

1月18日: 今日の前半では「陰関数定理」の残りを簡単にまとめます。後半では「条件付き極値問題」に入ります。

重要な連絡: 演習の方で一問も解いていない人は、後1回ありますから、ともかく問題を解いて下さい!

この科目の期末試験は2月8日(水)、演習の試験は2月10日(金)のようですが、各自、確かめて下さい。

先週は2変数関数 $f(x, y) = 0$ を y について「解いて」 $y = \varphi(x)$ の形に書く(書ける条件を求める)問題を考えた。ここでは簡単に、これらの一般化を考えておこう。自然な一般化の方向は二つある:

- 3変数以上の関数 $f(x_1, x_2, \dots, x_n, y) = 0$ を y について「解く」。つまり、条件は一つだけで、考えている関数のいる空間の次元が高くなっている。
- 多変数の問題で、条件が2つ以上ある。例えば、 $f(x_1, x_2, \dots, x_n, y, z) = 0$ と $g(x_1, x_2, \dots, x_n, y, z) = 0$ を y, z について「解け」、など。

これらの話は教科書に載っている(大半は付録に追いやられているが)。厳密な事は教科書にまわして、このプリントでは大体の感じをつかみやすくなるような手助けを目指す事にする。

まず、3変数に条件が一つの場合を考えよう。つまり、 $f(x, y, z) = 0$ を z について(または x または y について)解くわけだ。この問題の例として、地球の内部の点の座標を (x, y, z) とし、 $f(x, y, z)$ はその点の温度から $1000^\circ C$ を引いたものを考える。とすると、 $f(x, y, z) = 0$ は温度が丁度 $1000^\circ C$ の点の集合を表す。今は3次元空間内に1つの条件がついているから、これは次元が一つ下がって2次元の曲面(等温面)になることが予想される。

それで陰関数定理の問題とするところは、「どのような条件のもとで、「等温面」が普通の意味での曲面になるか」ということだ。 $\mathbf{x} = (x_1, x_2, \dots, x_n)$ とまとめて書くと、以下のようになる。

定理 5.3.4 ($n+1$ 変数 1 条件の陰関数定理, 教科書の定理 2.4.4) $(n+1)$ -変数の C^1 -級の関数 $f(x, y)$ がある。 $f(a, b) = 0$ かつ $f_y(a, b) \neq 0$ ならば、 $f(x, y) = 0$ は (a, b) の近傍で y について解ける。すなわち、

$$b = \varphi(a) \quad \text{かつ} \quad (a, b) \text{ の近傍で } f(x, \varphi(x)) = 0 \quad (5.3.3)$$

かなりたつような C^1 -級の関数 $\varphi(x)$ が一意に存在する。更に (a, b) の近傍では

$$\frac{\partial}{\partial x_i} \varphi(x) = - \frac{f_{x_i}(x, y)}{f_y(x, y)} \Big|_{y=\varphi(x)} \quad (5.3.4)$$

もなりたつ。なお、 f が C^r -級 ($r \geq 1$) なら、 $\varphi(x)$ も C^r -級である。

条件が2つ以上の時はもっと大変だし、教科書でも大半は「付録」に追いやられている。結果は大事だから、一応、以下に掲げておく。

まず、言葉の定義

定義 5.3.5 (ヤコビ行列) n -変数の C^1 -級の関数が m 個ある状況を考える： $f_1(x), f_2(x), \dots, f_m(x)$ 。このとき、行列

$$\begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \dots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \dots & \frac{\partial f_2}{\partial x_n} \\ \vdots & \vdots & \dots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \frac{\partial f_m}{\partial x_2} & \dots & \frac{\partial f_m}{\partial x_n} \end{bmatrix} \quad (5.3.5)$$

つまり、その ij 成分が $\frac{\partial f_i}{\partial x_j}$ であるような行列を、 f のヤコビ行列といい、 $\frac{Df}{Dx}$ と書く。

一般の陰関数定理は以下のようになる ($x = (x_1, x_2, \dots, x_n), y = (y_1, y_2, \dots, y_m)$ とまとめて書く):

定理 5.3.6 ($n + m$ 変数 m 条件の陰関数定理, 教科書の定理 A.3.3) ($n + 1$ -変数の C^1 -級の関数 $f_i(x, y)$ がある ($i = 1, 2, \dots, m$). f_i を x の関数とみなして (y は固定して) 定義したヤコビ行列を $\frac{Df_i}{Dx}$, f_i を y の関数とみなして (x は固定して) 定義したヤコビ行列を $\frac{Df_i}{Dy}$ と書こう. すると, $f(a, b) = 0$ かつ $\frac{Df_i}{Dy}(a, b)$ が正則ならば, $f(x, y) = 0$ は (a, b) の近傍で y について解ける. すなわち,

$$b = \varphi(a) \quad \text{かつ} \quad (a, b) \text{ の近傍で } f_i(x, \varphi(x)) = 0 \quad (i = 1, 2, \dots, m) \quad (5.3.6)$$

がなりたつような m この C^1 -級の関数 $\varphi = (\varphi_1, \varphi_2, \dots, \varphi_m)$ が一意に存在する. 更に (a, b) の近傍では

$$\frac{D\varphi}{Dx} = - \left(\frac{Df}{Dy} \right)^{-1} \frac{Df}{Dx} \Big|_{y=\varphi(x)} \quad (5.3.7)$$

もなりたつ. なお, f が C^r -級 ($r \geq 1$) なら, $\varphi(x)$ も C^r -級である.

5.4 条件付き極値問題: ラグランジュの未定乗数法

今学期 (そしてこの科目) 最後の主な話題です.

わかりやすいように 2 変数の場合をまず考え, 一般の場合は後で簡単に触れるにとどめる.

(問 1) 関数 $f(x, y)$ を, 条件 $g(x, y) = 0$ の下で最大・最小 (極大・極小) にするような (x, y) と, その時の $f(x, y)$ の値を求めよ.

ここで「条件 $g(x, y) = 0$ の下に (a, b) で極小」の意味は以下の 2 つが成り立つ事である.

- $g(a, b) = 0$ である.
- $g(x, y) = 0$ かつ $(x, y) \neq (a, b)$ であるような, (a, b) に十分近い (x, y) に対しては $f(x, y) > f(a, b)$ である.

このような問題を「条件付き極値 (最大最小) 問題」という.

(注) 今までにも意識的に避けてきたのだが, 最大・最小の問題は極大・極小の問題よりも難しい—— 極大・極小点をすべて求めた上で, 考えている領域の境界での値とも比べる必要があるから. そこで, 最大・最小問題については来週, 簡単にコメントする事にして, ここでは極大・極小問題に注力する.

このような問題がいままで極大・極小問題と異なるのは, $g(x, y) = 0$ などの条件 (拘束条件, constraint) がついていることだ. この条件のため, x, y は独立に動く事ができない. 従って, 「2 変数関数の極値問題」のように単純に偏微分してやる訳にはいかない.

少し気をつければ, 今までの知識だけでも「愚直に」解く事は大体, 可能だ. つまり $g(x, y) = 0$ を y について解いて y を x の関数として表し, それを $f(x, y)$ に代入して $f(x, y)$ を x だけの関数として表す. こうすれば x は自由に動けるから, 問題は (高校でやった) 1 変数関数の極値問題になる. 従って, 普通に x で微分してやればよい.

(例 1) $f(x, y) = x^4 + y^4$ の極値を, 条件 $x^2 + y^2 = 1$ の下で求めよ.

これなら $y = \pm\sqrt{1-x^2}$ と解いて $f = x^4 + (1-x^2)^2 = 2x^4 - 2x^2 + 1 = 2(x^2 - \frac{1}{2})^2 + \frac{1}{2}$ となるから, $x = \pm 1/\sqrt{2}$ で極小 (この場合は最小) になる. 極小値は $\frac{1}{2}$. 極値をとる (x, y) は $(\pm \frac{1}{\sqrt{2}}, \pm \frac{1}{\sqrt{2}})$ (複合任意).

ところが, このようなやり方は往々にして非常に面倒になる. 上の例では $g(x, y)$ が簡単だから助かったけど, 例えば, $g(x, y) = x^6 + 3xy - y^2$ だったらどうだろう? $g(x, y)$ が多項式でなく, \sin, \cos, \log などで書かれていたら?

と言うわけで, 応用上, もっと簡便な方法がないとやってられない. これを与えてくれるのが「Lagrange の未定乗数法」である. そのやり方をまず説明しよう (理由はあとで).

(Lagrange の未定乗数法; 教科書の定理 2.6.2) 上の (問 1) の条件付き極値問題を考える. まず, 天降りではあるが, 新しい変数 λ を導入して

$$F(x, y, \lambda) = f(x, y) - \lambda g(x, y) \quad (5.4.1)$$

を定義する. すると, この条件付き極値問題において, 極値を取る点の候補 (x, y) は, 以下の (i), (ii) のいずれかである.

(i) $g(x, y) = 0$ の特異点,

(ii) 未知変数を x, y, λ とする以下の連立方程式の解.

$$0 = \frac{\partial F}{\partial x} = \frac{\partial f}{\partial x}(x, y) - \lambda \frac{\partial g}{\partial x}(x, y), \quad 0 = \frac{\partial F}{\partial y} = \frac{\partial f}{\partial y}(x, y) - \lambda \frac{\partial g}{\partial y}(x, y) \quad (5.4.2)$$

$$0 = \frac{\partial F}{\partial \lambda} = g(x, y) \quad (5.4.3)$$

つまり ($g(x, y) = 0$ の特異点を除けば) 形式的には, この条件付き極値問題は新しく定義した関数 $F(x, y, \lambda)$ の普通の極値問題 — x, y と λ が自由に動く — のように見える.

λ を考案者の名前をとって Lagrange の未定乗数 (Lagrange multiplier) という. なお, この方法では極値をとる (x, y) の候補が見つかるだけであって, それらが実際の極値を与えるか否かを定める一般論は存在しない (より正確には, そのような一般論がない訳ではないが, 実用的なものはほとんどない.) ただし, 極値点の候補が見つければ, その点の周りでのテイラー展開などを用いて, 実際に極値になっているかどうかの判定は可能な事が多いから, これは実用上は大した問題ではない (少なくとも計算機の助けを借りれば何とかなる). また, 方程式 (5.4.2) と (5.4.3) (やその多変数の場合の該当物) を解くのは大変だと強調している本が多いが, これも計算機の助けを借りればそんなに大した問題ではない (事も多い). というわけで, 未定乗数法はやはり偉大なのである.

具体例: 上の (例 1) なら, $g(x, y) = 0$ の特異点はないので, 解くべきは $F(x, y, \lambda) = x^4 + y^4 + \lambda(x^2 + y^2 - 1)$ を考えて

$$0 = 4x^3 + 2\lambda x, \quad 0 = 4y^3 + 2\lambda y, \quad 0 = x^2 + y^2 - 1 \quad (5.4.4)$$

の 3 つである. これを解くと,

$$\begin{bmatrix} x \\ y \\ \lambda \end{bmatrix} = \begin{bmatrix} 0 \\ \pm 1 \\ -2 \end{bmatrix}, \quad \begin{bmatrix} \pm 1 \\ 0 \\ -2 \end{bmatrix}, \quad \begin{bmatrix} \pm \frac{1}{\sqrt{2}} \\ \pm \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix}, \quad \begin{bmatrix} \pm \frac{1}{\sqrt{2}} \\ \mp \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix}, \quad (\text{ベクトルの中では複合同順}) \quad (5.4.5)$$

となる. 後ろの 2 つは変数を消去して解いたものと同じでメダタシメダシ.

(未定乗数法がうまく行く理由) これがうまく行く「証明」は教科書の pp.66-67 に載っている. もう少し粗っぽい説明を書いておこう.

条件 $g(x, y) = 0$ が嫌らしいわけだから, 「愚直」な方法で解くつもりになって, y を x で表してやろう. これを $y = \varphi(x)$ と書く (実際にこのように表せるかどうかは自明ではないが, 「陰関数定理」によって, $g(x, y) = 0$ の特異点以外では可能である — 場合によっては $x = \psi(y)$ の形にしか解けない事もあるが). これを元の f に代入して $h(x) = f(x, \varphi(x))$ を作る.

この $h(x)$ は x のみの関数だから極値の条件は

$$0 = h'(x) = f_x + f_y \varphi'(x) \quad (5.4.6)$$

となっている (偏微分は $(x, \varphi(x))$ での値). ところが, $g(x, \varphi(x)) = 0$ であるから, この両辺を x で微分すると

$$0 = \frac{d}{dx} g(x, \varphi(x)) = g_x + g_y \varphi'(x) \quad (5.4.7)$$

この 2 つから,

$$\varphi'(x) = -\frac{f_x}{f_y} = -\frac{g_x}{g_y} \quad (5.4.8)$$

が導かれるが、これは見方を変えれば

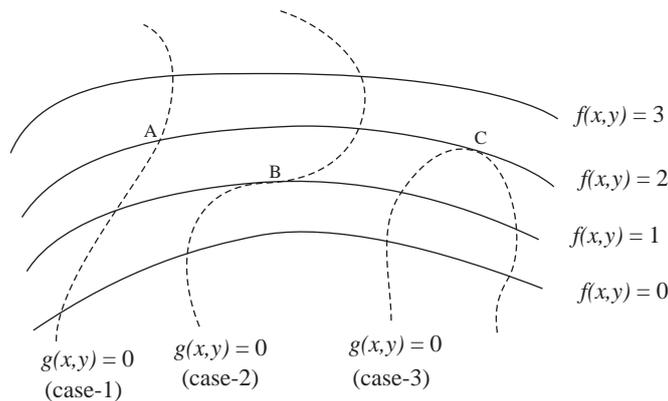
$$\frac{g_x}{f_x} = \frac{g_y}{f_y} \tag{5.4.9}$$

ということであり、この値を λ と書けば、これは (5.4.2) に他ならない (以上では g_y や f_y などがゼロでないと仮定して分数の形に書いたが、これらがゼロの場合は個別に扱えば大丈夫である事はわかる — 教科書の正しい証明を参照のこと) . □

(未定乗数法の直感的意味) 上の「証明」は愚直な方法で計算してみたらこうなった、というもので、どうも直感的ではない。ここではその直感的な説明を試みる (以下は「解析概論」などを参考にした) .

陰関数定理を扱ったとき、 $g(x, y) = 0$ は $g(x, y) = 0$ の「等高線」を表していることを指摘した。同様に c を定数として、 $f(x, y) = c$ は $f = c$ の等高線を表している。我々の問題は、 $g(x, y) = 0$ の等高線上で $f(x, y)$ の値を極大(極小)にすること、言い換えれば $g(x, y) = 0$ の等高線と $f(x, y) = c$ の等高線の交わりが存在するような c の値を探す事である。

以下に $f(x, y) = c$ の等高線と $g(x, y) = 0$ の等高線の様子を模式的に描いてみた。 $f(x, y) = 0, 1, 2, 3$ の4本の等高線が図の実線、 $g(x, y) = 0$ の等高線が図の点線である (ただし、3つの典型的な場合を同じ図の中に描きこんだ) .



通常、 $f(x, y) = c$ の等高線と $g(x, y) = 0$ の等高線は (接しないで) 交わり図の case-1 のようになっている。この場合、 $g(x, y) = 0$ の等高線 (点線) に沿って進むと、 $f(x, y)$ の値は $0, 1, 2, 3$ と増えてくるので、極値はない。

しかし、case-3 の場合には $g(x, y) = 0$ に沿って進むと、始めは $f(x, y) = 0, 1$ と増えて行くが、 $f(x, y) = 2$ になったのを最高にして、 f の値が減少してしまう。つまり、この場合には $f = 2$ が極大になっているわけだ。この場合、図でも明らかのように、 $f(x, y) = 2$ と $g(x, y) = 0$ の曲線が点 C で接している。

一方、case-2 の場合にも2つの曲線が点 B で接しているが、点 B では極値にはなっていない。つまり、接する事は必要条件ではあるが、十分条件ではない。

以上から、点 (a, b) で極値になるための必要条件は、 $f(x, y) = c$ と $g(x, y) = 0$ の曲線が (a, b) で接する事だとわかった (もちろん、接線がひけないような曲線の場合には話は別だが) . そこで、2つの曲線が接する条件を具体的に書き下してみよう。そのためには、 $f(x, y) = c$ の接線の傾きを知る必要があるが、その答えは既に陰関数定理 5.3.3 の (5.3.2) で与えられている。つまり

$$f(x, y) = c \text{ の接線の傾きは } -\frac{f_x}{f_y}, \quad g(x, y) = 0 \text{ の接線の傾きは } -\frac{g_x}{g_y} \tag{5.4.10}$$

なのだ。従って、両者が接する条件は

$$-\frac{f_x}{f_y} = -\frac{g_x}{g_y} \quad \text{つまり} \quad \frac{g_x}{f_x} = \frac{g_y}{f_y} \tag{5.4.11}$$

であるが、これは (5.4.2) に他ならない。 □

1月25日: いよいよ最終回になりました。今日は条件付き極値問題をまとめた後で、言い残した事(最大最小, 包絡線?)を少しやりましょう。

重要な連絡: 演習の方で一問も解いていない人は、後1回ありますから、ともかく問題を解いて下さい!

この科目の期末試験は2月8日(水)、演習の試験は2月10日(金)のようですが、各自、確かめて下さい。期末試験は「今学期にやった事全部」ですが、主な題材は「中間テストの範囲」に加えて、「条件付き極値問題」と「陰関数定理」などです。ただ正直、陰関数定理で何を出題するかは難しく、教科書の練習問題のような、偏微分の計算問題になってしまうかも。まあ考えてみます...

(条件付き極値問題の続き)

より一般の条件付き極値問題は以下ようになる(今までのように $x = (x_1, x_2, \dots, x_n)$ と書く):

(問2) n -変数の関数 $f(x), g_1(x), g_2(x), \dots, g_m(x)$ がある。 $m < n$ として、 m の条件 $g_i(x) = 0$ ($i = 1, 2, \dots, m$) の下で $f(x)$ を最大・最小(極大・極小)にする $x = (x_1, x_2, \dots, x_n)$ と、その時の $f(x)$ の値を求めよ。

(Lagrange の未定乗数法; 教科書の定理 A.3.7) 上の(問2)の条件付き極値問題を考える。ただし、 f, g_i は C^1 -級の関数とする。このとき、新しい変数 $\lambda_1, \lambda_2, \dots, \lambda_m$ を導入して

$$F(x, \lambda_1, \lambda_2, \dots, \lambda_m) = f(x) - \{ \lambda_1 g_1(x) + \lambda_2 g_2(x) + \dots + \lambda_m g_m(x) \} \quad (5.4.12)$$

を定義する。すると、この条件付き極値問題において、極値を取る点の候補 x は、以下の (i), (ii) のどちらかを満たす。

(i) x で のヤコビ行列 $\frac{Dg}{Dx}$ の階数が m より小さい。

(ii) x は未知変数を x および $\lambda_1, \lambda_2, \dots, \lambda_m$ とする以下の連立方程式を満たす。

$$\begin{aligned} 0 &= \frac{\partial F}{\partial x_j} = \frac{\partial f}{\partial x_j}(x) - \sum_{k=1}^m \lambda_k \frac{\partial g_k}{\partial x_j}(x), & (j = 1, 2, \dots, n) \\ 0 &= \frac{\partial F}{\partial \lambda_k} = g_k(x), & (k = 1, 2, \dots, m) \end{aligned} \quad (5.4.13)$$

大雑把に言えば、 m 個の条件があった場合には、 m 個の未定乗数を導入して、条件が1個のときと同じように解けば良いのである。ただし、条件が1個の時と同様に、このようにして求めたものはあくまで「極値を取る点の候補」である。これらの候補で実際に極値になっているかどうかの簡単な判定条件はない。

5.5 最大・最小問題についての注意(おまけ)

今まで、関数の最大・最小を扱う問題は意識的に避け、代わりに極大・極小問題を扱ってきた。1変数の場合には両者にはそれほど差がないが、2変数以上では少し差がでてくる。この節ではこの講義のおまけとして、2変数以上でどのような困難が生じるのか、および、それでも成り立つ一般的な定理について述べる。

まず、1変数の場合を思い出そう。ある閉区間 $[a, b]$ で定義された関数 $f(x)$ の最大値 M とは、

ある $x_0 \in [a, b]$ にて $f(x_0) = M$, かつ

すべての $x \in [a, b]$ にて $f(x) \leq M$

となる数の事であった。関数 $f(x)$ が C^1 -級の場合、この問題は要するに

- もし x_0 が開区間 (a, b) の点なら、 x_0 は $f(x)$ の極大点であるので、今までの極大・極小を調べるやり方で扱える。
- x_0 が (a, b) の点ではない、つまり $x_0 = a$ または $x_0 = b$ ならば、 $f(a), f(b)$ の値を具体的に調べて、上の極大点での値と比べれば良い。

として解く事ができたのは、受験数学で皆さんよくご存知の通りだ。だから、1変数(で考える区間が単純な閉区間)の場合は最大最小の問題は極大極小問題とほとんど変わらない。考える区間が開区間や半开区間の場合も、要するに端点での $f(x)$ の値(またはその極限值)と区間内の点の比べあいだから、大した問題ではなかった。

ところが、2変数以上になると事情はそう単純ではない。簡単のため、領域 $x^2 + y^2 \leq 1$ 上で定義された C^1 -級の関数 $f(x, y)$ の最大最小を考えてみよう。考え方は1変数の時と同じで(1)この領域の内点での極大極小を調べ(2)それと境界での $f(x, y)$ の値を比べて、真の最大や最小を探す、となるわけだ。これは原理的には1変数の場合と同じだが、境界が点ではなく曲線になっているため、それなりに難しい事が多く、個々の問題で工夫が必要となる。

さて、もうすこし一般的に言えることを考えておこう。

1変数の場合は(グラフを書けば直感的には明らかだが)閉区間で定義された連続関数はかならず最大値と最小値を持った(前期の定理 2.6.1, 教科書の定理 1.5.10)。同じ事は多変数の場合も成り立つのだろうか? この問いに答えるには、「閉区間」に相当するものを多変数の場合に定義する必要がある。今までと同じように $a = (a_1, a_2, \dots, a_n)$ と書く。また、 $\|a\| = (\sum_{j=1}^n a_j^2)^{1/2}$ でベクトル a の長さを表す。

定義 5.5.1 (開集合, 閉集合) \mathbb{R}^n の部分集合 A が 開集合 とは,

$$\forall a \in A \exists \delta(a) > 0 \quad D(a, \delta) := \{x \in \mathbb{R}^n \mid \|x - a\| < \delta\} \subset A \quad (5.5.1)$$

がなりたつこと。また、 $B \subset \mathbb{R}^n$ が 閉集合 とは、 $B^c = \mathbb{R}^n \setminus B$ が開集合であること。

(注) 世の中の集合には、開集合でも閉集合でもないものは一杯ある! 時々、4年生くらいになって「この集合は開集合ではないので閉集合です」と言ってる人がいるから、要注意!

実は、教科書での閉集合の定義は上のものと違うのだ。でも、以下の定理があるから、大丈夫!

定理 5.5.2 (閉集合の性質) \mathbb{R}^n の部分集合 B が閉集合である事は、以下の命題と同値である。

B 内の点列 a_1, a_2, a_3, \dots が、極限 $b = \lim_{n \rightarrow \infty} a_n$ を持つならば、 $b \in B$ である。

さて、以上の準備の下に、最大最小に関する重要な定理を述べて、締めくくりとしよう。

定理 5.5.3 (教科書の定理 2.7.9, 3.1.8) \mathbb{R}^n の有界閉集合上の連続関数は、必ず最大値、最小値を持つ。

(最後に: この講義を終えるにあたって)

この講義では平均的な学生さんを主な対象としたため、かなりの題材が扱えませんでした。微積の基本的なところは何とか入れようとしたものの、以下のような題材が手薄、または全くやらないままになっています。

- 陰関数定理と逆関数定理。特に逆関数と逆関数定理はやってませんね。
- 曲線、曲面に関すること(教科書の 2.5 節など)。接線、接平面など。
- 包絡線(教科書, p.76 の問 7)。
- 積分の応用としての曲線の長さなど(これは「微積続論」で少しやるはず。)

これらについては適宜、これから補って下さい。

(講義を終えた感想を書こうかと思ったけど、時間切れになってしまった。)