ĐẠO HÀM VÀ CỰC TRỊ HÀM SỐ NHIỀU BIẾN

Nguyễn Minh Huy

HCMUS-Khoa Toán Tin

Ngày 31 tháng 7 năm 2019

Nội dung

1. Đạo hàm của hàm số nhiều biến
   - Hàm số nhiều biến
   - Đạo hàm riêng
   - Đạo hàm theo hướng
   - Gradient
   - Đạo hàm riêng của hàm số hợp
   - Đạo hàm riêng cấp cao
2. Cực trị hàm số nhiều biến
   - Cực trị
   - Cực trị có điều kiện

Định nghĩa 1.

Hàm số nhiều biến f *gán vào mỗi* $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ *một điểm* $y = (y_1, \ldots, y_m) \in \mathbb{R}^m$. *Mỗi thành phần của* y *là một hàm trả về giá trị thực với* $x \in \mathbb{R}^n$, *được viết dưới dạng*

$$\begin{aligned} y_1 &= f_1(x_1, \ldots, x_n) \\ y_2 &= f_2(x_1, \ldots, x_n) \\ &\vdots \\ y_m &= f_m(x_1, \ldots, x_n). \end{aligned}$$

Ví dụ 1.

$$f\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} f_1(x,y,z) \\ f_2(x,y,z) \end{pmatrix} = \begin{pmatrix} x-y+z \\ xyz \end{pmatrix}$$

xác định một hàm số từ $\mathbb{R}^3$ *vào* $\mathbb{R}^2$.

Định nghĩa 2.

Cho $f : D \subset \mathbb{R}^n \to \mathbb{R}$ *và* $x^0 = \left(x_1^0, x_2^0, \ldots, x_n^0\right) \in D$. *Giới hạn*

$$\lim_{h \to 0} \frac{f\left(x_1^0, \ldots, x_i^0 + h, x_{i+1}^0, \ldots, x_n^0\right) - f\left(x_1^0, \ldots, x_i^0, x_{i+1}^0, \ldots, x_n^0\right)}{h}$$

nếu có, được gọi là đạo hàm riêng theo biến thứ i *của* f *tại* x*, ký hiệu là* $\dfrac{\partial f}{\partial x_i}\left(x^0\right)$ *hay* $f_{x_i}\left(x^0\right)$.

Đạo hàm riêng

Ví dụ 2.

Cho $f(x,y) = x^2y + y^3 - 1$. *Muốn tính* $\dfrac{\partial f}{\partial x}$, *ta xem* y *như hằng số và biến* x, *ta có* $\dfrac{\partial f}{\partial x} = 2xy$. *Tương tự, ta có* $\dfrac{\partial f}{\partial y} = x^2 + 3y^2$.

Ví dụ 3.

Tìm các đạo hàm riêng của hàm số $f(x, y, z) = x^y$.

Ví dụ 4.

Cho $f : \mathbb{R}^2 \to \mathbb{R}$ *xác định bởi*

$$f(x, y) = \begin{cases} xy\dfrac{x^2 - y^2}{x^2 + y^2} & , (x, y) \neq (0, 0) \\ 0 & , (x, y) = (0, 0). \end{cases}$$

Chứng minh rằng $\dfrac{\partial f}{\partial y}(x_0, 0) = x_0$ *và* $\dfrac{\partial f}{\partial x}(0, y_0) = -y_0$.

Định nghĩa 3.

Xét hàm $f : D \subset \mathbb{R}^n \to \mathbb{R}$ *và vectơ* $a = (a_1, a_2, \ldots, a_n) \in \mathbb{R}^n$ *khác vectơ không. Giới hạn*

$$\lim_{t\to 0} \frac{f(x_1 + ta_1, x_2 + ta_2 \ldots, x_n + ta_n) - f(x_1, x_2, \ldots, x_n)}{t}$$

nếu có, được gọi là đạo hàm theo hướng a *của* f *tại* x*, ký hiệu* $D_a f(x)$*.*

Nhận xét 1. Như vậy, đạo hàm theo biến thứ i của f tại x chính là đạo hàm theo hướng $\overrightarrow{e_i}$, trong đó $\overrightarrow{e_i}$ là vectơ mà mọi thành phần đều bằng 0 trừ thành phần thứ i bằng 1.

Nhận xét 2. Đặt $f_a(t) = f(x + ta)$, ta có $D_a f(x) = f_a'(0)$. (nếu có)

Ví dụ 5.

Cho $f(x,y) = x^3y + y$ *và vectơ* $a = (1,-2)$. *Tính đạo hàm của* f *theo hướng* a *tại* $(0,1)$.

Chứng minh.

Đặt $x = (0,1)$ và

$$f_a(t) = f(x + ta) = f(t, 1-2t) = t^3(1-2t) + 1 - 2t.$$

Ta có

$$f_a'(t) = -8t^3 + 3t^2 - 2.$$

Suy ra

$$D_af(x) = f_a'(0) = -2.$$

□

Định nghĩa 4.

Cho D *là một miền và* $f : D \subset \mathbb{R}^n \to \mathbb{R}$*. Khi* f *có đạo hàm riêng theo tất cả các biến tại* $x \in D$*, gradient của* f *tại* x*, ký hiệu grad* $f(x)$ *(hay vắn tắt* $\nabla f(x)$*) là vectơ*

$$\nabla f(x) = \left(\frac{\partial f}{\partial x_1}(x), \frac{\partial f}{\partial x_2}(x), \ldots, \frac{\partial f}{\partial x_n}(x) \right).$$

Ví dụ 6.

Xét hàm $f : \mathbb{R}^2 \to \mathbb{R}$ *xác định bởi*

$$f(x, y) = x^2 y.$$

Tính $\nabla f(1, 2)$.

Chứng minh.

Ta có $\dfrac{\partial f}{\partial x}(x, y) = 2xy$ nên $\dfrac{\partial f}{\partial x}(1, 2) = 2.1.2 = 4$.

Tương tự, $\dfrac{\partial f}{\partial y}(x, y) = x^2$ nên $\dfrac{\partial f}{\partial y}(x, y) = 1^2 = 1$.

Vậy $\nabla f(1, 2) = \left(\dfrac{\partial f}{\partial x}(1, 2), \dfrac{\partial f}{\partial y}(1, 2)\right) = (4, 1)$. $\square$

Mệnh đề 1.

Cho $f, g : D \subset \mathbb{R}^n \to \mathbb{R}$*. Nếu* f, g *có đạo hàm riêng theo mọi biến tại* $x \in D$ *thì*

i) $\nabla (f + g)(x) = \nabla f(x) + \nabla g(x)$;

ii) $\nabla (f.g)(x) = g(x) . \nabla f(x) + f(x) . \nabla g(x)$.

Hơn nữa, nếu $g(x) \neq 0$ *và hàm* $\dfrac{f}{g}$ *xác định trên một lân cận của* x *thì nó có đạo hàm riêng theo mọi biến tại* x *và*

iii) $\nabla \left(\dfrac{f}{g}\right)(x) = \dfrac{1}{g^2(x)} \left(g(x) . \nabla f(x) - f(x) . \nabla g(x)\right)$.

Mệnh đề 2.

Cho $f : D \subset \mathbb{R}^n \to \mathbb{R}$ *là hàm có đạo hàm riêng theo các biến tại mọi điểm của* D *và tất cả các đạo hàm riêng này đều liên tục tại điểm* $x \in D$. *Khi đó, với mọi vectơ* $a = (a_1, a_2, \ldots, a_n)$ *khác vectơ không trong* $\mathbb{R}^n$, f *có đạo hàm theo hướng* a *tại* x *và*

$$D_a f(x) = a_1.\frac{\partial f}{\partial x_1}(x) + \ldots + a_n.\frac{\partial f}{\partial x_n}(x) = a.\nabla f(x).$$

Từ công thức trên, ta có

$$D_a f(x) = a.\nabla f(x) \leq \left(\nabla f(x)\right)^2,$$

Điều này có nghĩa là theo hướng của vectơ $\nabla f(x)$ thì f sẽ thay đổi với tốc độ nhanh nhất.

Định lý 1.

(Quy tắc mắc xích) *Cho hàm số khả vi* $f : \mathbb{R}^n \to \mathbb{R}$ *và các hàm số khả vi* $x_1, x_2, \ldots, x_n : \mathbb{R}^m \to \mathbb{R}$ *đều là hàm theo các biến* $t_1, t_2, \ldots, t_m$*. Khi đó,*

$$\frac{\partial f}{\partial t_i} = \frac{\partial f}{\partial x_1}.\frac{\partial x_1}{\partial t_i} + \frac{\partial f}{\partial x_2}.\frac{\partial x_2}{\partial t_i} + \ldots + \frac{\partial f}{\partial x_n}.\frac{\partial x_n}{\partial t_i}, \quad \forall i \in \{1, 2, \ldots, m\}.$$

Ví dụ 7.

Cho $f(x, y, z) = x^2 + yz$ *với*

$$\begin{aligned} x(s,t) &= t + s, \\ y(s,t) &= 3t - s, \\ z(s,t) &= t^2. \end{aligned}$$

Khi đó,

$$\frac{\partial f}{\partial s} = 2x.1 + z.(-1) + y.0 = 2(t+s) - t^2,$$

$$\frac{\partial f}{\partial t} = 2x.1 + z.3 + y.2t = 2(t+s) + 3t^2 + 2t(3t-s).$$

Ví dụ 8.

Tìm đạo hàm riêng của f theo s và t với

$$\begin{aligned} f(x,y) &= x^3 \sin(xy), \\ x(s,t) &= t\cos s, \\ y(s,t) &= t\sin s. \end{aligned}$$

Đạo hàm riêng cấp cao

Định nghĩa 5.

Cho $f : D \subset \mathbb{R}^n \to \mathbb{R}$ *là hàm có đạo hàm riêng theo tất cả các biến tại mọi điểm của* D. *Khi đó, đạo hàm riêng theo các biến của các hàm*

$$\begin{array}{cccc} \dfrac{\partial f}{\partial x_i} & : & D & \longrightarrow & \mathbb{R} \\ & & x & \longmapsto & \dfrac{\partial f}{\partial x_i}(x) \end{array}$$

với $i = \overline{1, n}$ *tại điểm* $x \in D$, *nếu có, được gọi là các* **đạo hàm riêng cấp 2** *của* f *tại điểm* x, *ký hiệu là* $\dfrac{\partial^2 f}{\partial x_j \partial x_i}(x)$ *hoặc* $f_{x_i x_j}(x)$ *với* $i, j = \overline{1, n}$.

Ví dụ 9.

Cho $f(x,y) = x^2y + x$. *Tính các đạo hàm riêng cấp hai của* f.

Định lý 2.

Nếu $f : D \subset \mathbb{R}^n \to \mathbb{R}$ *có các đạo hàm riêng cấp 2 liên tục trên* D *thì*

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

trên D, *với mọi* $i, j = \overline{1, n}$.

Định nghĩa 6.

Cho f là hàm nhiều biến xác định trên một miền $D \subset \mathbb{R}^n$ và $x^0 \in D$.

i) *f được gọi là đạt cực đại địa phương tại x^0 nếu tồn tại lân cận Ω chứa trong D của x^0 sao cho $f(x) \leq f(x^0)$ với mọi $x \in \Omega$.*

ii) *f được gọi là đạt cực tiểu địa phương tại x^0 nếu tồn tại lân cận Ω chứa trong D của x^0 sao cho $f(x) \geq f(x^0)$ với mọi $x \in \Omega$.*

iii) *f được gọi là đạt cực trị địa phương tại x^0 khi nó đạt cực đại địa phương hoặc cực tiểu địa phương tại x^0.*

Chú ý rằng khi $x = (x_1, \ldots, x_n), x^0 = \left(x_1^0, \ldots, x_n^0\right) \in D \subset \mathbb{R}^n$ thì với mỗi $i = \overline{1, n}$, hàm $t \to f\left(x_1^0, \ldots, x_{i-1}^0, t, x_{i+1}^0, \ldots, x_n^0\right)$ là hàm một biến xác định trên $D_i = \left\{t \in \mathbb{R} : \left(x_1^0, \ldots, t, \ldots, x_n^0\right) \in D\right\}$ và đạt cực trị địa phương tại $t = x_i^0$ khi f đạt cực trị địa phương tại x^0. Do đó, ta có

Định lý 3.

Cho $f : D \subset \mathbb{R}^n \to \mathbb{R}$ *là hàm có đạo hàm riêng theo các biến trên* D.
Điều kiện cần để f *đạt cực trị địa phương tại* $x^0 \in D$ *là*

$$\frac{\partial f}{\partial x_i}\left(x^0\right) = 0, \forall i = \overline{1, n}.$$

Đẳng thức trên có thể viết lại dưới dạng vectơ là

$$\nabla f\left(x^0\right) = 0$$

và điểm x^0 *được gọi là một điểm dừng của* f.

Ví dụ 10.

Tìm điểm dừng của $f(x, y) = x^2 - 12y^2 + 4y^3 + 3y^4$.

Chứng minh.

Ta có $\nabla f(x, y) = \Big(2x, 12\left(y^3 + y^2 - 2y\right)\Big)$.

Suy ra $\nabla f(x, y) = 0 \Leftrightarrow \begin{cases} x = 0 \\ y = 0 \vee y = 1 \vee y = -2 \end{cases}$

Vậy $(0, 0), (0, 1)$ và $(0, -2)$ là các điểm dừng của f. $\square$

Vậy có phải tất cả các điểm dừng trên đều là các cực trị địa phương của f không?

Chú ý rằng ngay cả khi f là hàm một biến thì không nhất thiết f đạt cực trị địa phương tại các điểm dừng của nó. Tuy nhiên, giống như trường hợp hàm một biến, người ta có thể dùng các đạo hàm cấp 2 để xây dựng một điều kiện đủ cho các cực trị tại các điểm dừng.

Với hàm $f : D \subset \mathbb{R}^n \to \mathbb{R}$ thuộc lớp C^2 và điểm dừng $x \in D$. Đặt

$$\varphi(h,h) = \sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(x)\, h_i h_j$$

$$= \begin{pmatrix} h_1 & \dots & h_n \end{pmatrix} \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(x) & \frac{\partial^2 f}{\partial x_1 \partial x_2}(x) & \dots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(x) \\ \frac{\partial^2 f}{\partial x_2 \partial x_1}(x) & \frac{\partial^2 f}{\partial x_2^2}(x) & \dots & \frac{\partial^2 f}{\partial x_2 \partial x_n}(x) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(x) & \frac{\partial^2 f}{\partial x_n \partial x_2}(x) & \dots & \frac{\partial^2 f}{\partial x_n^2}(x) \end{pmatrix} \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix}$$

$$\equiv h^T . H_f(x) . h,$$

trong đó $H_f(x)$ được gọi là *ma trận Hessian* của f tại x.

Định lý 4.

i) *Nếu* $\varphi(h,h) > 0, \forall h \neq 0$ *thì* f *đạt cực tiểu địa phương tại* x.

ii) *Nếu* $\varphi(h,h) < 0, \forall h \neq 0$ *thì* f *đạt cực đại địa phương tại* x.

iii) *Nếu tồn tại* $h, k \in \mathbb{R}^n$ *sao cho* $\varphi(h,h) > 0$ *và* $\varphi(k,k) < 0$ *thì* x *không là cực trị địa phương của* f.

Ví dụ 11.

Tìm cực trị của $f(x,y) = x^2 - 12y^2 + 4y^3 + 3y^4$.

Chứng minh.

Ta có $f_{xx} = 2, f_{xy} = f_{yx} = 0, f_{yy} = 12\left(y^3 + y^2 - 2y\right)$ nên với $h = (h_1, h_2)$ thì

$$\varphi(h,h) = 2h_1^2 + 12\left(y^3 + y^2 - 2y\right)h_2^2.$$

Trong ví dụ trước, ta đã tìm được các điểm dừng của f là $(0,0), (0,1)$ và $(0,-2)$. $\square$

Chứng minh.

*Tại $x = (0, 1)$, ta có

$$\varphi(h, h) = 2h_1^2 + 36h_2^2 > 0, \forall h \neq (0, 0).$$

Vậy f đạt cực tiểu tại $x = (0, 1)$.
*Tại $x = (0, -2)$, ta có

$$\varphi(h, h) = 2h_1^2 + 72h_2^2 > 0, \forall h \neq (0, 0).$$

Vậy f đạt cực tiểu tại $x = (0, -2)$.
*Tại $x = (0, 0)$, ta có

$$\varphi(h, h) = 2h_1^2 - 24h_2^2.$$

Khi đó, $\varphi(1, 0) > 0$ nhưng $\varphi(0, 1) < 0$ nên f không đạt cực trị tại $x = (0, 0)$. $\square$

Ví dụ 12.

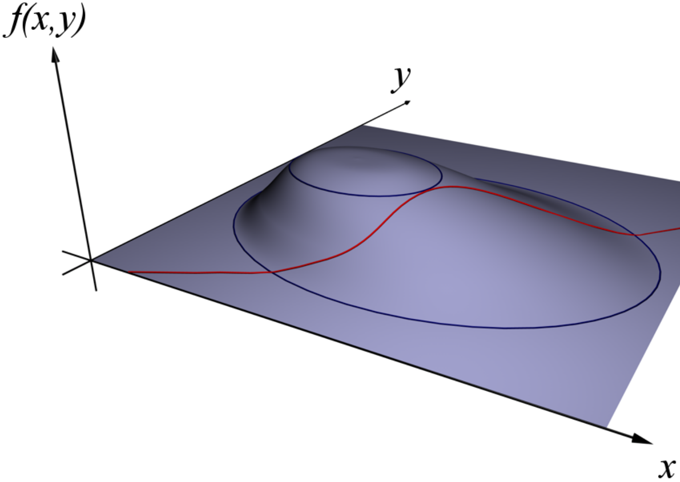
Khảo sát cực trị của hàm số $f(x,y) = x^2 - 2xy + \dfrac{y^3}{3} - 3y$.

Khi D không là một mở của $\mathbb{R}^n$, chẳng hạn khi D là biên của một miền trong $\mathbb{R}^n$, do không xét được tính khả vi của f trên D nên ta không thể dùng **Định lý 2**. Tuy nhiên nếu ta có phương trình xác định D, chẳng hạn

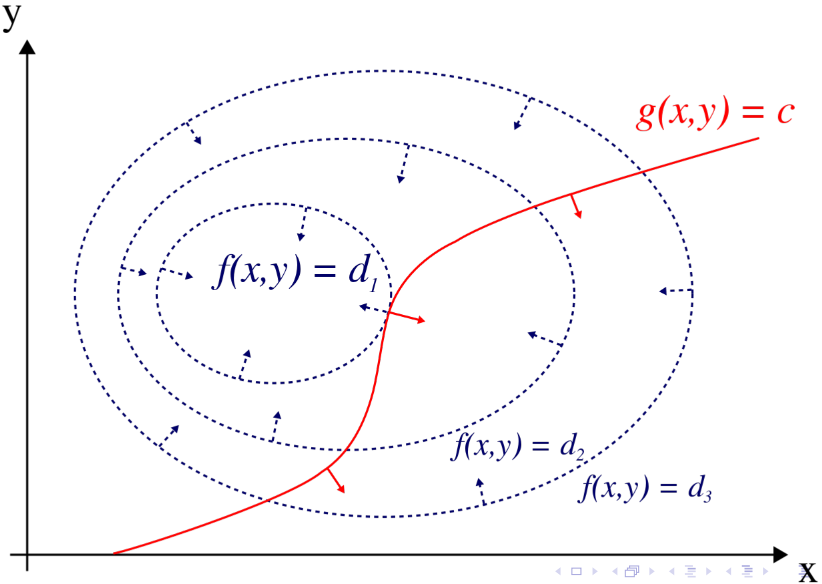
$$D = \left\{(x_1, \ldots, x_n) \in \mathbb{R}^n : \Phi\left(x_1, \ldots, x_n\right) = 0\right\},$$

với hàm khả vi $\Phi : \mathbb{R}^n \to \mathbb{R}$ và f là hàm xác định trên một mở chứa D. Khi đó, cực trị của f trên D còn được gọi là **cực trị có điều kiện**.

Cực trị có điều kiện



Cực trị có điều kiện



Định lý 5.

Cho $f : E \subset \mathbb{R}^n \to \mathbb{R}$ *là hàm thuộc lớp* C^1, *nghĩa là* f *có đạo hàm riêng theo các biến tại mọi điểm của* E *và tất cả các đạo hàm riêng này đều liên tục trên* E. *Xét*

$$D = \left\{(x_1, \ldots, x_n) \in E : g(x_1, \ldots, x_n) = 0\right\}$$

với $g : E \subset \mathbb{R}^n \to \mathbb{R}$ *là hàm thuộc lớp* C^1. *Nếu* f *đạt cực trị trên* D, *tại* $x^0 \in D$ *và* $\nabla g\left(x^0\right) \neq 0$ *thì tồn tại* $\lambda \in \mathbb{R}$ *sao cho*

$$\nabla f\left(x^0\right) + \lambda \nabla g\left(x^0\right) = 0.$$

Hơn nữa, việc tìm cực trị của f *trên* D *chính là tìm cực trị của* $f + \lambda g$.

Ví dụ 13.

Tìm cực đại, cực tiểu của $f(x, y, z) = x + y - z$ *với điều kiện* $x^2 + y^2 + z^2 = 1.$

Chứng minh.

Đặt $g(x, y, z) = x^2 + y^2 + z^2 - 1$ và $\phi = f + \lambda g$. Ta giải hệ phương trình sau

$$\begin{cases} \nabla f + \lambda \nabla g = 0 \\ g(x, y, z) = 0 \end{cases} \Leftrightarrow \begin{cases} (1, 1, -1) + \lambda(2x, 2y, 2z) = 0 \\ x^2 + y^2 + z^2 = 1 \end{cases}$$

$$\Leftrightarrow \begin{cases} \lambda = \dfrac{\sqrt{3}}{2} \\ x = y = -\dfrac{1}{\sqrt{3}} \\ z = \dfrac{1}{\sqrt{3}} \end{cases} \textbf{(1)} \vee \begin{cases} \lambda = -\dfrac{\sqrt{3}}{2} \\ x = y = \dfrac{1}{\sqrt{3}} \\ z = -\dfrac{1}{\sqrt{3}} \end{cases} \textbf{(2)}$$

Chứng minh.

Ta tính được $\varphi(h,h) = h^T.H_\phi(x,y,z).h = 2\lambda\left(h_1^2+h_2^2+h_3^2\right)$.

*Với **(1)**, ta có $\varphi(h,h) > 0, \forall h \neq 0$ nên ϕ đạt cực tiểu tại $\left(-\frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right)$ hay f đạt cực tiểu trên D tại $\left(-\frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right)$ với $D = \left\{(x,y,z) \in \mathbb{R}^3 : g(x,y,z) = 0\right\}$.

*Với **(2)**, ta có $\varphi(h,h) < 0, \forall h \neq 0$ nên ϕ đạt cực đại tại $\left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{3}}\right)$ hay f đạt cực đại trên D tại $\left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{3}}\right)$. □

Ví dụ 14.

Tìm cực trị của $f(x, y, z) = x^2 + xy + y^2 + yz + z^2$ *với điều kiện* $x^2 + y^2 + z^2 = 1$.

Định lý 6.

Cho $f : E \subset \mathbb{R}^n \to \mathbb{R}$ là hàm thuộc lớp C^1, nghĩa là f có đạo hàm riêng theo các biến tại mọi điểm của E và tất cả các đạo hàm riêng này đều liên tục trên E. Xét

$$D = \left\{(x_1, \ldots, x_n) \in E : g_i\left(x_1, \ldots, x_n\right) = 0, \forall i = \overline{1, m}\right\}$$

với $g : E \subset \mathbb{R}^n \to \mathbb{R}$ là hàm thuộc lớp C^1. Nếu f đạt cực trị trên D, tại $x^0 \in D$ và $\nabla g_i\left(x^0\right) \neq 0, \forall i = \overline{1, m}$ thì tồn tại $\lambda_i \in \mathbb{R}$ sao cho

$$\nabla f\left(x^0\right) + \sum_{i=1}^{m} \lambda_i \nabla g_i\left(x^0\right) = 0.$$

Hơn nữa, việc tìm cực trị của f trên D chính là tìm cực trị của $f + \sum\limits_{i=1}^{m} \lambda_i g_i$.

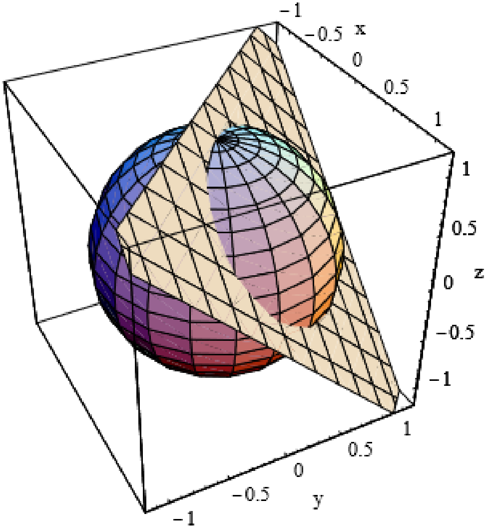
Ví dụ 15.

Giả sử nhiệt độ trong một quả cầu đơn vị $x^2 + y^2 + z^2 = 1$ *được cho bởi hàm số sau*

$$T(x, y, z) = 80 + 50(x + z).$$

Tìm cực trị của T *trên phần giao của quả cầu đơn vị và mặt phẳng* $x + y + z = 1$.

Cực trị có điều kiện



Chứng minh.

Đặt $g_1(x,y,z) = x^2 + y^2 + z^2 - 1$, $g_2(x,y,z) = x + y + z - 1$ và $\phi = f + \lambda_1 g_1 + \lambda_2 g_2$. Ta giải hệ phương trình sau

$$\begin{cases} \nabla f + \lambda_1 \nabla g_1 + \lambda_2 \nabla g_2 = 0 \\ g_1(x,y,z) = 0 \\ g_2(x,y,z) = 0 \end{cases}$$

$\square$