Esercitazioni del 15 aprile 2013

Polinomio di Taylor del secondo ordine per funzioni di due variabili

Sia $f : A \subseteq \mathbb{R}^2 \to \mathbb{R}$ una funzione di classe C^2. Fissato un p unto $(x_0, y_0) \in A$ consideriamo il seguente polinomio di secondo grado nelle variabili x, y:

$$\begin{aligned} P_2(x,y) &= f(x_0,y_0) + \frac{\partial f(x_0,y_0)}{\partial x}(x-x_0) + \frac{\partial f(x_0,y_0)}{\partial y}(y-y_0) + \\ &+ \frac{1}{2}\left(\frac{\partial^2 f(x_0,y_0)}{\partial x^2}(x-x_0)^2 + \frac{\partial^2 f(x_0,y_0)}{\partial y^2}(y-y_0)^2 + 2\frac{\partial^2 f(x_0,y_0)}{\partial y \partial x}(x-x_0)(y-y_0)\right) \end{aligned}$$

P_2 è detto *polinomio di Taylor del secondo ordine della funzione f nel punto* (x_0, y_0). Rappresenta la migliore funzione polinomiale di grado 2 che approssima la funzione f in un intorno di (x_0, y_0), infatti si ha che

$$f(x_0,y_0) = P_2(x_0,y_0), \quad \frac{\partial f(x_0,y_0)}{\partial x} = \frac{\partial P_2(x_0,y_0)}{\partial x}, \quad \frac{\partial f(x_0,y_0)}{\partial y} = \frac{\partial P_2(x_0,y_0)}{\partial y},$$

$$\frac{\partial^2 f(x_0,y_0)}{\partial x^2} = \frac{\partial^2 P_2(x_0,y_0)}{\partial x^2}, \quad \frac{\partial^2 f(x_0,y_0)}{\partial y^2} = \frac{\partial^2 P_2(x_0,y_0)}{\partial y^2}$$
$$\frac{\partial^2 f(x_0,y_0)}{\partial x \partial y} = \frac{\partial^2 P_2(x_0,y_0)}{\partial x \partial y} \quad \frac{\partial^2 f(x_0,y_0)}{\partial y \partial x} = \frac{\partial^2 P_2(x_0,y_0)}{\partial y \partial x}$$

Inoltre si ha che

$$f(x,y) = P_2(x,y) + o(\|(x-x_0, y-y_0\|^2)$$

cioé

$$\lim_{(x,y)\to(x_0,y_0)} \frac{f(x,y) - P_2(x,y)}{\|(x-x_0, y-y_0\|^2} = 0$$

L'ultima uguaglianza significa che quando (x, y) tende a (x_0, y_0) lo scarto $f(x, y) - P_2(x, y)$ tra il valore $f(x, y)$ e il valore $P_2(x, y)$ tende a 0 più velocemente di $\|(x - x_0, y - y_0\|^2$, la distanza al quadrato tra (x, y) e (x_0, y_0). Introdotto il vettore gradiente $\nabla f(x_0, y_0)$

$$\nabla f(x_0,y_0) = \left(\frac{\partial f(x_0,y_0)}{\partial x}, \frac{\partial f(x_0,y_0)}{\partial y}\right)$$

e la matrice 2×2 contenente le derivate parziali seconde di f in (x_0, y_0), detta matrice *Hessiana*

$$H_f(x_0,y_0) = \begin{pmatrix} \frac{\partial^2 f(x_0,y_0)}{\partial x^2} & \frac{\partial^2 f(x_0,y_0)}{\partial y \partial x} \\ \frac{\partial^2 f(x_0,y_0)}{\partial y \partial x} & \frac{\partial^2 f(x_0,y_0)}{\partial y^2} \end{pmatrix}$$

allora possiamo scrivere il polinomio P_2 nel modo seguente:

$$P_2(x,y) = f(x_0,y_0)+\nabla f(x_0,y_0)\cdot(x-x_0,y-y_0)+\frac{1}{2}(x-x_0,y-y_0)H_f(x_0,y_0)\begin{pmatrix} x-x_0 \\ y-y_0 \end{pmatrix}$$

o, indicato con $\mathbf{v}$ il vettore $\mathbf{v} = (x-x_0, y-y_0)$, allora:

$$P_2((x_0,y_0)+\mathbf{v}) = f(x_0,y_0) + \nabla f(x_0,y_0)\cdot \mathbf{v} + \frac{1}{2}\mathbf{v}H_f(x_0,y_0)\mathbf{v}$$

e inoltre

$$f((x_0,y_0)+\mathbf{v}) = P_2((x_0,y_0)+\mathbf{v}) + o(\|\mathbf{v}\|^2).$$

Notiamo inoltre che, nell'ipotesi che f sia di classe C^2, allora, per il teorema di Schwartz, si ha che

$$\frac{\partial^2 f(x_0,y_0)}{\partial y\partial x} = \frac{\partial^2 f(x_0,y_0)}{\partial y\partial x}$$

e quindi la matrice Hessiana è una matrice simmetrica.

Esempio: Si consideri la funzione $f : \mathbb{R}^2 \to \mathbb{R}$ definita da $f(x,y) = x^2y^3$ e il punto $(x_0,y_0) = (1,1)$. Il polinomio di Taylor del secondo ordine di f in $(x_0,y_0) = (1,1)$ è dato da:

$$P_2(x,y) = f(x_0,y_0)+\nabla f(x_0,y_0)\cdot(x-x_0,y-y_0)+\frac{1}{2}(x-x_0,y-y_0)H_f(x_0,y_0)\begin{pmatrix} x-x_0 \\ y-y_0 \end{pmatrix},$$

dove $f(x_0,y_0) = 1$, $\frac{\partial f(x_0,y_0)}{\partial x} = 2$, $\frac{\partial f(x_0,y_0)}{\partial y} = 3$, $\frac{\partial^2 f(x_0,y_0)}{\partial x^2} = 2$, $\frac{\partial^2 f(x_0,y_0)}{\partial y^2} = 6$, $\frac{\partial^2 f(x_0,y_0)}{\partial x\partial y} = \frac{\partial^2 f(x_0,y_0)}{\partial y\partial x} = 6$, quindi

$$\begin{aligned} P_2(x,y) &= 1 + 2(x-1) + 3(y-1) + \frac{1}{2}\left(2(x-1)^2 + 12(x-1)(y-1) + 6(y-1)^2\right) \\ &= 1 + 2(x-1) + 3(y-1) + (x-1)^2 + 6(x-1)(y-1) + 3(y-1)^2 \end{aligned}$$

Massimi e minimi locali

Sia $f : A \subseteq \mathbb{R}^2 \to \mathbb{R}$ una funzione di due variabili a valori reali. Ricordiamo che un punto (x_0,y_0) è detto

- **massimo locale** se esiste un intorno $U(x_0,y_0)$ del punto (x_0,y_0) tale che per ogni $(x,y) \in U(x_0,y_0)$ si ha che $f(x,y) \leq f(x_0,y_0)$,
- **minimo locale** se esiste un intorno $U(x_0,y_0)$ del punto (x_0,y_0) tale che per ogni $(x,y) \in U(x_0,y_0)$ si ha che $f(x,y) \geq f(x_0,y_0)$.

Un criterio particolarmente utile per la ricerca dei punti di massimo e di minimo locale di una funzione f sufficientemente regolare è dato dal *teorema di Fermat.* Se f è derivabile (esistono le derivate parziali) in (x_0, y_0) e (x_0, y_0) è un punto di massimo o minimo locale, allora si ha che $\nabla f(x_0, y_0) = (0, 0)$.
I punti candidati ad essere punti di massimo o minimo locale sono quei punti in cui si annullano tutte le derivate parziali. Tali punti vengono chiamati *punti critici* o *punti stazionari.*

Una volta individuati i punti critici, il passo successivo consiste nel capire se sono punti di minimo locale, di massimo locale oppure *punti di sella.* Ricordiamo che un punto stazionario (x_0, y_0) è detto punto di sella se per ogni suo intorno $U(x_0, y_0)$ è sempre possibile trovare un punto $(x, y) \in U(x_0, y_0)$ tale che $f(x, y) > f(x_0, y_0)$ e un punto $(x', y') \in U(x_0, y_0)$ tale che $f(x', y') < f(x_0, y_0)$.
Esempi:

1. se consideriamo la funzione $f(x, y) = x^2 + 3y^2$, l'origine $(0, 0)$ è un punto di minimo locale

2. se consideriamo la funzione $f(x, y) = -x^2 - 3y^2$, l'origine $(0, 0)$ è un punto di massimo locale

3. se consideriamo la funzione $f(x, y) = x^2 - y^2$, l'origine $(0, 0)$ è un punto di sella

In generale la natura di un punto stazionario (x_0, y_0) dipende dal segno di $f(x, y) - f(x_0, y_0)$, lo scarto tra il valore assunto dalla funzione f nei punti (x, y) in un intorno di (x_0, y_0) e il valore $f(x_0, y_0)$. In particolare

- se esiste un intorno $U(x_0, y_0)$ di (x_0, y_0) tale che per ogni $(x, y) \in U(x_0, y_0)$ si ha che $f(x, y) - f(x_0, y_0) \leq 0$ allora (x_0, y_0) è un punto di massimo locale;

- se esiste un intorno $U(x_0, y_0)$ di (x_0, y_0) tale che per ogni $(x, y) \in U(x_0, y_0)$ si ha che $f(x, y) - f(x_0, y_0) \geq 0$ allora (x_0, y_0) è un punto di minimo locale;

- se per ogni intorno $U(x_0, y_0)$ di (x_0, y_0) è sempre possibile trovare un punto $(x, y) \in U(x_0, y_0)$ tale che $f(x, y) - f(x_0, y_0) > 0$ e un punto $(x', y') \in U(x_0, y_0)$ tale che $f(x', y') - f(x_0, y_0) < 0$ allora (x_0, y_0) è un punto di sella.

Per stimare il segno di $f(x, y) - f(x_0, y_0)$ in un intorno di (x_0, y_0) utilizziamo l'approssimazione di f tramite il polinomio di Taylor del secondo ordine in (x_0, y_0). Dato che $\nabla f(x_0, y_0) = (0, 0)$ allora abbiamo

$$f(x, y) = f(x_0, y_0) + \frac{1}{2}(x - x_0, y - y_0) H_f(x_0, y_0) \begin{pmatrix} x - x_0 \\ y - y_0 \end{pmatrix} + o(\|(x - x_0, y - y_0)\|^2),$$

quindi, se $\|(x-x_0, y-y_0)\|$ è piccola, abbiamo che

$$f(x,y)-f(x_0,y_0) \sim \frac{1}{2}(x-x_0, y-y_0)H_f(x_0,y_0)\begin{pmatrix} x-x_0 \\ y-y_0 \end{pmatrix}$$

e quindi il segno di $f(x,y)-f(x_0,y_0)$ in un intorno di (x_0,y_0) puøssere studiato analizzando il segno della forma quadratica nelle componenti (h,k) del vettore $\mathbf{v} = (x-x_0, y-y_0) \equiv (h,k)$:

$$(x-x_0, y-y_0)H_f(x_0,y_0)\begin{pmatrix} x-x_0 \\ y-y_0 \end{pmatrix} = \mathbf{v}H_f(x_0,y_0)\mathbf{v} = ah^2 + 2bhk + ck^2$$

dove $H_f(x_0,y_0) = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$, è la matrice Hessiana di f in (x_0,y_0) che ricordiamo essere simmetrica e quindi diagonalizzabile.
Se indichiamo con λ_1, λ_2 gli autovalori di $H_f(x_0,y_0)$, allora esiste un cambiamento di coordinate (una rotazione degli assi) $(h,k) = U(h',k')$ con U matrice 2×2 con $U^{-1} = U^T$, tale per cui nelle nuove coordinate la forma quadratica diviene

$$ah(h',k')^2 + 2bh(h',k')k(h',k') + ck(h',k')^2 = \lambda_1(h')^2 + \lambda_2(k')^2$$

In altre parole, indicando con $\mathbf{v}_1$ e $\mathbf{v}_2$ gli autovettori (normalizzati) di $H_f(x_0,y_0)$ relativi agli autovalori λ_1 e λ_2 rispettivamente, se scriviamo il vettore incremento $\mathbf{v} = (x-x_0, y-y_0)$ come combinazione lineare degli autovettori $\mathbf{v}_1$ e $\mathbf{v}_2$

$$\mathbf{v} = (x-x_0, y-y_0) = h'\mathbf{v}_1 + k'\mathbf{v}_2,$$

allora nelle variabili (h',k') abbiamo:

$$\mathbf{v}H_f(x_0,y_0)\mathbf{v} = \lambda_1(h')^2 + \lambda_2(k')^2.$$

A seconda del segno dei due autovalori λ_1, λ_2, possiamo trovarci in una situazione simile a quelle descritte negli esempi 1,2,3.

Una matrice simmetrica $n \times n$ è detta:

- *definita positiva* se tutti gli autovalori sono strettamente positivi,
- *definita negativa* se tutti gli autovalori sono strettamente negativi,
- *semidefinita positiva* se tutti gli autovalori sono positivi e almeno uno di questi è nullo,

- *semidefinita negativa* se tutti gli autovalori sono negativi e almeno uno di questi è nullo,
- *indefinita* se esistono almeno due autovalori di segno discorde,

Dai ragionamenti fatti sopra possiamo dunque concludere che, dato un punto stazionario (x_0, y_0), allora

- se $H_f(x_0, y_0)$ è definita positiva allora (x_0, y_0) è un punto di minimo locale,
- se $H_f(x_0, y_0)$ è definita negativa allora (x_0, y_0) è un punto di massimo locale,
- se $H_f(x_0, y_0)$ è indefinita allora (x_0, y_0) è un punto di sella,
- se $H_f(x_0, y_0)$ è semidefinita (positiva o negativa) allora la matrice Hessiana non dà sufficienti informazioni per capire la natura del punto stazionario (x_0, y_0) ed è necessario utilizzare altri metodi.

Esercizio 1:

Si consideri la funzione $f : \mathbb{R}^2 \to \mathbb{R}$

$$f(x, y) = 2x^3 - 6xy + 3y^2$$

Calcolare i punti critici di f e determinare se sono punti di massimo locale, di minimo locale o di sella.
Soluzione: Cerchiamo i punti in cui $\nabla f(x, y) = (6x^2 - 6y, -6x + 6y) = (0, 0)$, cioé le soluzioni del sistema:

$$\begin{cases} 6x^2 - 6y = 0 \\ -6x + 6y = 0 \end{cases}$$

che sono i punti $(0, 0)$ e $(1, 1)$.
La matrice Hessiana di f è

$$H_f(x, y) = \begin{pmatrix} 12x & -6 \\ -6 & 6 \end{pmatrix}$$

Consideriamo il punto $(0, 0)$. Abbiamo che

$$H_f(0, 0) = \begin{pmatrix} 0 & -6 \\ -6 & 6 \end{pmatrix}$$

dato che $\det H_f(0, 0) = -36 < 0$, allora, dato che il determinante di una matrice è uguale al prodotto dei suoi autovalori, possiamo concludere che $H_f(0, 0)$ ha due

autovalori di segno discorde e quindi $H_f(0,0)$ è indefinita. Il punto $(0,0)$ è quindi un punto di sella.
Consideriamo il punto $(1,1)$. Abbiamo che

$$H_f(1,1) = \begin{pmatrix} 12 & -6 \\ -6 & 6 \end{pmatrix}$$

dato che $\det H_f(1,1) = 36 > 0$, allora, dato che il determinante di una matrice è uguale al prodotto dei suoi autovalori, possiamo concludere che $H_f(1,1)$ ha due autovalori di segno concorde. Inoltre la traccia di $H_f(1,1)$ è positiva, infatti $\operatorname{Tr} H_f(1,1) = 18$. Dato che la traccia di una matrice è uguale alla somma dei suoi autovalori e, dall'analisi del segno del determinante, sapendo che i due autovalori hanno lo stesso segno, possiamo concludere che $H_f(1,1)$ è definita positiva e quindi $(1,1)$ è un punto di minimo locale.