CLASSIFICAÇÃO DOS PONTOS CRÍTICOS DE UMA FUNÇÃO DE $\mathbb{R}^2$ EM $\mathbb{R}$

MAT-2454 – CÁLCULO DIFERENCIAL E INTEGRAL II – 2015

RESUMO. Estas notas têm por objetivo fornecer uma demonstração para o critério de classificação de pontos críticos de funções de classe $\mathcal{C}^2$, $f : A \to \mathbb{R}$, onde A é um conjunto aberto de $\mathbb{R}^2$. A ideia é estabelecer uma conexão entre este resultado com os conteúdos vistos em MAT-2458 – Álgebra Linear II.

1. INTRODUÇÃO

Recordemos que se $f : A \to \mathbb{R}$ é de classe $\mathcal{C}^2$, com A aberto de $\mathbb{R}^2$, então para cada $(x_0, y_0) \in A$ podemos escrever

$$f(x,y) = P_1(x,y) + E(x,y), \tag{1.1}$$

onde $P_1(x,y)$ é o polinômio de Taylor de f em torno de (x_0,y_0) e $E(x,y)$ é o erro cometido na aproximação de $f(x,y)$ por $P_1(x,y)$. Sabemos que $P_1(x,y)$ e $E(x,y)$ são dados por

$$P_1(x,y) = f(x_0,y_0) + f_x(x_0,y_0)(x-x_0) + f_y(x_0,y_0)(y-y_0) \tag{1.2}$$

$$E(x,y) = \frac{1}{2}\left(f_{xx}(\overline{x},\overline{y})(x-x_0)^2 + 2f_{xy}(\overline{x},\overline{y})(x-x_0)(y-y_0) + f_{yy}(\overline{x},\overline{y})(y-y_0)^2\right), \tag{1.3}$$

onde $(\overline{x},\overline{y})$ está no interior do segmento ligando (x_0,y_0) a (x,y).

Se (x_0,y_0) é um ponto crítico de f, então $\nabla f(x_0,y_0) = (0,0)$. Usando este fato e substituindo a expressão (1.2) em (1.1) temos que

$$f(x,y) = f(x_0,y_0) + E(x,y). \tag{1.4}$$

Com isto observamos que se existe uma bola $B_\epsilon(x_0,y_0)$, de raio $\epsilon > 0$ e centrada em (x_0,y_0), tal que $E(x,y) \geq 0$ para todo $(x,y) \in B_\epsilon(x_0,y_0)$ então $f(x,y) \geq f(x_0,y_0)$, ou seja, (x_0,y_0) é *ponto de mínimo local* de f. De modo análogo, se $E(x,y) \leq 0$ para todo $(x,y) \in B_\epsilon(x_0,y_0)$ então $f(x,y) \leq f(x_0,y_0)$, e portanto (x_0,y_0) é *ponto de máximo local* de f.

Isto nos mostra que o sinal de $E(x,y)$ é quem carateriza a natureza do ponto crítico (x_0,y_0).

Além disso, podemos escrever a expressão (1.2) na forma matricial da seguinte maneira

$$E(x,y) = \frac{1}{2}\begin{bmatrix} x-x_0 & y-y_0 \end{bmatrix}\begin{bmatrix} f_{xx}(\overline{x},\overline{y}) & f_{xy}(\overline{x},\overline{y}) \\ f_{xy}(\overline{x},\overline{y}) & f_{yy}(\overline{x},\overline{y}) \end{bmatrix}\begin{bmatrix} x-x_0 \\ y-y_0 \end{bmatrix}. \tag{1.5}$$

Observe que a matriz quadrada na expressão (1.5) acima é simétrica e portanto diagonalizável, com uma base ortonormal de autovetores.

Na seção seguinte estabelecemos alguns resultados da Álgebra Linear utilizando este fato.

Preparado e redigido pelo prof. Alexandre Lymberopoulos – MAT – IME-USP.

2. Um Pouco de Álgebra Linear

Lema 2.1. *Sejam* $A = \begin{bmatrix} a & b \\ b & d \end{bmatrix}$ *e* $D = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$ *matrizes semelhantes, ou seja,* $A = PDP^{-1}$, *onde* P *é a matriz (ortogonal) de mudança de base entre a base canônica e a base de autovetores de* A. *Denotando* $Q(x, y)$ *por*

$$Q(x,y) = \begin{bmatrix} x & y \end{bmatrix} A \begin{bmatrix} x \\ y \end{bmatrix} \tag{2.1}$$

temos que

(i) se $\lambda_1 \geq 0$ *e* $\lambda_2 \geq 0$ *então* $Q(x, y) \geq 0$ *para todos* $(x, y) \in \mathbb{R}^2$.
(ii) se $\lambda_1 \leq 0$ *e* $\lambda_2 \leq 0$ *então* $Q(x, y) \leq 0$ *para todos* $(x, y) \in \mathbb{R}^2$.
(iii) se $\lambda_1\lambda_2 < 0$ *então existem* (x, y) *e* (u, v) *em* $\mathbb{R}^2$ *tais que* $Q(x, y) < 0$ *e* $Q(u, v) > 0$.

Observação 2.1. $Q(x, y)$ é a *forma quadrática* associada à matriz simétrica A.

Demonstração: Denotando por $\begin{bmatrix} \tilde{x} & \tilde{y} \end{bmatrix}$ as coordenadas (na base de autovetores de A) do vetor $P \begin{bmatrix} x & y \end{bmatrix}$ temos que

$$Q(x,y) = \begin{bmatrix} x & y \end{bmatrix} A \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x & y \end{bmatrix} PDP^{-1} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \tilde{x} & \tilde{y} \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} \tilde{x} \\ \tilde{y} \end{bmatrix} = \lambda_1\tilde{x}^2 + \lambda_2\tilde{y}^2.$$

Com isso é claro que

(i) se $\lambda_1 \geq 0$ e $\lambda_2 \geq 0$ então $Q(x, y) \geq 0$ para todos $(x, y) \in \mathbb{R}^2$.
(ii) se $\lambda_1 \leq 0$ e $\lambda_2 \leq 0$ então $Q(x, y) \leq 0$ para todos $(x, y) \in \mathbb{R}^2$.
(iii) se $\lambda_1\lambda_2 < 0$, vamos supor $\lambda_1 > 0$ e $\lambda_2 < 0$. Basta tomar (x, y) como um autovetor associado a λ_1 (que terá segunda coordenada nula na base de autovetores de A) e então $Q(x, y) = \lambda_1\tilde{x}^2 > 0$. Analogamente, tomando (u, v) como um autovetor associado a λ_2 e então $Q(u, v) = \lambda_2\tilde{v}^2 \leq 0$.

□

O seguinte lema será útil para simplificar a demonstração do lema que o sucede.

Lema 2.2. *Sejam* A, B *matrizes quadradas. Então* $\mathrm{Tr}\,(AB) = \mathrm{Tr}\,(BA)$, *onde* $\mathrm{Tr}\,(M)$ *é o* traço *de* M, *ou seja, a soma dos elementos da diagonal principal de* M.

Demonstração: Denotando $A = (a_{ij})$ e $B = (b_{i,j})$, com $1 \leq i, j \leq n$ temos que os elementos c_{ii}, da diagonal principal de AB são dados por

$$c_{ii} = \sum_{j=1}^{n} a_{ij}b_{ji},$$

e os elementos d_{ii}, da diagonal principal de BA são dados por

$$d_{jj} = \sum_{i=1}^{n} b_{ji}a_{ij}.$$

Claramente $\mathrm{Tr}\,(AB) = \sum_{i=1}^{n} c_{ii} = \sum_{j=1}^{n} d_{jj} = \mathrm{Tr}\,(BA)$. □

O resultado a seguir faz uso da tradicional fórmula para o determinante do produto de matrizes, $\det(AB) = \det(A)\det(B)$. Uma demonstração para este fato pode ser encontrada em [Cal, pp.218].

Lema 2.3. *Sejam A e B matrizes semelhantes, ou seja, existe P invertível tal que $A = PBP^{-1}$. Então*

(i) $\det(A) = \det(B)$;
(ii) $\operatorname{Tr}(A) = \operatorname{Tr}(B)$.

Demonstração:

(i) $\det(A) = \det(PBP^{-1}) = \det(P)\det(B)\det(P^{-1}) = \det(P)\det(B)\det(P)^{-1} = \det(B)$.
(ii) $\operatorname{Tr}(A) = \operatorname{Tr}(PBP^{-1}) = \operatorname{Tr}(P^{-1}PB) = \operatorname{Tr}(IB) = \operatorname{Tr}(B)$.

□

Procure entender o significado geométrico da invariância do determinante e do traço mediante mudança de base.

Agora estamos em condições de demonstrar um critério para determinar o sinal de $Q(x,y)$, definido em (2.1), em termos de cálculos simples com os elementos da matriz A. Claramente $Q(0,0) = 0$.

Teorema 2.4 (Critério de Sylvester). *Sejam $A = \begin{bmatrix} a & b \\ b & d \end{bmatrix}$ e $D = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$ semelhantes e $Q(x,y)$ como em (2.1). Então*

(i) se $a > 0$ e $\det A \geq 0$ então $Q(x,y) > 0$ para todo $(x,y) \in \mathbb{R}^2$, $(x,y) \neq (0,0)$.
(ii) se $a < 0$ e $\det A \geq 0$ então $Q(x,y) < 0$ para todo $(x,y) \in \mathbb{R}^2$, $(x,y) \neq (0,0)$.
(iii) se $\det(A) < 0$ existem (x,y) e (u,v) em $\mathbb{R}^2$ tais que $Q(x,y) > 0$ e $Q(u,v) < 0$.

Demonstração: Como A e D são semelhantes, segue-se do lema 2.3 que

$$\det(A) = ad - b^2 = \lambda_1\lambda_2$$
$$\operatorname{Tr}(A) = a + d = \lambda_1 + \lambda_2.$$

(i) se $\det(A) \geq 0$ então $ad - b^2 \geq 0$, ou seja $ad \geq b^2 \geq 0$. Como $a > 0$, temos $d \geq 0$ e portando $a + d > 0$. Assim $\lambda_1\lambda_2 = \det(A) \geq 0$ e $\lambda_1 + \lambda_2 = \operatorname{Tr}(A) > 0$, o que só é possível quando $\lambda_1 \geq 0$ e $\lambda_2 \geq 0$, não se anulando simultaneamente.

Pela parte (i) do lema 2.1, temos que $Q(x,y) > 0$ para todo $(x,y) \in \mathbb{R}^2$, $(x,y) \neq (0,0)$.

(ii) se $\det(A) \geq 0$ e $a < 0$, obtemos de maneira análoga que $d \leq 0$ e portando $a + d < 0$. Assim $\lambda_1\lambda_2 = \det(A) \geq 0$ e $\lambda_1 + \lambda_2 = \operatorname{Tr}(A) < 0$, o que só é possível quando $\lambda_1 \leq 0$ e $\lambda_2 \leq 0$, não se anulando simultaneamente.

Pela parte (ii) do lema 2.1, temos que $Q(x,y) < 0$ para todo $(x,y) \in \mathbb{R}^2$, $(x,y) \neq (0,0)$.

(iii) se $\det(A) < 0$ então $\lambda_1\lambda_2 < 0$ e segue-se a afirmação da parte (iii) do lema 2.1.

□

3. Da Álgebra Linear de Volta para o Cálculo

Agora podemos usar o teorema 2.4 para determinar o sinal de $E(x,y)$, definido em (1.3), em termos de cálculos simples com as segundas derivadas de f, calculadas apenas no ponto crítico (x_0, y_0). Observe que $E(x,y)$ depende de $(\overline{x}, \overline{y})$, que não é obtido explicitamente.

Definição 3.1. Se $f : A \subset \mathbb{R}^2 \to \mathbb{R}$ é de classe $\mathcal{C}^2$ no aberto A então

$$H(x_0, y_0) = \begin{bmatrix} f_{xx}(x_0, y_0) & f_{xy}(x_0, y_0) \\ f_{xy}(x_0, y_0) & f_{yy}(x_0, y_0) \end{bmatrix}$$

é o *hessiano* de f em (x_0, y_0).

Teorema 3.1 (Classificação dos pontos críticos). *Sejam $A \subset \mathbb{R}^2$ um aberto, $f : A \to \mathbb{R}$ uma função de classe $\mathcal{C}^2$ e (x_0, y_0) um ponto crítico de f. Então,*

(i) se $f_{xx}(x_0, y_0) > 0$ e $\det H(x_0, y_0) > 0$ então (x_0, y_0) é mínimo local de f;
(ii) se $f_{xx}(x_0, y_0) < 0$ e $\det H(x_0, y_0) > 0$ então (x_0, y_0) é máximo local de f;
(iii) se $\det H(x_0, y_0) < 0$ então (x_0, y_0) é um ponto de sela de f;

Demonstração: Como f é de classe $\mathcal{C}^2$ então $f_{xx}(x, y)$ e $\det H(x, y)$ são funções contínuas e portanto se estão funções não se anulam em (x_0, y_0) existe $\epsilon > 0$ tal que o sinal de $\det H(x, y)$ é o mesmo de $\det H(x_0, y_0)$ assim como o sinal de $f_{xx}(x, y)$ é o mesmo de $f_{xx}(x_0, y_0)$, para todo $(x, y) \in B_\epsilon(x_0, y_0)$.

Se $(x, y) \in B_\epsilon(x_0, y_0)$ então o sinal de $E(x, y)$, determinado por $f_{xx}(\overline{x}, \overline{y})$ e $\det H(\overline{x}, \overline{y})$ é o mesmo da forma quadrática dada pela matriz $H(x_0, y_0)$. Em vista da expressão (1.4), segue-se o resultado. $\square$

Observação 3.1. O critério acima não prevê o caso em que $\det H(x_0, y_0) = 0$ pois o teorema da conservação do sinal não se aplica neste caso. Mas algo ainda pode ser dito neste caso se sabemos o comportamento de $\det H$ e $\operatorname{Tr} H$ numa vizinhança do ponto (x_0, y_0).

Por exemplo, se $\det H(x, y) = 0$ e $\operatorname{Tr} H(x, y) > 0$ numa vizinhança do ponto (x_0, y_0) então podemos dizer, usando o teorema 2.4 que $E(x, y) \geq 0$ e portanto (x_0, y_0) é ponto de mínimo local de f.

Exercício 1. Tente estabelecer critério semelhante para funções de classe $\mathcal{C}^2$ definidas em abertos do $\mathbb{R}^3$.

Referências

[Cal] Callioli C.A., Domingues H.H., Costa, R.C.F., *Álgebra Linear e Aplicações*, 6a. edição, Ed. Atual, 1990.