VV 285
MID 2 Review

SUN YAN

Content

- The Second Derivative
- Extrema of Potential Functions

The Second Derivative

2.5.1. Definition. Let X, V be finite-dimensional normed vector spaces and $\Omega \subset X$ an open set. A function $f\colon \Omega \to V$ is said to be ***twice differentiable at*** $x \in \Omega$ if

- f is differentiable in an open ball $B_\varepsilon(x)$ around x and
- the function $Df\colon B_\varepsilon(x) \to \mathscr{L}(X, V)$ is differentiable at x.

We say that f is twice differentiable on Ω if f is twice differentiable at every $x \in \Omega$.

Actually, the second derivative is a map:

$$D(Df) =: D^2 f\colon \Omega \to \mathscr{L}(X, \mathscr{L}(X, V)).$$

And is found by:

$$Df|_{x+h} = Df|_x + D^2 f|_x h + o(h) \qquad \text{as } h \to 0.$$

Example

- Calculate the second, third and forth derivatives of the map,
- $\Phi : \mathrm{Mat}(n \times n; \mathbb{R}) \to \mathrm{Mat}(n \times n; \mathbb{R}),\ \Phi(A) = A^3$

$D\Phi|_A = AHA+AAH+HAA$

$D^2\Phi|_A(H,J)= JHA + AHJ + AJH + JAH + AHJ + AJH$

$D^3\Phi|_A(H, J, K)= JHK + KHJ + KJH + JKH + KHJ + KJH$

$D^4\Phi|_A = 0$

The Hessian

The Hessian. For a differentiable potential function $f : \mathbb{R}^n \to \mathbb{R}$. The derivative is given by the Jacobian:

$$Df|_x = \begin{pmatrix} \left.\frac{\partial f}{\partial x_1}\right|_x & \cdots & \left.\frac{\partial f}{\partial x_n}\right|_x \end{pmatrix} \quad , \qquad Df|_x \in \mathcal{L}(\mathbb{R}^n, \mathbb{R})$$

and the second derivative is found by ***Hessian*** where

$$\mathrm{Hess}\, f(x) = D(\nabla f)|_x = \begin{pmatrix} \left.\frac{\partial^2 f}{\partial x_1 \partial x_1}\right|_x & \left.\frac{\partial^2 f}{\partial x_2 \partial x_1}\right|_x & \cdots & \left.\frac{\partial^2 f}{\partial x_n \partial x_1}\right|_x \\ \vdots & \vdots & & \vdots \\ \left.\frac{\partial^2 f}{\partial x_1 \partial x_n}\right|_x & \left.\frac{\partial^2 f}{\partial x_2 \partial x_n}\right|_x & \cdots & \left.\frac{\partial^2 f}{\partial x_n \partial x_n}\right|_x \end{pmatrix}$$

$$D^2 f|_x h = \mathrm{Hess}\, f(x) h$$

Bilinear Maps

Note that the expression $\langle \operatorname{Hess} f(x)h, \tilde{h}\rangle$ is linear in both h and $\tilde{h}$; hence we can also regard the second derivative as a ***bilinear map***

$$D^2 f|_x \colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}, \qquad (h, \tilde{h}) \mapsto \langle \operatorname{Hess} f(x)h, \tilde{h}\rangle.$$

2.5.3. Definition. Let X, V be finite-dimensional normed vector spaces. The set of multilinear maps from X to V is denoted by

$$\mathscr{L}^{(n)}(X, V) := \left\{ L \colon X \times \cdots \times X \to V \colon L \text{ linear in each component} \right\}.$$

In the special case $V = \mathbb{R}$ an element of $\mathscr{L}^{(n)}(X, V)$ is called a ***multilinear form***.

Bilinear form in R^n

- Every linear map $L \in (R^n)^*$ has the form $L = \langle z, \cdot \rangle$ for some $z \in R^n$.
- Interpret an element $\boldsymbol{A \in L(R^n, L(R^n, R))}$ as a linear map, $A: y \to L_y := \langle z_y, \cdot \rangle, z_y = A(y)$.
- A is actually a matrix $A: y \to z_y$.
- For every $y \in R^n$ we obtain a linear map $\langle Ay, \cdot \rangle \in L(R^n, R)$.
- Then $L_y x = \langle Ay, x \rangle = L(x, y)$.

Schwarz's Theorem

2.5.5. Schwarz's Theorem. Let X, V be normed vector spaces and $\Omega \subset X$ an open set. Let $f \in C^2(\Omega, V)$. Then $D^2 f|_x \in \mathscr{L}^{(2)}(X \times X, V)$ is symmetric for all $x \in \Omega$, i.e.,

$$D^2 f(u, v) = D^2 f(v, u), \qquad \text{for all } u, v \in X.$$

- This implies that if *f* is **twice continuously differentiable**, the Hessian of *f* at *x* is symmetric.

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}.$$

In other words, if f is twice continuously differentiable, the order of differentiation in the second-order partial derivatives does not matter.

Schwarz's Theorem

- Note: The Schwarz's theorem will hold if all the second-order partial derivatives **are continuous**.
- It will break if the function fails to differentiable partial derivative.
- Example:
- f(x,y)= $\begin{cases} \frac{xy(x^2-y^2)}{x^2+y^2}. & for(x,y)! = (0,0) \\ 0 & for(x,y) = (0,0) \end{cases}$
- Since the second partial derivative is **not continuous** at (0,0)
- Try to see that it is not symmetric

Schwarz's Theorem

- Try to see that it is not symmetric
- f(x,y)= $\begin{cases} \frac{xy(x^2-y^2)}{x^2+y^2}. & for(x,y)! = (0,0) \\ 0 & for(x,y) = (0,0) \end{cases}$
- Try to calculate

$$\frac{\partial^2}{\partial x \partial y} f|_{(0,0)} \qquad \frac{\partial^2}{\partial y \partial x} f|_{(0,0)}$$

Exercise 1-1

- Find the first and second derivatives of the following functions:
- 1. f :Mat($n \times n$;R)$\rightarrow$R, f(A)=tr(A^2)
- 2. g :Mat($n \times n$;R)$\rightarrow$R, f(A)=$(\mathrm{tr}A)^2$

Exercise 1-2

- Find the second derivative of the determinant of an invertible matrix A.

Det: Mat(n x n;R)->R

The Frenet–Serret formulas (in R^3)

$$\frac{d\mathbf{T}}{ds} = \kappa\mathbf{N},$$
$$\frac{d\mathbf{N}}{ds} = -\kappa\mathbf{T} + \tau\mathbf{B},$$
$$\frac{d\mathbf{B}}{ds} = -\tau\mathbf{N},$$

$$T(t) = \frac{\gamma'(t)}{\|\gamma'(t)\|} \qquad \kappa = \frac{\|\gamma'(t) \times \gamma''(t)\|}{\|\gamma'(t)\|^3}$$

$$N(t) = \frac{(T \circ \gamma)'(t)}{\|(T \circ \gamma)'(t)\|} \qquad \tau = \frac{\det(\gamma'(t), \gamma''(t), \gamma'''(t))}{\|\gamma'(t) \times \gamma''(t)\|^2}$$

$$B(t) = T \times N$$

Content

- The Second Derivative
- Extrema of Potential Functions

Quadratic Approximation of Potential Functions

2.6.1. Lemma. Let $\Omega \subset \mathbb{R}^n$ be an open set and $f \in C^2(\Omega, \mathbb{R})$. Then for any $h \in \mathbb{R}^n$ small enough that $x + h \in \Omega$,

$$f(x+h) = f(x) + \langle \nabla f(x), h \rangle + \int_0^1 (1-t)\langle \operatorname{Hess} f(x+th)h, h \rangle \, dt. \tag{2.6.2}$$

2.6.2. Corollary. Let $\Omega \subset \mathbb{R}^n$ be an open set and $f \in C^2(\Omega, \mathbb{R}))$. Then, as $h \to 0$,

$$f(x+h) = f(x) + \langle \nabla f(x), h \rangle + \frac{1}{2}\langle \operatorname{Hess} f(x)h, h \rangle + o(h^2). \tag{2.6.3}$$

Quadratic Form

2.6.3. Definition. Let $A \in \mathrm{Mat}(n \times n, \mathbb{R})$. Then the ***quadratic form induced by A*** is defined as the map

$$Q_A := \langle \,\cdot\,, A(\,\cdot\,)\rangle, \qquad x \mapsto \langle x, Ax\rangle = \sum_{j,k=1}^{n} a_{jk} x_j x_k, \qquad x \in \mathbb{R}^n.$$

Clearly, $Q_A(\lambda x) = \lambda^2 Q_A(x)$ for any $\lambda \in \mathbb{R}$. Note also that Q_A is continuous, because it is a polynomial in $x_1, \dots, x_n$.

Quadratic Form

2.6.4. Definition. A quadratic form Q_A induced by a matrix $A \in \mathrm{Mat}(n \times n, \mathbb{R})$ is called

- ***positive definite*** if $Q_A(x) > 0$ for all $x \neq 0$,
- ***negative definite*** if $Q_A(x) < 0$ for all $x \neq 0$,
- ***indefinite*** if $Q_A(x_0) > 0$ for some $x_0 \in \mathbb{R}^n$ and $Q_A(y_0) < 0$ for some $y_0 \in \mathbb{R}^n$.

A matrix A is said to be negative definite / positive definite / indefinite if the induced quadratic form Q_A has the corresponding property.

Criteria for Definiteness

2.6.7. Lemma. Let $A \in \mathrm{Mat}(2 \times 2, \mathbb{R})$ be symmetric, i.e.,

$$A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}.$$

Let $\Delta = \det A$. Then

(i) A positive definite $\Leftrightarrow a > 0$ and $\Delta > 0$

(ii) A negative definite $\Leftrightarrow a < 0$ and $\Delta > 0$

(iii) A indefinite $\Leftrightarrow \Delta < 0$

2.6.8. Lemma. The matrix $A \in \mathrm{Mat}(n \times n, \mathbb{R})$ is

$$\text{positive definite} \quad \Leftrightarrow \quad \underset{\alpha>0}{\exists} \ \underset{x\in\mathbb{R}^n}{\forall} \ Q_A(x) \geq \alpha \|x\|^2$$

$$\text{negative definite} \quad \Leftrightarrow \quad \underset{\alpha>0}{\exists} \ \underset{x\in\mathbb{R}^n}{\forall} \ Q_A(x) \leq -\alpha \|x\|^2$$

Extrema of Real Functions

2.6.10. Definition. Let $\Omega \subset \mathbb{R}^n$ and $f\colon \Omega \to \mathbb{R}$.

(i) f is said to have a ***local maximum*** at $\xi \in \Omega$ if there exists a $\delta > 0$ such that

$$x \in \Omega \cap B_\delta(\xi) \qquad \Rightarrow \qquad f(x) \le f(\xi).$$

(ii) f is said to have a ***strict local maximum*** at $\xi \in \Omega$ if there exists a $\delta > 0$ such that

$$x \in \Omega \cap B_\delta(\xi) \setminus \{\xi\} \qquad \Rightarrow \qquad f(x) < f(\xi).$$

Extrema of Real Function

2.6.11. Theorem. Let $\Omega \subset \mathbb{R}^n$, $f \colon \Omega \to \mathbb{R}$ and $\xi \in \operatorname{int} \Omega$. Assume that all partial derivatives of f exist at ξ and that f has a local extremum (maximum or minimum) in ξ. Then

$$\nabla f(\xi) = 0.$$

If f is differentiable in ξ, this implies $Df|_\xi = 0$.

2.6.12. Theorem. Let $\Omega \subset \mathbb{R}^n$ be open, $f \in C^2(\Omega)$ and $\xi \in \Omega$. Let $\nabla f(\xi) = 0$ (i.e., $Df|_\xi = 0$).

(i) If $\operatorname{Hess} f|_\xi$ is positive definite, f has a strict local minimum at ξ.

(ii) If $\operatorname{Hess} f|_\xi$ is negative definite, f has a strict local maximum at ξ.

(iii) If $\operatorname{Hess} f|_\xi$ is indefinite, f has no extremum at ξ.

Extrema of Real Functions

2.6.13. Corollary. Let $\Omega \subset \mathbb{R}^2$ be open, $f \in C^2(\Omega)$ and $\xi \in \Omega$ with $\nabla f(\xi) = 0$. Set

$$\Delta := \det \operatorname{Hess} f|_{\xi} = \frac{\partial^2 f}{\partial x_1^2}\bigg|_{\xi} \frac{\partial^2 f}{\partial x_2^2}\bigg|_{\xi} - \left(\frac{\partial^2 f}{\partial x_1 \partial x_2}\bigg|_{\xi}\right)^2$$

Then $f(\xi)$ is

- a local minimum if $\frac{\partial^2 f}{\partial x_1^2}\big|_{\xi} > 0$ and $\Delta > 0$,
- a local maximum if $\frac{\partial^2 f}{\partial x_1^2}\big|_{\xi} < 0$ and $\Delta > 0$,
- no extremum if $\Delta < 0$.

Note that if $\Delta = 0$, Corollary 2.6.13 yields no information.

Finding Extrema

In searching for extrema of functions $f \in C^2(\Omega, \mathbb{R})$, we follow a four-step process:

1. Check for critical points $\xi \in \operatorname{int} \Omega$, i.e., those were $Df|_\xi = 0$.
2. Use Theorem 2.6.12 or Corollary 2.6.13 to check which of the critical points is an extremum.
3. Check the boundary $\partial\Omega$ separately for local extrema.
4. Identify the global extrema. Any finite global extremum must also be a local extremum, so it will be included among those found above.

Reference

Hohberger, Horst. “VV285_main.pdf”

Serge Lang, Calculus of Several Variables, 1987

Chen, Xiwen “RC_7.pdf”,2018

Yuan, Jian “RC_7.pdf”,2018

Thank You

Have fun in Mid 2!