Laboratorio di Matematica, A.A. 2009-2010; I modulo;

Lezione XV.

1. Nel seguito, quando non ci sia pericolo di fraintendimento, ometteremo le parentesi nella scrittura delle valutazioni di una funzione: ad esempio scriveremo fx al posto di $f(x)$.

   Ricordiamo che una funzione $f : U \to \mathbb{R}$ definita su un insieme convesso $U \subseteq \mathbb{R}^n$ si dice convessa su U se per ogni $x, y \in U$ l'arco di grafico di estremi (x, fx) e (y, fy) sta' sotto il segmento con gli stessi estremi, cioe' se

   $$f\left((1-t)x + ty\right) \leq (1-t)fx + tfy,$$

   per ogni $0 \leq t \leq 1$. Se la disuguaglianza vale sempre in senso stretto, diciamo che f e' *strettamente convessa.*

   Di seguito riportiamo una serie di esempi di classi di funzioni convesse.

   (a) Ciascuna funzione costante

   $$f : \mathbb{R}^n \to \mathbb{R}, \qquad f(x) = q \quad \forall x \in \mathbb{R}^n$$

   e' convessa. Cio' si verifica immediatamente come segue.
   Per ogni $x, y \in \mathbb{R}^n$ ed ogni $t \in \mathbb{R}$, da un lato si ha

   $$f\left((1-t)x + ty\right) = q$$

   e dall'altro

   $$(1-t)fx + tfy = (1-t)q + tq = q,$$

   cosi'

   $$f\left((1-t)x + ty\right) = (1-t)fx + tfy.$$

   (b) Ciascuna funzione lineare

   $$L : \mathbb{R}^n \to \mathbb{R},$$

   e' convessa. Cio' si verifica immediatamente come segue.
   Per ogni $x, y \in \mathbb{R}^n$ ed ogni $t \in \mathbb{R}$, si ha

   $$L\left((1-t)x + ty\right) = (1-t)L(x) + tL(y).$$

   (c) Una forma quadratica

   $$Q : \mathbb{R}^n \to \mathbb{R}, \qquad Q(x) = x^T Ax, \qquad \forall x \in \mathbb{R}^n,$$

   dove A e' una matrice simmetrica semidefinita positiva e' convessa. Cio' si verifica come segue.
   Siano $x, y \in \mathbb{R}^n$ e sia $0 < t < 1$. La disuguaglianza

   $$Q\left((1-t)x + ty\right) \leq (1-t)Qx + tQy$$

puo' essere riscritta via via come segue

$$((1-t)x+ty)^T A\,((1-t)x+ty) \leq (1-t)\ x^T Ax + t\ y^T Ay$$

$$(1-t)^2 x^T Ax + 2(1-t)t\ x^T Ay + t^2 y^T Ay \leq (1-t)\ x^T Ax + t\ y^T Ay$$

$$0 \leq \left((1-t)-(1-t)^2\right)\ x^T Ax - 2(1-t)t\ x^T Ay + \left(t-t^2\right)\ y^T Ay$$

$$0 \leq (1-t)t\ x^T Ax - 2(1-t)t\ x^T Ay + (1-t)t\ y^T Ay$$

Essendo $0 < t < 1$, si ha $(1-t)t > 0$, cosi' questa disuguaglianza e' equivalente alla disuguaglianza

$$0 \leq x^T Ax - 2x^T Ay + y^T Ay,$$

che si puo' riscrivere

$$0 \leq (x+y)^T A(x+y),$$

che e' verificata poiche' A e' definita positiva. Vale anche il viceversa: se la forma quadratica Q e' convessa, allora la matrice A e' semidefinita positiva. Dagli stessi conti si puo' ricavare che la forma quadratica Q e' strettamente convessa se e solo se la matrice A e' definita positiva.

(d) La funzione norma

$$\|\ \| : \mathbb{R}^n \to \mathbb{R}, \qquad x \mapsto \|x\|, \qquad \forall x \in \mathbb{R}^n,$$

e' convessa. Cio' si verifica come segue.
Siano $x, y \in \mathbb{R}^n$ e $0 < t < 1$. Per le proprieta' della norma si ha

$$\begin{aligned}\|(1-t)x+ty\| &\leq \|(1-t)x\| + \|ty\| \\ &= |1-t|\|x\| + |t|\|y\| = (1-t)\|x\| + t\|y\|.\end{aligned}$$

Le operazioni lineari sulle funzioni si comportano abbastanza bene rispetto alla convessita', nel senso della seguente Proposizione, la cui verifica e' immediata.

Proposizione 1. *Siano $f, q : U \to \mathbb{R}$ funzioni convesse su un convesso U, e sia $r \in \mathbb{R}^+$ un numero reale non negativo. Allora la funzione somma $f+g$ e la funzione prodotto rf sono convesse su U.*

Da questa Proposizione segue in particolare che ogni funzione polinomiale di secondo grado

$$P : \mathbb{R}^n \to \mathbb{R}, \qquad P(x) = q + a'x + x^T Ax,$$

dove $q \in \mathbb{R}$, $a' \in \mathbb{R}^{n\times 1}$, $A \in \mathbb{R}^{n\times n}$, sotto la condizione che A sia semidefinita positiva, e' convessa.

2. Le funzioni convesse derivabili con continuita' sono caratterizzate dal seguente Teorema, che non dimostriamo.

Teorema 1. *Una funzione $f : U \to \mathbb{R}$ di classe $\mathcal{C}^1$ su un insieme aperto convesso $U \subseteq \mathbb{R}^n$ e' convessa se e solo se per ogni $x \in U$ la funzione f e' maggiore-uguale alla sua approssimazione del I ordine in x, cioe'*

$$f(y) \geq f(x) + \nabla f(x)(y - x), \qquad \forall y \in U,$$

in altri termini

$$f(y) - f(x) \geq \nabla f(x)(y - x), \qquad \forall y \in U.$$

Le funzioni convesse derivabili due volte con continuita' sono caratterizzate dal seguente Teorema, che non dimostriamo.

Teorema 2. *Una funzione $f : U \to \mathbb{R}$ di classe $\mathcal{C}^2$ su un insieme aperto convesso $U \subseteq \mathbb{R}^n$ e' convessa se e solo se per ogni $x \in U$ la matrice Hessiana*

$$D^2 f(x)$$

e' semidefinita positiva.

3. **Esempi**

- Consideriamo la funzione $f : U \to \mathbb{R}$, definita da

  $$f(x, y) = \frac{1}{xy},$$

  dove $U = \{(x, y) : \ x, y > 0\}$. Il dominio di f e' convesso, ed f e' di classe $\mathcal{C}^k$ su U, per ogni $k = 0, 1, 2, \ldots$.
  La matrice Hessiana di f in e' data da

  $$D^2 f(x, y) = \begin{bmatrix} \frac{2}{x^3 y} & \frac{1}{x^2 y^2} \\ \frac{1}{x^2 y^2} & \frac{2}{x y^3} \end{bmatrix}.$$

  Osserviamo che per ogni $(x, y) \in U$ si ha

  $$\frac{\partial^2}{\partial x^2} (f(x, y)) = \frac{2}{x^3 y} > 0$$
  $$\left| D^2 f(x, y) \right| = \frac{3}{x^4 y^4} > 0,$$

  e dunque $D^2 f(x, y)$ e' definita positiva.
  Per il Th. 2, possiamo affermare che f e' convessa su U.
- Consideriamo la funzione $f : \mathbb{R}^2 \to \mathbb{R}$, definita da

  $$f(x, y) = -\exp\left(-\frac{x^2 + y^2}{2}\right).$$

Il dominio di f e' convesso, ed f e' di classe $\mathcal{C}^k$ su $\mathbb{R}^2$, per ogni $k = 0, 1, 2, \ldots$. La matrice Hessiana di f e' data da

$$D^2 f(x,y) = \begin{bmatrix} \exp\left(-\frac{x^2+y^2}{2}\right)(1-x^2) & -\exp\left(-\frac{x^2+y^2}{2}\right)xy \\ -\exp\left(-\frac{x^2+y^2}{2}\right)xy & \exp\left(-\frac{x^2+y^2}{2}\right)(1-y^2) \end{bmatrix}.$$

Questa matrice e' semidefinita positiva se e solo se

$$\begin{aligned} \frac{\partial^2}{\partial x^2}(f(x,y)) &= \exp\left(-\frac{x^2+y^2}{2}\right)(1-x^2) \geq 0 \\ \frac{\partial^2}{\partial y^2}(f(x,y)) &= \exp\left(-\frac{x^2+y^2}{2}\right)(1-y^2) \geq 0 \\ \left|D^2 f(x,y)\right| &= \exp\left(-(x^2+y^2)\right)(1-x^2-y^2) \geq 0, \end{aligned}$$

cioe' se e solo se

$$\begin{aligned} 1-x^2 &\geq 0 \\ 1-y^2 &\geq 0 \\ 1-x^2-y^2 &\geq 0, \end{aligned}$$

cioe'

$$1-x^2-y^2 \geq 0.$$

Per il Th. 2, possiamo affermare che f non e' convessa su $\mathbb{R}^2$; e' pero' convessa su ogni sottinsieme convesso del cerchio unitario centrato nell'origine.

4. Le funzioni convesse posseggono notevoli proprieta' rispetto ai loro estremi relativi.

Teorema 3. *Sia $f : U \to \mathbb{R}$ una funzione convessa di classe $\mathcal{C}^1$, e sia x^* un punto interno ad U. Sono equivalenti le seguenti proprieta'*

(a) x^* e' un punto di minimo assoluto per f;

(b) x^* e' un punto di minimo relativo per f;

(c) x^* e' un punto stazionario per f.

Dim. E' chiaro che la (a) implica la (b) che a sua volta implica la (c). Proviamo che la (c) implica la (a).

Supponiamo che $\nabla f(x^*) = 0$. Essendo f convessa, si ha

$$f(y) - f(x^*) \geq \nabla f(x^*)(y - x^*), \qquad \forall y \in U,$$

dunque

$$f(y) - f(x^*) \geq 0, \qquad \forall y \in U,$$

cioe'

$$f(y) \geq f(x^*), \qquad \forall y \in U.$$

Da questo teorema discende che, sotto le stesse ipotesi, *Se f non e' costante, ed x^* e' un punto di massimo assoluto per f su U, allora x^* non e' interno ad U (informalmente, sta' sul "bordo" di U).*

5. La struttura dell'insieme dei punti di minimo per una funzione convessa e' data dal seguente

   Teorema 4. *Sia $f : U \to \mathbb{R}$ una funzione convessa.*

   L'insieme dei punti di minimo assoluto per f su U e' convesso.

   Se f e' strettamente convessa, allora l'insieme dei punti di minimo assoluto per f contiene al piu' un punto.

   Dim. Supponiamo che ci sia in U un punto x^* di minimo assoluto per f, e poniamo $f(x^*) = m$. Proviamo solo la prima affermazione. Per ogni due punti x, y di minimo assoluto per f, anche ciascun punto

   $$(1-t)x + ty, \qquad 0 < t < 1$$

   del segmento che li unisce e' un punto di minimo assoluto per f, in quanto si ha

   $$m \leq f\left((1-t)x + ty\right) \leq (1-t)\ fx + t\ fy = (1-t)\ m + t\ m = m,$$

   da cui segue

   $$f\left((1-t)x + ty\right) = m.$$

   La seconda affermazione si puo' provare in modo simile.

   Riferimenti.

   Nella lezione precedente ed in questa si e' data una presentazione ridotta e semplificata degli argomenti che si trovano in: SB, Cap. 11, par. 11.1 e 11.2