Condizioni di Karush-Kuhn-Tucker e Programmazione Lineare

A. Agnetis[*]

1 Richiami su condizioni di Karush-Kuhn-Tucker e convessità

Si consideri il problema di ottimizzazione vincolata:

$$\begin{array}{r} \min \; f(x) \\ x \in X \subset R^n \end{array} \tag{1}$$

dove $X = \{x|h(x) = 0, g(x) \geq 0\}$, in cui h e g sono vettori di funzioni di $x \in R^n$. Si è visto in una precedente dispensa che, se f, g e h hanno derivate parziali continue in x^*, se x^* è un punto di minimo locale ed è un punto regolare, esiste un vettore λ^* tale che valgono le seguenti condizioni del primo ordine (condizioni di Karush-Kuhn-Tucker):

$$\begin{aligned}
\nabla_x L(x^*, \lambda^*) &= 0 && (2)\\
h_i(x^*) &= 0 \quad i = 1, \ldots, m && (3)\\
g_j(x^*) &\geq 0 \quad j = m+1, \ldots, m+p && (4)\\
\lambda_j^* &\geq 0 \quad j = m+1, \ldots, m+p && (5)\\
\lambda_j^* g_j(x^*) &= 0 \quad j = m+1, \ldots, m+p && (6)
\end{aligned}$$

dove L è la *funzione lagrangiana*:

$$L(x, \lambda) = f(x) - \sum_{i=1}^{m} \lambda_i h_i(x) - \sum_{j=m+1}^{m+p} \lambda_j g_j(x).$$

Richiamiamo ora alcuni risultati relativi al caso convesso. Due proprietà delle funzioni convesse, stabilite nel caso non vincolato, sussistono anche nel caso vincolato (con dimostrazioni praticamente identiche a quelle del caso non vincolato, per cui verranno omesse):

[*]Dipartimento di Ingegneria dell'Informazione - Università di Siena

TEOREMA 1 *Si consideri una funzione f convessa in R^n, e un insieme ammissibile X convesso. Se x^* è un punto di minimo locale, è anche un punto di minimo globale.* $\square$

TEOREMA 2 *Si consideri una funzione f, convessa in R^n, e un insieme ammissibile X convesso. L'insieme dei punti di minimo è convesso.* $\square$

Alla luce di questi due risultati, risulta allora interessante stabilire delle condizioni sui vettori h e g tali che la regione ammissibile X sia convessa.

TEOREMA 3 *Dato un problema del tipo (1), se tutte le funzioni $h_i(x) = a_i^T x + b_i$ sono lineari ($i = 1, \ldots, m$) e tutte le $-g_j(x)$ sono convesse ($j = m+1, \ldots, m+p$), allora la regione ammissibile X è convessa.*

Dim.– Siano x e y due punti ammissibili, ossia tali che

$$
\begin{aligned}
h(x) = 0, \;\; g(x) \geq 0 \\
h(y) = 0, \;\; g(y) \geq 0
\end{aligned}
$$

e si consideri un punto $z = \lambda x + (1-\lambda)y$, $0 \leq \lambda \leq 1$, ossia ottenuto come combinazione convessa di x e y. Dalla definizione delle h_i si ha

$$
\begin{aligned}
h_i(z) = h_i(\lambda x + (1-\lambda)y) = a_i^T(\lambda x + (1-\lambda)y) + b = \\
\lambda a_i^T x + (1-\lambda)a_i^T y + b = -\lambda b - (1-\lambda)b + b = 0
\end{aligned}
$$

mentre per la convessità delle $-g_j$:

$$-g_j(z) = -g_j(\lambda x + (1-\lambda)y) \leq -\lambda g_j(x) - (1-\lambda)g_j(y)$$

dal momento che $0 \leq \lambda \leq 1$, si ha $-g(z) \leq 0$ e dunque $z \in X$. $\square$

2 Programmazione Lineare

Un problema di Programmazione Lineare (PL) è un problema del tipo (1) in cui f, g e h sono funzioni lineari. Ricordando che le funzioni lineari sono sia concave che convesse, siamo nelle ipotesi del Teorema 3, e quindi la regione ammissibile X è convessa. Dunque, la PL è un caso particolare di programmazione convessa, e quindi un minimo locale è anche un minimo globale. Senza perdita di generalità, com'è noto, possiamo porre un problema di Programmazione Lineare in *forma standard*:

$$
\begin{aligned}
\min \;\; & c^T x \\
Ax &= b \\
x &\geq 0
\end{aligned}
\tag{7}
$$

dove A è una matrice $m \times n$, $c \in R^n$, $b \in R^m$ e $x \in R^n$. Supporremo qui che $m < n$ e che la A abbia rango pieno, ossia m. Quello che vogliamo fare è applicare le condizioni di KKT (2)-(6) a un problema di PL in forma standard.

Nello scrivere la funzione lagrangiana, esprimiamo separatamente i moltiplicatori associati ai vincoli di uguaglianza (che indicheremo con $u \in R^m$) da quelli associati ai vincoli di non negatività (che indicheremo con $s \in R^n$):

$$L(x, u, s) = c^T x - u^T (Ax - b) - s^T x$$

dunque, dalle (2)-(6), abbiamo che, se x^* è un punto di minimo, allora devono esistere due vettori u^* e s^* tali che:

$$\begin{aligned}
\nabla_x L(x^*, u^*, s^*)^T &= c^T - u^{*T} A - s^{*T} = 0^T && (8)\\
Ax^* &= b && (9)\\
x^* &\geq 0 && (10)\\
s^* &\geq 0 && (11)\\
s_j^* x_j^* &= 0 \text{ per ogni } j = 1, \ldots, n && (12)
\end{aligned}$$

Peraltro, poiché $s^* \geq 0$ e $x^* \geq 0$, la condizione (12) può equivalentemente scriversi come $s^{*T} x^* = 0$. Inoltre, dalla (8) possiamo ricavare $s^* = c^T - u^{*T} A$ e in definitiva eliminare s^* riscrivendo le condizioni come:

$$\begin{aligned}
Ax^* &= b && (13)\\
x^* &\geq 0 && (14)\\
c^T - u^{*T} A &\geq 0 && (15)\\
(c^T - u^{*T} A) x^* &= 0 && (16)
\end{aligned}$$

Ricordiamo che, in un generico problema di programmazione non lineare, le condizioni KKT sono necessarie soltanto per punti regolari. Nel caso lineare, è possibile invece mostrare che le condizioni di KKT *continuano a essere necessarie* anche in punti non regolari. Vale dunque il seguente teorema:

TEOREMA 4 *Dato un problema di PL (7), se x^* è un punto di minimo, allora esistono vettori u^* e s^* tali da soddisfare (13)-(16).*

Una importante osservazione riguarda il confronto tra due quantità:

TEOREMA 5 *Se x^* e u^* soddisfano le (13)-(16), allora $c^T x^* = b^T u^*$.*

Dim.– Dalla (16)

$$(c^T - u^{*T} A)x^* = c^T x^* - u^{*T} A x^* = 0$$

ossia, dal momento che $Ax^* = b$,

$$c^T x^* - u^{*T} b = 0.$$

□

3 Il problema duale

Le condizioni di KKT con riferimento a un certo problema di PL possono essere interpretate in modo più efficace introducendo un altro problema di PL strettamente legato, come vedremo, al primo.

Definizione 1 *Dato un problema di PL in forma standard (7), definiamo* problema duale *il PL:*

$$\begin{aligned} \max\ & b^T y \\ & A^T y \leq c \end{aligned} \tag{17}$$

□

Nel contesto di questa definizione, il problema (7) viene indicato come problema *primale*. Per capire il significato di questa definizione, scriviamo le condizioni di KKT per il problema duale. Si osservi che, essendo (17) un problema di massimizzazione, nella funzione lagrangiana la funzione obiettivo è cambiata di segno. Indicando con v i moltiplicatori di Lagrange, abbiamo allora

$$L(y, v) = -b^T y - v^T (c - A^T y)$$

e andando a scrivere le condizioni di KKT:

$$\begin{aligned} \nabla_y L(y^*, v^*)^T &= -b^T + v^{*T} A^T = 0^T && (18) \\ c - A^T y^* &\geq 0 && (19) \\ v^* &\geq 0 && (20) \\ (c - A^T y^*)_j v_j^* &= 0 \ \text{ per ogni } j = 1, \ldots, n && (21) \end{aligned}$$

in considerazione delle (19) e (20), la (21) può scriversi anche come $(c^T - y^{*T}A)v^*$, e dunque in definitiva, se u^* è soluzione ottima del problema (17), deve esistere un vettore v^* tale che

$$\begin{aligned}
Av^* &= b && (22)\\
c^T - y^{*T}A &\geq 0 && (23)\\
v^* &\geq 0 && (24)\\
(c^T - y^{*T}A)v^* &= 0 && (25)
\end{aligned}$$

Teorema 6 *Dato un problema di PL (7), se e solo se esistono una soluzione x^* ammissibile e un vettore u^* tali da soddisfare (13)-(16), allora esistono un vettore y^* ammissibile per il problema duale e un vettore v^* tali da soddisfare le (22)-(25).*

Dim.– Confrontando le (13)-(16) con le (22)-(25), osserviamo che sono le stesse, se identifichiamo x^* con v^* e u^* con y^*. Dunque, se e solo se x^* e u^* sono tali da soddisfare le (13)-(16), ponendo $v^* := x^*$ e $y^* := u^*$, risultano soddisfatte anche le (22)-(25). In particolare, il vettore y^* così definito è ammissibile per il problema duale (per la (23)), e soddisfa le condizioni di KKT con valori dei moltiplicatori v^* dati da x^* (condizioni (22),(24) e (25)). □

In definitiva, i valori dei moltiplicatori di Lagrange u^* associati alla soluzione ottima del problema primale *sono* i valori delle variabili duali all'ottimo, e simmetricamente, i valori dei moltiplicatori di Lagrange v^* associati alla soluzione ottima del problema duale sono i valori delle variabili primali all'ottimo. Per questo motivo, si può parlare di condizioni di KKT associate a una *coppia* di problemi primale/duale.

Quanto detto mostra che tra problema primale e problema duale sussista una profonda simmetria. In effetti, non è difficile provare il seguente risultato:

Teorema 7 *Dato un problema di PL, il duale del problema duale coincide con il primale.*

Dim.– Vedi ad esempio il libro di Fischetti, pag.55. □

Possiamo a questo punto individuare ancora meglio gli stretti legami che intercorrono tra i due problemi. Anzitutto, vale la seguente proprietà, nota col nome di *dualità debole*:

Teorema 8 *Si consideri una coppia primale-duale di problemi di PL, e siano $\bar{x}$ e $\bar{y}$ soluzioni ammissibili dei due problemi rispettivamente. Allora, $c^T\bar{x} \geq b^T\bar{y}$.*

Dim.– Dalla definizione dei due problemi: $c^T\bar{x} \geq (\bar{y}^TA)\bar{x} = \bar{y}^Tb$. □

Dunque, data una *qualsiasi* coppia di soluzioni *ammissibili* per i due problemi, il valore della funzione obiettivo del problema primale è sempre maggiore o uguale del valore della funzione obiettivo del problema duale, e di conseguenza in nessun caso $b^T y$ può superare $c^T x$, se x e y sono soluzioni ammissibili dei due problemi. Una immediata conseguenza del Teorema 8 è quindi:

TEOREMA 9 *Siano $\bar{x}$ e $\bar{y}$ soluzioni ammissibili per una coppia primale-duale di problemi di PL. Se $c^T \bar{x} = b^T \bar{y}$, allora $\bar{x}$ e $\bar{y}$ sono ottime per i rispettivi problemi.*

A questo punto possiamo chiudere il cerchio. Infatti, dato un problema di PL (primale) in forma standard:

1. Se x^* è una soluzione ottima del problema primale, allora (Teorema 4) deve soddisfare le KKT;

2. Se un punto x^* soddisfa le KKT, allora il vettore u^* dei moltiplicatori è ammissibile per il problema duale (Teorema 6) ed è tale che $c^T x^* = b^T u^*$ (Teorema 5);

3. Se x^* e u^* sono ammissibili per i rispettivi problemi e si ha che $c^T x^* = b^T u^*$, allora x^* e u^* sono una coppia di soluzioni ottime per i rispettivi problemi (Teorema 9).

In definitiva, possiamo concludere che, nel caso della programmazione lineare, queste tre affermazioni sono *equivalenti* relativamente a un punto x^* :

- x^* è un punto di minimo (globale)
- esiste un vettore u^* tale che le (13)-(16) sono soddisfatte
- esiste un vettore u^*, ammissibile per il problema duale, tale che $c^T x^* = b^T u^*$

Ricapitolando, dato un problema di programmazione lineare in forma standard e il suo duale, si ha che (x^*, u^*) sono una coppia di soluzioni ottime per i rispettivi problemi se e solo se x^* e u^* soddisfano le condizioni di KKT. Dunque, mentre in generale le KKT sono solo necessarie, *nel caso della programmazione lineare sono anche sufficienti* a garantire che un punto sia di minimo (sia esso regolare o no); ossia, in altre parole, nel caso della programmazione lineare le KKT caratterizzano completamente i punti di minimo.

Il risultato espresso dal Teorema 9, di fondamentale importanza nella programmazione lineare, è noto come *Teorema della dualità forte*. Per completare il quadro, consideriamo il caso in cui uno dei due problemi non ammetta soluzione ottima.

TEOREMA 10 *Se, in una coppia primale-duale di problemi di PL, uno dei due problemi è illimitato, l'altro non ammette soluzione ammissibile.*

Dim.– Supponiamo che il primale sia illimitato. Se il duale avesse una soluzione ammissibile $\bar{u}$, per il teorema della dualità debole, il valore della funzione obiettivo del problema primale non potrebbe scendere al di sotto di $\bar{u}^T b$, contraddicendo cosi' l'ipotesi. $\square$

In definitiva, le condizioni di KKT per un problema di PL possono essere enunciate nel seguente modo: un punto $\bar{x}$ è soluzione ottima di un problema di PL in forma standard se e solo se è ammissibile ed esiste una soluzione $\bar{u}$, ammissibile per il problema duale, tale che

$$(c_j - \bar{u}^T A_j)\bar{x}_j = 0 \ \text{ per ogni } j = 1, \ldots, n$$

Il metodo del simplesso propone a ogni iterazione una soluzione ammissibile per il primale, ossia tale da soddisfare le (9) e (10). A ogni passo, è possibile associare alla soluzione corrente un vettore $\bar{u}$ costruito in modo tale da soddisfare le condizioni di complementarietà (12). Se tale vettore $\bar{u}$ è anche ammissibile per il problema duale, ossia soddisfa le (8) e (11), allora $\bar{x}$ e $\bar{u}$ sono una coppia di soluzioni ottime per i rispettivi problemi.