

Question 8.5.2.m 个变量情况下的勒让德变换. 从变量 x^1, \dots, x^m 和函数 $f(x^1, \dots, x^m)$ 到新变量 ξ_1, \dots, ξ_m 和新函数 $f^*(\xi_1, \dots, \xi_m)$ 的勒让德变换由下列关系给出

$$\begin{cases} \xi_i = \frac{\partial f}{\partial x^i}(x) \\ f^*(\xi) = \sum_{i=1}^m \xi_i x^i - f(x) \end{cases}$$

a): 给出勒让德变换的几何解释: 他把函数 $f(x)$ 的图像上的点的坐标 $(x^1, \dots, x^m, f(x))$ 变换到此处切平面方程的参数 $(\xi_1, \dots, \xi_m, f(\xi))$.

b): 证明: 如果 $f \in C^{(2)}$, $\det \frac{\partial^2 f}{\partial x^i \partial x^j} \neq 0$, 则勒让德变换在此处局部必然存在。

c): 对于函数 $f(x) = f(x^1, \dots, x^m)$, 定义凸函数, 并证明凸函数经过勒让德变换后仍然是凸函数。

d): 证明

$$df^* = \sum_{i=1}^m x^i d\xi_i + \sum_{i=1}^m \xi_i dx^i - df = \sum_{i=1}^m x^i d\xi_i$$

并由此推出勒让德变换是对合变换, 即

$$(f^*)^*(x) = f(x)$$

e): 将变换写为对称形式

$$\begin{cases} f^*(\xi) + f(x) = \sum_{i=1}^m \xi_i x^i \\ \xi_i = \frac{\partial f}{\partial x^i}(x) \\ x^i = \frac{\partial f^*}{\partial \xi_i}(\xi) \end{cases}$$

并简写为

$$f^*(\xi) + f(x) = \xi^T x, \quad \xi = \nabla f(x), \quad x = \nabla f^*(\xi)$$

f): 函数 f 与 f^* 的黑塞矩阵为

$$\begin{bmatrix} \frac{\partial^2 f}{\partial x^1 \partial x^1} & \cdots & \frac{\partial^2 f}{\partial x^1 \partial x^m} \\ \vdots & & \vdots \\ \frac{\partial^2 f}{\partial x^m \partial x^1} & \cdots & \frac{\partial^2 f}{\partial x^m \partial x^m} \end{bmatrix} (x), \quad \begin{bmatrix} \frac{\partial^2 f^*}{\partial \xi^1 \partial \xi^1} & \cdots & \frac{\partial^2 f^*}{\partial \xi^1 \partial \xi^m} \\ \vdots & & \vdots \\ \frac{\partial^2 f^*}{\partial \xi^m \partial \xi^1} & \cdots & \frac{\partial^2 f^*}{\partial \xi^m \partial \xi^m} \end{bmatrix} (\xi)$$

设 d_{ij} 与 d_{ij}^* 分别为上述矩阵的代数余子式, 而 d, d^* 代表其行列式, 认为 $d \neq 0$, 证明

$$dd^* = 1, \quad \frac{\partial^2 f}{\partial x^1 \partial x^1}(x) = \frac{d_{11}^*}{d^*}(\xi), \quad \frac{\partial^2 f^*}{\partial \xi^1 \partial \xi^1}(\xi) = \frac{d_{11}}{d}(x)$$

g): 一个线圈上形成的肥皂泡膜构成了通常所说的极小曲面, 它是以该线圈为边界的所有曲面中面积最小的曲面。如果在局部用函数 $z = f(x, y)$ 给出极小曲面, 则函数 f 满足以下方程

$$(1 + f_y'^2)f_{xx}'' - 2f_x'f_y'f_{xy}'' + (1 + f_x'^2)f_{yy}'' = 0$$

证明这个方程在勒让德变换下转化为

$$(1 + \eta^2)f_{\eta\eta}^{*''} - 2\xi\eta f_{\xi\eta}^{*''} + (1 + \xi^2)f_{\xi\xi}^{*''} = 0$$

Solution. a): $y = f(x^1, \dots, x^m)$ 是 $m+1$ 维中的 m 维曲面, 在 x 处的切平面可以写为 $y_t - y = \partial_i f(x)(x_t - x)^i$ 。化为

$$\partial_i f(x)x^i - y = \partial_i f(x)x_t^i - y_t^i$$

其中 (x_t, y_t) 是切空间中的点, 故对于切空间中的任何一点上述式子成立, 令 $\xi = \partial_i f(x)$, $f^*(\xi) = \xi_i x^i - f(x)$, 在给定的 x 处, 这几个参数是不会改变的。

b): 即验证满足 $\det \frac{\partial^2 f}{\partial x^i \partial x^j} \neq 0$ 时 $(x^1, \dots, x^m, f(x)) \mapsto (\xi_1, \dots, \xi_m, f^*(\xi))$ 在局部同胚。将变换

$$\begin{cases} \xi_1 = \partial_1 f(x) \\ \vdots \\ \xi_m = \partial_m f(x) \\ \eta = \partial_i f(x)x^i - y \end{cases}$$

视作 $\Xi = g(X)$, 改写为 $G(\Xi, X) = \Xi - g(X) = 0$, 对于曲面上的任何一点, 首先 $G(\Xi, X) = 0$ 成立, 其次易证 $g \in C^{(1)}$, 故 $G \in C^{(1)}$, 最后要使变换可逆, 只需要 $G'_X(\Xi, X)$ 可逆

$$G'_X = \begin{bmatrix} \partial_{11}f & \cdots & \partial_{1m}f & 0 \\ \vdots & & \vdots & \vdots \\ \partial_{m1}f & \cdots & \partial_{mm}f & 0 \\ x^i \partial_{i1}f & \cdots & x^i \partial_{im}f & 1 \end{bmatrix} (x)$$

可见 $\det \frac{\partial^2 f}{\partial x^i \partial x^j} \neq 0$ 时 G'_X 可逆。

c): 该问会采取两种方法计算。第一种是使用二阶导数, 第二种会用到勒让德变换的另一种表达形式。第一种方法会复杂许多, 根据一元函数的凸函数定义写出多元下的定义: $f: \mathbb{R}^m \rightarrow \mathbb{R}$

$$f(\theta x_1 + (1-\theta)x_2) \leq \theta f(x_1) + (1-\theta)f(x_2)$$

首先证明: $\frac{\partial^2 f}{\partial x^i \partial x^j}$ 半正定等价于 $f(x)$ 在 x 的领域内是凸函数。令 $z = \theta x_1 + (1-\theta)x_2 = x_2 + \theta(x_1 - x_2) = x_1 + (1-\theta)(x_2 - x_1)$ 。将 f 泰勒展开

$$f(x+h) = f(x) + f'(x)h + h^T f''(x)h + o(\|h\|^2)$$

由 $f''(x)$ 半正定, 可得

$$\begin{aligned} f(x_1) &= f(z) + f'(z)(x_1 - z) + (x_1 - z)^T f''(z)(x_1 - z) + o(\|x_1 - z\|^2) \\ f(x_1) &\geq f(z) + f'(z)(\theta - 1)(x_2 - x_1) \\ f(x_2) &\geq f(z) + f'(z)\theta(x_2 - x_1) \end{aligned}$$

因此

$$\begin{aligned} \theta f(z) &\leq \theta f(x_1) - f'(z)\theta(\theta - 1)(x_2 - x_1) \\ (1-\theta)f(z) &\leq (1-\theta)f(x_2) + f'(z)\theta(\theta - 1)(x_2 - x_1) \end{aligned}$$

最后得出

$$f(z) \leq \theta f(x_1) + (1-\theta)f(x_2)$$

从而证得函数是凸函数。另一方面, 假设 $f(x)$ 是凸函数, x_1 是定义域任意一点, x_2 是 x_1 邻域内的任意一点。根据凸函数定义, 有

$$f(\theta x_1 + (1-\theta)x_2) \leq \theta f(x_1) + (1-\theta)f(x_2)$$

变换一下

$$\frac{f(x_2 + \theta(x_1 - x_2)) - f(x_2)}{\theta} \leq f(x_1) - f(x_2)$$

取 $\theta \rightarrow 0$ 的极限, 根据符合函数求导法则, 写出

$$f'(x_2)(x_1 - x_2) \leq f(x_1) - f(x_2)$$

现在可以说对任意 x_1 以及其领域内的任意 x_2 都有

$$f(x_1) \geq f(x_2) + f'(x_2)(x_1 - x_2)$$

再一次写出泰勒展开

$$f(x_1) = f(x_2) + f'(x_2)h + h^T f''(x_2)h + o(\|h\|^2)$$

得知 $h^T f''(x_2)h + o(\|h\|^2) \geq 0$, 让 h 充分小以至于 $o(\|h\|^2) \leq \inf_x (h^T f''(x)h)$, 可得 $h^T f''(x_2)h \geq 0$.

接着, 我们证明 $\frac{\partial^2}{\partial \xi_i \partial \xi_j} f^*(\xi)$ 是半正定矩阵, 复述一遍变换的定义 $f^*(\xi) = x^i \xi_i - f(x)$, 现在认为 x 是与 ξ 有关的函数。实际上确实如此, 因为 $\xi = f'(x)$, 所以 $x = (f')^{-1}(\xi)$ 。

$$\partial_k f^*(\xi) = \sum_{i=1}^m \xi_i \partial_k x^i(\xi) + x^i \delta_{ik} - \partial_i f(x(\xi)) \partial_k x^i(\xi)$$

注意到 $\partial_i f(x(\xi)) = \xi_i$, 所以消去后有 $\partial_k f^* = x^k$, 即 $(f^*)'(\xi) = x = (f')^{-1}(\xi)$ 。现在要再对 ξ 求导, 根据反函数定理

$$\begin{cases} \xi = f'(x) \\ f'(x) \in C^{(1)}(U; \mathbb{R}^m) \\ f''(x) \text{可逆} \end{cases}$$

存在 $(f')^{-1}(\xi) = x$ 且 $[(f')^{-1}]'(\xi) = [f''(x)]^{-1}$ 。故 $(f^*)''(\xi) = [f''(x)]^{-1}$ 。已经知道了 f 是凸函数, $f''(x)$ 就是半正定矩阵。根据线性代数的知识可以证明 $[f''(x)]^{-1}$ 也是半正定矩阵。

下面给出一种更加简单的证法。当 f 是凸函数时, 题目中的变换表达和下面的表达是等价的

$$f^*(\xi) = \sup_x \{ \langle \xi, x \rangle - f(x) \}$$

此时在 $\langle \xi, x \rangle - f(x)$ 中, 我们认为 ξ 作为一个固定的参数是与 x 无关的。但是在 $\sup_x \{ \langle \xi, x \rangle - f(x) \}$ 中,

他们是相关的, 因为在 $\langle \xi, x \rangle - f(x)$ 取到极大值时需满足充要条件: $\xi - \frac{\partial f}{\partial x}(x) = 0$ 和 $-\frac{\partial^2 f}{\partial x^i \partial x^j}$ 半负定。后者显而易见, 前者就表明了 $\sup_x \{ \langle \xi, x \rangle - f(x) \} = \langle \xi, x \rangle - f(x)|_{x=(f')^{-1}(\xi)}$ 。

在这样的定义下, 直接写出

$$\begin{aligned} \theta f^*(\xi_1) &= \theta \sup_x \{ \langle \xi_1, x \rangle - f(x) \} \\ (1 - \theta) f^*(\xi_2) &= (1 - \theta) \sup_x \{ \langle \xi_2, x \rangle - f(x) \} \end{aligned}$$

因此

$$\begin{aligned} \theta f^*(\xi_1) + (1 - \theta) f^*(\xi_2) &= \theta \sup_x \{ \langle \xi_1, x \rangle - f(x) \} + (1 - \theta) \sup_x \{ \langle \xi_2, x \rangle - f(x) \} \\ &\geq \sup_x \{ \theta \langle \xi_1, x \rangle - \theta f(x) + (1 - \theta) \langle \xi_2, x \rangle - (1 - \theta) f(x) \} \\ &= f^*(\theta \xi_1 + (1 - \theta) \xi_2) \end{aligned}$$

d): 上一问已经证明了 $\partial_k f^*(\xi) = x^k(\xi)$, 两端乘上坐标微分得 $\partial_k f^*(\xi) d\xi_k = x^k(\xi) d\xi_k$ 。这样考虑是因为“任何微分都可以写为坐标微分的线性组合”。进一步

$$df^*(\xi) = \sum_{i=1}^m \partial_i f^*(\xi) d\xi_i = \sum_{i=1}^m x^i(\xi) d\xi_i$$

于是

$$\begin{aligned} (f^*)^*(x) &= \sum_{i=1}^m \frac{\partial f^*}{\partial \xi_i}(\xi) \xi_i - f^*(\xi) \\ &= \sum_{i=1}^m x^i \xi_i - \left(\sum_{i=1}^m x^i \xi_i - f(x) \right) \\ &= f(x) \end{aligned}$$

e): 显而易见, 就不写了

$$\boxed{f^*(\xi) + f(x) = \xi^T x, \quad \xi = \nabla f(x), \quad x = \nabla f^*(\xi)}$$

f): 已经证得 $(f^*)''(\xi) = [f''(x)]^{-1}$, 故 $d^*d = 1$ 。为了简单分别用 R, S 表示 f 和 f^* 的黑塞矩阵。将 S 写为列的形式

$$R \begin{bmatrix} S_1 & \cdots & S_j & \cdots & S_m \end{bmatrix} = I$$

利用矩阵乘法, 有 $RS_j = e_j$, 利用克拉默法则, $S_j^i = \frac{|R_i(e_j)|}{|R|}$ 其中 $R_i(e_j)$ 表示 R 的第 i 列替换为 e_j 。则 $|R_i(e_j)| = (-1)^{(i+j)}C_i^j = d_{ji}$, C_j^i 表示 j 行 i 列的余子式。 R 是对称矩阵, 所以 R^T 去掉 i 行 j 列的余子式的行列式值和 R 是相等的, 故 $(-1)^{(i+j)}C_i^j = (-1)^{(i+j)}C_j^i = d_{ij}$, 由此得出

$$\frac{\partial^2 f^*}{\partial \xi_i \partial \xi_j} = \frac{d_{ij}}{d}$$

g):

$$(1 + f_y'^2)f_{xx}'' - 2f_x'f_y'f_{xy}'' + (1 + f_x'^2)f_{yy}'' = 0$$

做变换

$$\begin{cases} \xi = \frac{\partial f}{\partial x}(x, y) \\ \eta = \frac{\partial f}{\partial y}(x, y) \\ f^*(\xi, \eta) = \frac{\partial f}{\partial x}(x, y)\xi + \frac{\partial f}{\partial y}(x, y)\eta - f(x, y) \end{cases}$$

根据上一问的证明, $\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{d_{ij}^*}{d^*}$, 列出变换 $f_{xx}'' = \frac{d_{11}^*}{d^*} = \frac{1}{d^*}f_{\eta\eta}^{*''}(\xi, \eta)$, $f_{xy}'' = \frac{d_{12}^*}{d^*} = \frac{1}{d^*}f_{\xi\eta}^{*''}(\xi, \eta)$, $f_{yy}'' = \frac{d_{22}^*}{d^*} = \frac{1}{d^*}f_{\xi\xi}^{*''}(\xi, \eta)$, 最终得

$$(1 + \eta^2)f_{\eta\eta}^{*''} - 2\xi\eta f_{\xi\eta}^{*''} + (1 + \xi^2)f_{\xi\xi}^{*''} = 0$$

□