分条件である。

2-5 制約つき最適化問題まとめ

（補足：KKT条件式のイメージ）

制約なし問題では、∇f(x)=0が最適であるための必要条件のひとつだが、制約ありの場合は∇f(x)=0ではなくても実行可能領域の境界上などで最適解がみつかる場合が多いので、∇f(x)という情報だけでは最適解は見つからない。

そこで、最適解である点がいくつもの制約を満たすために制約を満たすような方向へ”引っ張られて釣り合った結果”の点であると考えたら、

$$\nabla f(\bar{x}) - (\lambda_1 \nabla g_1(\bar{x}) + \cdots + \lambda_m \nabla g_m(\bar{x}) + \mu_1 \nabla h_1(\bar{x}) + \cdots + \mu_l \nabla h_l(\bar{x}))$$

=0 を満たす実数 λ と μ が存在するといえる。

おわりに

・制約付きの場合も、ラグランジュ乗数法やKKT条件式をつかって制約なし最適化問題に帰着させて解く。

・結局浅い数学的理解とパターン化に終わってしまった。

・凸計画問題ではないときにKKT条件式の十分性を示す方法についてはよくわからなかった。

・これらのスライドを作るのに丸3日もかかってしまったので、次回からはもっと効率よくやりたい。

参考文献

関口良之「はじめての最適化」近代科学社、2014年

・きちんと基礎から学ぶことができる。

・文系でもほぼ理解できた。

その他、多数のウェブサイトにお世話になりました。