Estremi vincolati ed hessiano orlato

Riccardo Dossena

10 aprile 2012

1 Richiami sulle forme quadratiche

Una *forma quadratica* è un'applicazione $q : \mathbb{R}^n \to \mathbb{R}$ che può essere scritta nella forma

$$q(\mathbf{h}) = \mathbf{h}^t A \mathbf{h} \quad \forall \mathbf{h} \in \mathbb{R}^n$$

dove A è una matrice simmetrica, detta *matrice associata* alla forma q ($\mathbf{h}^t$ è invece il vettore trasposto di $\mathbf{h}$).

1.1 Definizione. *Una forma quadratica q su $\mathbb{R}^n$ (o la corrispondente matrice simmetrica associata) è detta semidefinita positiva o semidefinita negativa quando*

$$q(\mathbf{h}) \geq 0 \quad \textit{o rispettivamente} \quad q(\mathbf{h}) \leq 0 \quad \forall \mathbf{h} \in \mathbb{R}^n.$$

Essa è detta definita positiva o definita negativa quando

$$q(\mathbf{h}) > 0 \quad \textit{o rispettivamente} \quad q(\mathbf{h}) < 0 \quad \forall \mathbf{h} \in \mathbb{R}^n \setminus \{\mathbf{0}\}.$$

La forma è detta indefinita quando non è semidefinita positiva, né semidefinita negativa.

Diamo ora un criterio per stabilire se una matrice A simmetrica è definita positiva o negativa. Per questo è necessario considerare le cosiddette *sottomatrici principali di nord-ovest*:

$$A_1 = (a_{11}) \qquad A_2 = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \qquad A_3 = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \qquad \dots \qquad A_n = A.$$

I determinanti di queste matrici si chiamano *minori principali di nord-ovest.*

1.2 Criterio di Sylvester. *Sia A una matrice simmetrica di ordine n. Allora:*

i) *A è definita positiva se e solo se $|A_k| > 0$ per ogni $k = 1, \dots, n$;*

ii) *A è definita negativa se e solo se $(-1)^k |A_k| > 0$ per ogni $k = 1, \dots, n$.*

Il prossimo teorema mostra una proprietà che si rivelerà utile per lo studio di massimi e minimi di funzioni di più variabili.

1.3 Teorema. *Sia* $q(\mathbf{h}) = \mathbf{h}^t A\mathbf{h}$ *una forma quadratica definita positiva su* $\mathbb{R}^n$. *Allora esiste una costante* $c > 0$ *tale che*

$$q(\mathbf{h}) = \mathbf{h}^t A\mathbf{h} \geq c|\mathbf{h}|^2 \quad \forall \mathbf{h} \in \mathbb{R}^n.$$

Analogamente, se $q(\mathbf{h}) = \mathbf{h}^t A\mathbf{h}$ *è una forma quadratica definita negativa su* $\mathbb{R}^n$, *esiste una costante* $c < 0$ *tale che*

$$q(\mathbf{h}) = \mathbf{h}^t A\mathbf{h} \leq c|\mathbf{h}|^2 \quad \forall \mathbf{h} \in \mathbb{R}^n.$$

Le dimostrazioni dei risultati precedenti possono essere trovate in qualsiasi libro di algebra lineare, ad esempio in [Strang, 2006] e in [Hoffman and Kunze, 1971].

2 Considerazioni sugli estremi liberi

Supponiamo di voler cercare gli estremi liberi di una funzione $f \in C^2(\Omega)$ con $\Omega \subseteq \mathbb{R}^2$. La condizione necessaria affinché un punto $(x_0, y_0) \in \Omega$ sia di massimo o minimo è che il gradiente in tale punto si annulli (cioè che il piano tangente abbia pendenza nulla; un tale punto viene detto *critico*)

$$\frac{\partial f}{\partial x}(x_0, y_0) = 0 \quad \text{e} \quad \frac{\partial f}{\partial y}(x_0, y_0) = 0.$$

Approssimiamo $f(x + dx, y + dy)$ mediante la formula di Taylor al secondo ordine nell'intorno di tale punto:

$$\begin{aligned} &f(x_0 + dx, y_0 + dy) = \\ &= f(x_0, y_0) + \frac{1}{2}\left[\frac{\partial^2 f}{\partial x^2}(x_0, y_0)dx^2 + 2\frac{\partial^2 f}{\partial x \partial y}(x_0, y_0)dxdy + \frac{\partial^2 f}{\partial y^2}(x_0, y_0)dy^2\right] + o(dx^2 + dy^2) \end{aligned}$$

per $|(dx, dy)| \to 0$. L'espressione tra parentesi quadre è il valore che la forma quadratica $d^2 f_{(x_0,y_0)}$ (differenziale secondo di f calcolato in (x_0, y_0)) associata alla matrice hessiana (simmetrica)

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

assume in (dx, dy). Per studiare la natura del punto critico (x_0, y_0) occorre dunque studiare e classificare la forma quadratica $d^2 f_{(x_0,y_0)}$ nelle variabili dx e dy. Infatti:

1. Se la forma quadratica $d^2 f_{(x_0,y_0)}$ è definita positiva, per il Teorema 1.3 esiste una costante $c > 0$ tale che

$$d^2 f_{(x_0,y_0)}(dx, dy) \geq c|(dx, dy)|^2$$

da cui, ricordando che $|(dx, dy)|^2 = dx^2 + dy^2$

$$\begin{aligned} d^2 f_{(x_0,y_0)}(dx, dy) + o(dx^2 + dy^2) &\geq c(dx^2 + dy^2) + o(dx^2 + dy^2) \\ &\geq (dx^2 + dy^2)(c + o(1)) \end{aligned}$$

e per $|(dx, dy)|$ abbastanza piccolo, $c + o(1) > 0$, dunque il secondo membro della seguente uguaglianza

$$f(x_0 + dx, y_0 + dy) - f(x_0, y_0) = \frac{1}{2} d^2 f_{(x_0,y_0)}(dx, dy) + o(dx^2 + dy^2)$$

è positivo per ogni dx, dy in un opportuno intorno I di 0. Quindi

$$f(x_0 + dx, y_0 + dy) - f(x_0, y_0) \geq 0 \quad \forall dx, dy \in I$$

ovvero (x_0, y_0) è un punto di *minimo relativo*.

2. Analogamente, se la forma quadratica $d^2 f_{(x_0,y_0)}$ è definita negativa, dal Teorema 1.3 segue che in un opportuno intorno di (x_0, y_0) l'incremento di f è negativo, e quindi (x_0, y_0) è un punto di *massimo relativo*.

3. Se la forma quadratica $d^2 f_{(x_0,y_0)}$ è indefinita, in modo analogo si dimostra che in ogni intorno di (x_0, y_0) l'incremento di f cambia segno. Pertanto (x_0, y_0) non è un punto né di massimo né di minimo, ma è un *punto di sella*.

4. Nel caso in cui l'hessiano sia solo semidefinito, non si hanno conclusioni e occorre uno studio di tipo diverso.

3 Estremi vincolati: moltiplicatori di Lagrange

Si vogliano cercare gli estremi di una funzione $f \in C^2(\Omega)$, con $\Omega \subseteq \mathbb{R}^2$, soggetta al vincolo $g(x, y) = c$, con $g \in C^2(\Omega)$. Supponiamo che il nostro punto di estremo (x_0, y_0) sia *regolare* per il vincolo, cioè

$$\nabla g(x_0, y_0) \neq (0, 0). \tag{3.1}$$

in modo che per il Teorema del Dini l'equazione del vincolo definisca implicitamente un arco di curva regolare passante per (x_0, y_0) (e quindi sia ben definito in quel punto il versore tangente al vincolo). La condizione necessaria affinché questo punto sia di estremo per la f vincolata, analogamente al caso degli estremi liberi, è che sia nulla la *derivata direzionale* di f lungo la direzione tangente alla curva descritta dal vincolo (cioè l'unica direzione ammissibile)

$$D_{\mathbf{t}} f(x_0, y_0) = 0$$

dove $\mathbf{t}$ è il versore tangente al vincolo. Da ciò si deduce che

$$\nabla f(x_0, y_0) \cdot \mathbf{t} = 0$$

cioè $\nabla f(x_0, y_0)$ e $\mathbf{t}$ sono ortogonali. Anche $\nabla g(x_0, y_0)$ e $\mathbf{t}$ sono ortogonali, dato che il gradiente di una funzione è ortogonale alle linee di livello. Dunque i due vettori $\nabla f(x_0, y_0)$ e $\nabla g(x_0, y_0)$ devono essere paralleli tra loro. Questo si riassume nel seguente teorema:

3.1 Teorema dei moltiplicatori di Lagrange. *Se* (x_0, y_0) *è un punto di estremo vincolato per* f *sotto il vincolo* $g(x, y) = c$ *e* $\nabla g(x_0, y_0) \neq (0, 0)$*, allora esiste un numero* λ*, detto moltiplicatore di Lagrange, tale che*

$$\nabla f(x_0, y_0) = \lambda \nabla g(x_0, y_0). \tag{3.2}$$

Se introduciamo la funzione *lagrangiana* definita da

$$Z(x, y, \lambda) = f(x, y) + \lambda[c - g(x, y)]$$

il teorema afferma che se (x_0, y_0) è un punto di estremo vincolato, allora esiste λ_0 tale che il punto (x_0, y_0, λ_0) è un punto critico libero per Z. Infatti i punti critici di Z sono le soluzioni del sistema

$$\begin{cases} Z_x(x, y, \lambda) = f_x(x, y) - \lambda g_x(x, y) = 0 \\ Z_y(x, y, \lambda) = f_y(x, y) - \lambda g_y(x, y) = 0 \\ Z_\lambda(x, y, \lambda) = c - g(x, y) = 0 \end{cases} \tag{3.3}$$

nel quale le prime due equazioni coincidono con la (3.2), mentre la terza esprime la condizione del vincolo.

4 Estremi vincolati: alla ricerca di condizioni sufficienti

Anche in questo caso, per una funzione $z = f(x, y)$ soggetta al vincolo $g(x, y) = c$ dobbiamo studiare il segno del differenziale secondo d^2z valutato in un punto stazionario. Da notare, però, che non tutte le direzioni sono ammesse, ma solo quelle $(dx, dy) \neq (0, 0)$ tali per cui

$$dg = g_x dx + g_y dy = 0$$

come si ottiene differenziando l'equazione del vincolo $g(x, y) = c$ ($dc = 0$ perché costante). Dunque dx e dy non possono essere scelti arbitrariamente, ma è necessario trattare ad esempio dy come variabile dipendente:

$$dy = -\frac{g_x}{g_y} dx.$$

Una volta assegnato un valore a dx, dy dipenderà solo da g_x e g_y, quindi da x e y. Dunque la formula per il differenziale d^2z, basata su valori arbitrari dx e dy, non può essere applicata così com'è. Per trovare un'espressione di d^2z che faccia al caso nostro bisogna trattare dy come variabile dipendente da x e y durante la differenziazione. Quindi,

$$\begin{aligned} d^2z &= d(dz) = \frac{\partial(dz)}{\partial x} dx + \frac{\partial(dz)}{\partial y} dy \\ &= \frac{\partial}{\partial x}(f_x dx + f_y dy) dx + \frac{\partial}{\partial y}(f_x dx + f_y dy) dy \\ &= \left[f_{xx} dx + \left(f_{xy} dy + f_y \frac{\partial dy}{\partial x} \right) \right] dx + \left[f_{yx} dx + \left(f_{yy} dy + f_y \frac{\partial dy}{\partial y} \right) \right] dy \\ &= f_{xx} dx^2 + f_{xy} dy dx + f_y \frac{\partial(dy)}{\partial x} dx + f_{yx} dx dy + f_{yy} dy^2 + f_y \frac{\partial(dy)}{\partial y} dy \end{aligned}$$

e dato che il terzo e il sesto termine possono essere riscritti

$$f_y \left[\frac{\partial(dy)}{\partial x} dx + \frac{\partial(dy)}{\partial y} dy \right] = f_y d(dy) = f_y d^2 y$$

l'espressione per $d^2 z$ diventa

$$d^2 z = f_{xx} dx^2 + 2 f_{xy} dx dy + f_{yy} dy^2 + f_y d^2 y \tag{4.4}$$

in cui compare in più il termine $f_y d^2 y$. Questa presenza tuttavia "squalifica" $d^2 z$ come forma quadratica. Comunque, $d^2 z$ può essere trasformato in una forma quadratica grazie al vincolo $g(x, y) = c$. Dal momento che il vincolo implica $dg = 0$ e anche $d^2 g = d(dg) = 0$, con lo stesso procedimento usato per arrivare alla (4.4) otteniamo

$$(d^2 g =)\ g_{xx} dx^2 + 2 g_{xy} dx dy + g_{yy} dy^2 + g_y d^2 y = 0$$

e ricavando da questa $d^2 y$ e sostituendola nella (4.4), possiamo scrivere $d^2 z$ come forma quadratica:

$$d^2 z = \left(f_{xx} - \frac{f_y}{g_y} g_{xx} \right) dx^2 + 2 \left(f_{xy} - \frac{f_y}{g_y} g_{xy} \right) dx dy + \left(f_{yy} - \frac{f_y}{g_y} g_{yy} \right) dy^2.$$

Dalla seconda delle (3.3) si ha che $\lambda = f_y / g_y$, dunque si può scrivere

$$d^2 z = (f_{xx} - \lambda g_{xx})\, dx^2 + 2\, (f_{xy} - \lambda g_{xy})\, dx dy + (f_{yy} - \lambda g_{yy})\, dy^2.$$

Osserviamo ora che derivando due volte la lagrangiana otteniamo

$$\begin{aligned} Z_{xx} &= f_{xx} - \lambda g_{xx} \\ Z_{xy} &= f_{xy} - \lambda g_{xy} = Z_{yx} \\ Z_{yy} &= f_{yy} - \lambda g_{yy} \end{aligned}$$

e possiamo esprimere finalmente $d^2 z$ in termini della lagrangiana stessa come segue:

$$d^2 z = Z_{xx} dx^2 + 2 Z_{xy} dx dy + Z_{yy} dy^2.$$

Le condizioni sufficienti per i massimi e minimi vincolati della funzione $z = f(x, y)$ si possono determinare studiando il segno del differenziale $d^2 z$, valutato in un punto critico (x_0, y_0, λ_0) della lagrangiana. Ricordiamo però che in questa situazione non dobbiamo considerare il segno di $d^2 z$ per tutti i possibili valori di $(dx, dy) \neq (0, 0)$, ma solo per quelli che soddisfano la condizione:

$$g_x dx + g_y dy = 0.$$

Dunque le *condizioni sufficienti del secondo ordine* sono:

- Per il massimo di z: $d^2 z$ definita negativa, soggetta a $dg = 0$;
- per il minimo di z: $d^2 z$ definita positiva, soggetta a $dg = 0$.

5 L'hessiano orlato

Come nel caso degli estremi liberi, è possibile esprimere la condizione sufficiente del secondo ordine in forma di determinante. Al posto del determinante hessiano, nel caso di estremi vincolati introdurremo il cosiddetto *hessiano orlato*.

Analizziamo dapprima le condizioni per il segno di una forma quadratica in due variabili, soggetta a un vincolo lineare:

$$q = au^2 + 2huv + bv^2 \quad \text{soggetta a} \quad \alpha u + \beta v = 0.$$

Dal momento che il vincolo implica $v = -(\alpha/\beta)u$, possiamo riscrivere q come funzione di una variabile soltanto

$$q = au^2 - 2h\frac{\alpha}{\beta}u^2 + b\frac{\alpha^2}{\beta^2}u^2 = (a\beta^2 - 2h\alpha\beta + b\alpha^2)\frac{u^2}{\beta^2}\,.$$

È quindi ovvio che q è definita positiva (negativa) se e solo se l'espressione tra parentesi è positiva (negativa). Ora, il determinante

$$\begin{vmatrix} 0 & \alpha & \beta \\ \alpha & a & h \\ \beta & h & b \end{vmatrix} = 2h\alpha\beta - a\beta^2 - b\alpha^2$$

è proprio l'*opposto* dell'espressione tra parentesi. Di conseguenza, possiamo affermare che

$$q \text{ soggetta a } \alpha u + \beta v = 0 \text{ è } \begin{Bmatrix} \text{definita positiva} \\ \text{definita negativa} \end{Bmatrix} \quad \text{se e solo se} \quad \begin{vmatrix} 0 & \alpha & \beta \\ \alpha & a & h \\ \beta & h & b \end{vmatrix} \begin{cases} < 0 \\ > 0 \end{cases}$$

Quando questo viene applicato alla forma quadratica d^2z, abbiamo naturalmente che le variabili u e v diventano dx e dy rispettivamente e per quanto riguarda il vincolo $\alpha = g_x$ e $\beta = g_y$. Così otteniamo il seguente criterio per il segno di d^2z:

$$d^2z \text{ soggetta a } dg = 0 \text{ è } \begin{Bmatrix} \text{definita positiva} \\ \text{definita negativa} \end{Bmatrix} \quad \text{se e solo se} \quad \begin{vmatrix} 0 & g_x & g_y \\ g_x & Z_{xx} & Z_{xy} \\ g_y & Z_{yx} & Z_{yy} \end{vmatrix} \begin{cases} < 0 \\ > 0 \end{cases}$$

dove l'ultima matrice prende il nome di *hessiano orlato*.

Abbiamo così dimostrato il seguente teorema.

5.1 Teorema. *Sia $f \in C^2(\Omega)$ con $\Omega \subseteq \mathbb{R}^2$, soggetta al vincolo $g(x, y) = c$ con $g \in C^2(\Omega)$. Sia $(x_0, y_0) \in \Omega$ tale che $g(x_0, y_0) = c$ e $\nabla g(x_0, y_0) \neq (0, 0)$. Si consideri la funzione lagrangiana*

$$Z(x, y, \lambda) = f(x, y) + \lambda[c - g(x, y)]$$

e sia (x_0, y_0, λ_0) una soluzione del sistema

$$\begin{cases} Z_x(x, y, \lambda) = f_x(x, y) - \lambda g_x(x, y) = 0 \\ Z_y(x, y, \lambda) = f_y(x, y) - \lambda g_y(x, y) = 0 \\ Z_\lambda(x, y, \lambda) = c - g(x, y) = 0. \end{cases}$$

Si consideri inoltre il determinante hessiano orlato

$$|\overline{H}| = \begin{vmatrix} 0 & g_x & g_y \\ g_x & Z_{xx} & Z_{xy} \\ g_y & Z_{yx} & Z_{yy} \end{vmatrix} \quad \textit{valutato in } (x_0, y_0, \lambda_0)$$

(i) se $|\overline{H}| > 0$, il punto (x_0, y_0) è un massimo vincolato per f;

(ii) se $|\overline{H}| < 0$, il punto (x_0, y_0) è un minimo vincolato per f;

(iii) se $|\overline{H}| = 0$, nessuna conclusione può essere tratta per (x_0, y_0).

Osserviamo che se il determinante hessiano orlato è positivo o negativo, allora è automaticamente soddisfatta la condizione $\nabla g(x_0, y_0) \neq (0, 0)$: in caso contrario il determinante sarebbe infatti nullo.

Precisazioni

Spesso, anziché considerare il vincolo nella forma $g(x, y) = c$, si usa scrivere $\varphi(x, y) = 0$ con $\varphi(x, y) = c - g(x, y)$. La lagrangiana si scrive perciò

$$Z(x, y, \lambda) = f(x, y) + \lambda\varphi(x, y).$$

In questo caso si può comunque considerare l'hessiano orlato nella forma

$$\begin{vmatrix} 0 & \varphi_x & \varphi_y \\ \varphi_x & Z_{xx} & Z_{xy} \\ \varphi_y & Z_{yx} & Z_{yy} \end{vmatrix}$$

mantenendo le stesse conclusioni. Infatti risulta

$$\begin{vmatrix} 0 & \varphi_x & \varphi_y \\ \varphi_x & Z_{xx} & Z_{xy} \\ \varphi_y & Z_{yx} & Z_{yy} \end{vmatrix} = \begin{vmatrix} 0 & -g_x & -g_y \\ -g_x & Z_{xx} & Z_{xy} \\ -g_y & Z_{yx} & Z_{yy} \end{vmatrix} = \begin{vmatrix} 0 & g_x & g_y \\ g_x & Z_{xx} & Z_{xy} \\ g_y & Z_{yx} & Z_{yy} \end{vmatrix}$$

dato che moltiplicando una riga e una colonna della matrice per -1, il determinante risulta moltiplicato per $(-1)^2$.

Riferimenti bibliografici

[Bramanti *et al.*, 2004] Bramanti, M., Pagani, C.D. e Salsa, S., *Matematica. Calcolo infinitesimale e algebra lineare*, seconda edizione, Zanichelli, Bologna, 2004.

[Chiang, 1984] Chiang, A.C., *Fundamental Methods of Mathematical Economics*, 3rd edition, McGraw-Hill, New York, 1984.

[Gilardi, 1992] Gilardi, G., *Analisi due*, McGraw-Hill, Milano, 1992.

[Hoffman and Kunze, 1971] Hoffman, K. and Kunze, R., *Linear Algebra*, 2nd edition, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1971.

[Marsden and Tromba, 2003] Marsden, J.E. and Tromba, A., *Vector Calculus*, W.H. Freeman, 2003.

[Simon and Blume, 1994] Simon, C.P. and Blume, L., *Mathematics for Economists*, W.W. Norton & Company, New York, 1994.

[Strang, 2006] Strang, G., *Linear Algebra and Its Applications*, Thomson, Brooks/Cole, 2006.