

二次形式と H の凸性について

湯元颯真

そもそも二次形式って必要性があるかを知らないし、数学は嫌だと思っている人が多いと思います。そこで、数学的な説明を除いて、軽くイメージで内容を掴めるように文章を作成してみます。理解するのが難しいとは思いますが頑張ってみてください。

今回は数学的な説明を軽くして直感的に理解できるように文章を作成しました。もし気になる部分があればそのまま chatgpt などにコピペしてください。

(原先生が授業で教えてくださったヘッシャンの計算は実はシルベスター判定法です。)

0.1 二次形式が大切な理由

二次形式とは $x^2 + y^2$ などを見たときみなさんこれがどのような関数かわかりますか？

これは x と y の変化に対して $x^2 + y^2 \geq 0$ となっています。

ここでこれを $\mathbf{x}^t A \mathbf{y} = x^2 + y^2$ とすると、 $A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ となっている。ここで A の形式自体でどのような形状になっているかを予想できたら **嬉しい**。そのためその予想をする方法を考えたいというのが 2 次形式の由来です。

1 変数の凸性について

微分積分のイメージ

1 変数で次の関数を考える

$$q(x) = ax^2 \quad (a \in \mathbb{R}).$$

これは判断基準は位置変数関数についての 2 階微分

$$q''(x) = 2a$$

の符号によるため。

- $a > 0 \Rightarrow$ 下に凸 (convex)
- $a < 0 \Rightarrow$ 上に凸 (concave)

ここでこの 2 回微分での判定とヘッシャンの関係をこれから考えたい。

2 ヘッシャン行列と局所形状

一般の関数 $f: \mathbb{R}^n \rightarrow \mathbb{R}$ (つまりここでは終域がスカラーとなる) を点 \mathbf{x}_0 近傍で 2 次までテイラー展開すると

$$f(\mathbf{x}) \approx f(\mathbf{x}_0) + \nabla f(\mathbf{x}_0)^\top (\mathbf{x} - \mathbf{x}_0) + \frac{1}{2} (\mathbf{x} - \mathbf{x}_0)^\top H(\mathbf{x}_0) (\mathbf{x} - \mathbf{x}_0),$$

ここで **ヘッシャン行列** は

$$H(\mathbf{x}_0) = [\partial^2 f / \partial x_i \partial x_j (\mathbf{x}_0)]_{i,j}$$

であり、対称行列である。二次項

$$q_H(\mathbf{x}) = \mathbf{x}^\top H(\mathbf{x}_0) \mathbf{x}$$

の符号性によって局所的な凹凸が決まる。

1. それぞれの変数について動かすときに必ず正をとるとき **正定値** (positive definite) という
2. それぞれの変数について動かすときに必ず負をとるとき **負定値** (negative definite) という
3. 上記いずれにも当てはまらない場合、**不定値** (indefinite)。

つまり、このヘッシャンが正定値では (極値から微小に動かすと増加する) 下に凸である。負定値では (極値から動かすと減少する) 上に凸となる。

ここで $f(x) = ax^2$ では $H = 2a$ である。

(ここでテイラー展開の 1 次の項についてを考えれば極値についても議論できます。ただ、ここでは凸性を考えたいためその議論を放置します。)

3 多変数二次形式と固有値

この内容は使いにくいので読み飛ばしてもらって構いません。

一応同値な関係の他の判定について軽く触れます。 n 変数ベクトル $\mathbf{x} \in \mathbb{R}^n$ に対し

$$q(\mathbf{x}) = \mathbf{x}^\top A \mathbf{x}, \quad A = A^\top \in \mathbb{R}^{n \times n}.$$

A の固有値	q の形状	凹凸
すべて > 0	お椀型	下に凸
すべて < 0	逆お椀型	上に凸
正負が混在	鞍型	不定 (非凸・非凹)

4 シルベスター判定法 (Sylvester's Criterion)

ヘッシャン H の固有値を直接計算せずに、主座小行列式を用いて正定値／負定値を判定する方法である。

定義

H_k を H の左上 $k \times k$ ブロック (主座部分行列), $\Delta_k = \det H_k$ とする.

判定条件

1. 正定値 (positive definite) $\Delta_1 > 0, \Delta_2 > 0, \dots, \Delta_n > 0$.
2. 負定値 (negative definite) $(-1)^k \Delta_k > 0 \quad (k = 1, 2, \dots, n)$.
3. 上記いずれにも当てはまらない場合, 不定値 (indefinite).

半正定値・半負定値の判定では $\Delta_k \geq 0$ (あるいは交互に符号を変えて ≥ 0) とするが, $\Delta_k = 0$ のみでは真の判別ができない場合があるので注意.

5 例: 2 変数関数

$$f(x, y) = -x^2 - 3y^2 + 2xy.$$

このとき

$$H = \begin{pmatrix} -2 & 2 \\ 2 & -6 \end{pmatrix}, \quad \Delta_1 = -2, \quad \Delta_2 = (-2)(-6) - 2 \times 2 = 8.$$

負定値判定:

$$(-1)^1 \Delta_1 = 2 > 0, \quad (-1)^2 \Delta_2 = 8 > 0.$$

よって H は負定値. 従って f は局所的に 上に凸, 点 $(0, 0)$ は局所最大点である.