Mathematical Analysis II

Felix Chen

Contents

§0.1 Convex functions

Definition 0.1.1 (Hesse matrix)**.** Let $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ be a C^2 function, we call the Jacobi matrix of ∇f to be the **Hesse matrix** of f. (Also called Hessian matrix)

$$H_f(p) = \nabla^2 f(p) = \left(\frac{\partial}{\partial x_i}\frac{\partial}{\partial x_j} f(p)\right)_{i,j}.$$

Since the partial derivatives commute, so H_f is a symmetrical matrix, hence diagonalizable.

Proposition 0.1.2

Let $f \in C^2(\Omega)$, let x_0 be a minimum of f, then $\nabla f(x_0) = 0$, and $H_f(x_0)$ is semi positive definite.

Proof. By Taylor's expansion,

$$f(x) = f(x_0) + \mathrm{d}f(x_0)(x - x_0) + \frac{1}{2}(x - x_0)^T H_f(x_0)(x - x_0) + o(|x - x_0|^2).$$

If $H_f(x_0)$ has a negative eigenvalue $-\lambda$, with eigenvector v, then $f(x_0 + tv) = f(x_0) - \frac{1}{2}\lambda t^2|v|^2 + o(|tv|^2)$, which contradicts with the minimality of x_0. □

Proposition 0.1.3

If $\nabla f(x_0) = 0$, $H_f(x_0)$ is positive definite, then x_0 is a local minimum of f.

Proof. Same as previous one. □

Definition 0.1.4 (Convex functions)**.** If f and Ω satisfies:

$$\forall x, y \in \Omega, tx + (1-t)y \in \Omega, \quad f(tx + (1-t)y) \le tf(x) + (1-t)f(y)$$

we say Ω is a **convex set** and f a **convex function**.

Theorem 0.1.5 (Jensen's inequality)

Let f be a convex function on Ω. Real numbers $t_i \geq 0, \sum_{i=1}^N t_i = 1$, for $x_i \in \Omega$,

$$f\left(\sum_{i=1}^N t_i x_i\right) \leq \sum_{i=1}^N t_i f(x_i).$$

Example 0.1.6 (Convex functions)

Linear functions $f(x) = Ax + b$ are convex.

The norm function $\|\cdot\| : \mathbb{R}^n \to \mathbb{R}$ is convex. Also let A be an $n \times n$ positive definite matrix, then $f(x) = x^T Ax$ is convex.

Just like the one dimensional case, convex functions have nice properties.

Theorem 0.1.7

Let f be a convex function on an open convex set Ω, then f is continuous, and Lipschitz continuous in any compact set, i.e.

$$|f(x) - f(y)| \leq M|x - y|, \quad x, y \in U$$

where U is a compact set.

Proof. WLOG $0 \in \Omega$, take an orthogonal basis $e_1, \ldots, e_n$. Let

$$x = \sum_{i=1}^n \lambda_i \bar{e}_i, \quad \bar{e}_i = e_i \text{ or } -e_i, \lambda_i \geq 0.$$

When $|x|$ sufficiently small, $\sum_{i=1}^n \lambda_i < 1$, so by Jensen's inequality,

$$f(x) \leq \sum_{i=1}^n \lambda_i f(\bar{e}_i) + \lambda f(0),$$

$$f(x) - f(0) \leq \sum_{i=1}^n \lambda_i (f(\bar{e}_i) - f(0)) \leq \left(\sum_{i=1}^n \lambda_i^2\right)^{\frac{1}{2}} \left(\sum_{i=1}^n (f(\bar{e}_i) - f(0))^2\right)^{\frac{1}{2}} \leq |x|C,$$

since we can change the length of e_i, and f is continuous on a straight line.

This means f is continuous. For the second part, let $\lambda_0 = \frac{1}{1+\sum_{i=1}^n \lambda_i}$, since $0 = \lambda_0 x + \sum_{i=1}^n \lambda_0 \lambda_i (-\bar{e}_i)$, by Jensen's inequality, we'll get the desired property. $\square$

Proposition 0.1.8

Let f be a differentiable function on a covex set Ω,

f is convex $\iff f(x) \geq f(x_0) + \mathrm{d}f(x_0)(x - x_0)$.

Proof. If f is convex, just use the definition and let $t \to 0$:

$$f(x_0) + f'(x_0)t(x - x_0) + o(t(x - x_0)) \le tf(x) + (1-t)f(x_0).$$

Conversely, let $z = tx + (1-t)x_0$,

$$f(x) \ge f(z) + f'(z)(1-t)(x - x_0), f(x_0) \ge f(z) + f'(z)t(x_0 - x).$$

Thus adding these together we get

$$tf(x) + (1-t)f(x_0) \ge f(z).$$

□

Theorem 0.1.9

Let $\Omega \subset \mathbb{R}^n$ be an open convex set, $f \in C^2(\Omega)$, f convex $\iff$ $H_f(x)$ semi positive definite.

Proof. One direction can be proved using Taylor's expansion.

On the other hand, let $H(t) = f(x_0 + t(x - x_0)) - f(x_0) - t\,\mathrm{d}f(x_0)(x - x_0)$, then $H'(t) = \mathrm{d}f(x_0 + t(x - x_0))(x - x_0) - \mathrm{d}f(x_0)(x - x_0)$,

$$H''(t) = (x - x_0)^T H_f(p)(x_0 + t(x - x_0))(x - x_0) \ge 0.$$

So $H(t)$ is a convex function, $H(0) = 0, H'(0) = 0$. □

§1 Integrals on surfaces

§1.1 Measures on manifolds

To define integrals, we need to define a measure on it first.

For example, let $v_1, \dots, v_d \in \mathbb{R}^n$ be linearly independent vectors, and unit vectors $v_{d+1}, \dots, v_n$ complete them to a basis, satisfying $v_j \perp v_i, j > d, j > i$.

Let A be a linear map s.t. $Ae_i = v_i$, then the volume of $A(E)$ is $|\det A| = \sqrt{\det(G \cdot G^T)}$, where $G = \begin{pmatrix} v_1 \\ \vdots \\ v_d \end{pmatrix}$ is a $d \times n$ matrix.

Since $AA^T = \begin{pmatrix} GG^T & 0 \\ 0 & I_{n-d} \end{pmatrix}$, $|\det A| = \sqrt{\det GG^T}$, we say GG^T is the **Gram matrix** of G.

Another example is the length of a curve. Recall that we have the formula

$$L(\gamma) = \int_0^1 |\gamma'(t)|\,\mathrm{d}t.$$

The length of a curve is essentially the "volume" of a 1-dimensional manifold, so the idea of higher dimensional manifold is the same.

Definition 1.1.1. Let M be a manifold in $\mathbb{R}^n$. Let $\Phi : V \subset \mathbb{R}^d \to U \subset M$ be a smooth homeomorphism, $\operatorname{rank}\Phi = d$. We can split U to many small regions and use the paraloids to approximate the volume of each regoin.

Thus we define:

$$m(U) = \int_V \sqrt{\det(\mathrm{d}\Phi(x)^T \, \mathrm{d}\Phi(x))} \, \mathrm{d}x_1 \, \mathrm{d}x_2 \cdots \mathrm{d}x_d.$$

The above differential form inside the integral is called the **volume form**:

$$\mathrm{d}\sigma(y) = \sqrt{\det(\mathrm{d}\Phi^T \, \mathrm{d}\Phi)} \, \mathrm{d}y.$$

Moreover, with this measure we can define the integral of general measurable function f (measurable means locally measurable on $\mathbb{R}^d$):

$$\int_U f \, \mathrm{d}\sigma = \int_V f(\Phi(x)) \sqrt{\det(\mathrm{d}\Phi^T \, \mathrm{d}\Phi)} \, \mathrm{d}x.$$

Just like the length of a curve, the volume defined above is also a geometric quantity, i.e. independent of coordinates.

Example 1.1.2

Let $d = 1$, $\gamma : (-1, 1) \to \mathbb{R}^n$, $\gamma'(0) \neq 0$. For fixed $-1 < a < b < 1$ and a function f on γ, let C_a^b denote the curve between $\gamma(a), \gamma(b)$,

$$\int_{C_a^b} f \, \mathrm{d}\sigma = \int_a^b f(\gamma(t)) |\gamma'(t)| \, \mathrm{d}t$$

is called the **curve integral of the first type**.

Example 1.1.3

Let $d = n - 1$, $f : \mathbb{R}^{n-1} \to \mathbb{R}$, the graph of f is a hyper-surface $\Gamma_f = \{(x, f(x)) \mid x \in \mathbb{R}^{n-1}\}$. It has a parametrization $\Phi(x) = (x, f(x))$, so

$$\mathrm{d}\Phi = \begin{pmatrix} I_{n-1} \\ \nabla f \end{pmatrix} \implies \mathrm{d}\Phi^T \, \mathrm{d}\Phi = I_{n-1} + \nabla f^T \cdot \nabla f.$$

Hence $\det(\mathrm{d}\Phi^T \, \mathrm{d}\Phi) = 1 + |\nabla f|^2$. (This can be obtained by looking at the eigenvectors)

Therefore for φ on $\mathbb{R}^n$, we have

$$\int_{\Gamma_f} \varphi \, \mathrm{d}\sigma = \int_{\mathbb{R}^{n-1}} \varphi(x, f(x)) \sqrt{1 + |\nabla f|^2} \, \mathrm{d}x.$$

Next we'll compute the surface area of unit sphere S^{n-1}.

Let c_n denote the volume of unit sphere in $\mathbb{R}^n$,

$$c_n = \frac{\pi^{\frac{n}{2}}}{\Gamma(\frac{n}{2} + 1)}.$$

We claim in advance that the surface area of unit sphere $\omega_{n-1} = n c_n$. Here we use the sphere coordinates:

$$x_1 = r \prod_{k=1}^{n-1} \sin\theta_k, \quad x_i = r \sin\theta_{n-1} \cdots \sin\theta_i \cos\theta_{i-1}, 2 \leq i \leq n.$$

Let $F_n(r, \theta_1, \dots, \theta_{n-1}) = (x_1, \dots, x_n)$.

$$\mathrm{d}F_n = \begin{pmatrix} \sin\theta_{n-1}\,\mathrm{d}F_{n-1} & \cos\theta_{n-1}F_{n-1}^T \\ (\cos\theta_{n-1}, 0, \dots, 0) & -r\sin\theta_{n-1} \end{pmatrix}.$$

So we can compute its determinant (expand using the last row), note that the first column of $\mathrm{d}F_{n-1}$ is $r^{-1}F_{n-1}^T$,

$$\begin{aligned}\det \mathrm{d}F_n = &- r\sin\theta_{n-1}(\sin\theta_{n-1})^{n-1}\det(\mathrm{d}F_{n-1}) \\ &+ (-1)^{n-1}(\cos\theta_{n-1})^2(-1)^{n-2}r(\sin\theta_{n-1})^{n-2}\det(\mathrm{d}F_{n-1}) \\ = &- r(\sin\theta_{n-1})^{n-2}\det(\mathrm{d}F_{n-1}).\end{aligned}$$

Hence $\mathrm{d}x = r^{n-1}\sin^{n-2}\theta_{n-1}\cdots\sin\theta_2\,\mathrm{d}r\,\mathrm{d}\theta_1\cdots\mathrm{d}\theta_{n-1} = r^{n-1}\,\mathrm{d}r\,\mathrm{d}\omega$.

Denote F_n^S to be the function F_n restricted to S^{n-1}. Then $\mathrm{d}F_n = (r^{-1}F_n^T, \mathrm{d}F_n^S)$. We can compute that the Gram determinant of $\mathrm{d}F_n^S$ is just $\det \mathrm{d}F_n$ with $r = 1$.

The rest is some integrals with gamma function and beta function, which is left out.

§1.2 Stolkes' formula

Intuitively, Stolkes' formula states that: Let D be a region, $\mathrm{d}\omega$ be a differential form, then

$$\int_D \mathrm{d}\omega = \int_{\partial D}\omega.$$

Here ∂D means the "boundary" of D.

Of course we need some "regularity" requirements of D and ω, and it's the generalization of Newton-Lebniz formula into higher dimensions.

Definition 1.2.1 (Bounded regions with boundary)**.** Let $\Omega \subset \mathbb{R}^n$ be a compact set, we say it's a **bounded region with boundary** if $\forall x \in \partial\Omega$, there exists open sets $U, V \subset \mathbb{R}^n$, $x \in U$ and a continuous homeomorphism $\Phi : U \to V$, such that

$$\Phi(U \cap \Omega) = V \cap \{x_n \geq 0\}, \quad \Phi(U \cap \partial\Omega) = V \cap \{x_n = 0\}.$$

If Φ is also C^1, we say $x \in \partial\Omega$ is a **regular point**, otherwise a **singular point**.

Lemma 1.2.2

Let Ω be a bounded region with boundary, for all regular $p \in \partial\Omega$, there exists a unique unit vector $\nu(p) \in \mathbb{R}^n$, and $\varepsilon > 0$, s.t.

$$\nu(p) \perp T_p\partial\Omega, |\nu(p)| = 1, p - t\nu(p) \in \Omega, \forall 0 < t < \varepsilon.$$

We call $\nu(p)$ the **outward unit normal vector** of p.

Proof. By the definition of regular points, we may assume that:

$$\Omega \cap V = \{x \in V \mid f(x) \geq 0\}, \quad \partial\Omega \cap V = \{x \in V \mid f(x) = 0\}.$$

Where f is a C^1 function.

Since ∇f is nonzero, the tangent space $T_p\partial\Omega = \{v \mid v \cdot \nabla f = 0\}$.

Let $\nu(p) = -\frac{\nabla f}{|\nabla f|}$, then it's obvious $\nu(p)$ points outside of Ω. $\square$

Now for a cuboid I and a C^1 function ϕ,

$$\int_I \frac{\mathrm{d}\phi}{\mathrm{d}x_n}\,\mathrm{d}x = \int_{I_{n-1}} \phi(x_1,\dots,x_{n-1},b_n)\,\mathrm{d}x_1\cdots\mathrm{d}x_{n-1} - \int_{I_{n-1}} \phi(x_1,\dots,x_{n-1},a_n)\,\mathrm{d}x_1\cdots\mathrm{d}x_{n-1}$$
$$= \int_{\partial I} \phi\cdot\nu_n\,\mathrm{d}\sigma.$$

Where σ is the measure on the boundary, ν is the outward unit normal vector.

Lemma 1.2.3

Let K be a compact set in $\mathbb{R}^n$, $U \supset K$ is open, there exists a smooth function f such that $\operatorname{supp} f \subset U$, and $f\big|_K > 0$.

Proof. Let $\rho(x)$ be a smooth function s.t. $\rho(x) = 1$ for $|x| \le 1$ and $\rho(x) = 0$ for $|x| \ge 2$. Let

$$g(x) = \int_{|y|\le 2} f(x-\delta y)\rho(y)\,\mathrm{d}y.$$

Then g is a smooth non-negative function. □

Theorem 1.2.4 (Unit decomposition on compact sets)

Let K be a compact set, $\{U_1,\dots,U_k\}$ is an open covering of K. There exists smooth functions $f_1,\dots,f_k$ s.t.

$$1 = f_1(x) + f_2(x) + \cdots + f_k(x), \quad \operatorname{supp} f_i(x) \subset U_i.$$

Proof. For $1 \le i \le k$, $\delta > 0$, define

$$K_i^\delta = \{x \in U_i \mid d(x, U_i^c) > \delta\}.$$

Note that $\{\bigcup_{i=1}^k K_i^{\frac{1}{m}}\}_{m=1}^\infty$ is also an open covering of K, thus there exists N s.t.

$$K \subset \bigcup_{i=1}^k K_i^{\frac{1}{N}}.$$

Hence by lemma we have g_i s.t. $\operatorname{supp} g_i \subset U_i$ and $g_i > 0$ on the closure of $K_i^{\frac{1}{N}}$. Similarly we have a smooth function g s.t. $g(x) = 0$ on K, and $g > 0$ outside of the closure of $\bigcup_{i=1}^k K_i^{\frac{1}{N}}$.

Let $G(x) = g_1(x) + \cdots + g_k(x) + g(x) > 0$ on $\bigcup_{i=1}^k U_i$, then we can define $f_i(x) = \frac{g_i(x)}{G(x)}$ which satisfy the condition. □

Theorem 1.2.5

Let Φ be a C^1 homeomorphism from a cuboid I to Ω, then Ω satisfies Stolkes' formula: $\forall \phi \in C^1(\mathring{D}) \cap C(\overline{D})$, we have

$$\int_D \nabla\phi\,\mathrm{d}x = \int_{\partial D} \phi\nu\,\mathrm{d}\sigma.$$

Proof. Since $\Omega = \Phi(I)$, let y be the coordinates on I, $x = \Phi(y)$,

$$\int_{\Phi(I)} \nabla\varphi \,\mathrm{d}x = \int_I \nabla\varphi(\Phi(y))(\mathrm{d}\Phi)^{-1} J_\Phi \,\mathrm{d}y.$$

Let $A = \mathrm{d}\Phi$, WLOG $J_\Phi > 0$. Using the index notation and Einstein summation,

$$A_{kj}A^{ji} = A^{kj}A_{ji} = \delta_{ki}.$$

Thus

$$\partial_{y_j}\varphi A^{ji}|A| = \partial_{y_j}(\varphi A^{ji}|A|) - \varphi\partial_{y_j}(A^{ji}|A|)$$

Since $|A| = A_{kl}A^{kl}|A|$, $A_{kl} = \frac{\partial\Phi_k}{\partial y_l}$.

$$\begin{aligned}
\partial_{y_j}(A^{ji}|A|) &= |A|\partial_{y_j}A^{ji} + A^{ji}\partial_{y_j}|A| \\
&= |A|\partial_{y_j}A^{ji} + A^{ji}|A|\partial_{y_j}A_{kl}A^{kl} \\
&= |A|(\partial_{y_j}A^{ji} + \partial_{y_l}A_{kj}A^{kl}) \\
&= |A|(\partial_{y_j}A^{ji} - \partial_{y_j}A^{ji}) = 0.
\end{aligned}$$

Hence by our previous work,

$$\int_I \partial_{y_j}(\varphi A^{ji}|A|)\,\mathrm{d}y = \int_{\partial I} \varphi A^{ji}|A|\nu_j \,\mathrm{d}\sigma.$$

Putting this together for all i's, note that $\widetilde{\nu} = \frac{\nabla\Phi_n^{-1}}{|\nabla\Phi_n^{-1}|}$, □