Capitolo 2

Trasformazione di Problemi Non Lineari

2.1 Trasformazione in problema di PL

In questa sezione, verranno presentati tre classi di problemi di programmazione non lineare che, attraverso l'uso di opportune tecniche, possono essere trasformati in problemi di programmazione lineare.

2.1.1 Problemi Min-Max

Consideriamo un problema avente la seguente forma:

$$\min_{x \in X} \quad \max\{f_1(x), \ldots, f_l(x)\}. \tag{2.1}$$

Assumiamo che X sia un poliedro e che le funzioni $f_i,\ i = 1, \ldots, l$ siano lineari:

$$f_i(x) = a_i^T x + b_i, \quad i = 1, \ldots, l.$$

Possiamo facilmente verificare che tale problema è non lineare. Introduciamo una variabile ausiliaria z e riscriviamo il problema come segue:

$$\begin{aligned}
&\min_{x,z} z \\
&z = \max\{f_1(x), \ldots, f_l(x)\} \\
&x \in X.
\end{aligned}$$

A questo punto, possiamo sostituire il vincolo della funzione max con un insieme di vincoli ed ottenere il problema:

$$\begin{aligned}
&\min_{x,z} z \\
&f_i(x) \leq z, \quad i = 1, \ldots, l \\
&x \in X
\end{aligned}$$

e, ricordando che le funzioni sono tutte lineari, scrivere:

$$\begin{aligned}
&\min_{x,z} z \\
&a_i^T x + b_i \leq z, \quad i = 1, \ldots, l \\
&x \in X.
\end{aligned}$$

Si possono ottenere soluzioni ammissibili del problema in cui abbiamo

$$z > \max\{f_i(x), i = 1, \dots, l\}.$$

L'obiettivo del problema ci assicura, però , che all'ottimo il valore z sia proprio pari a $\max\{f_i(x), i = 1, \dots, l\}$. Possiamo verificare facilmente che nel caso Max-Max il problema trasformato non ammette soluzione in quanto sarebbero ammissibili valori arbitrariamente grandi di z.

2.1.2 Problemi con valore assoluto

Consideriamo ora il problema seguente:

$$\begin{aligned}&\min_{x,y} \sum_j c_j |x_j| + \sum_k d_k y_k \\ &(x, y) \in C\end{aligned}$$

Assumiamo che C sia un poliedro e che $c_j \geq 0 \ \forall j$. Trasformiamo le variabili nel seguente modo:

$$x = x_j^+ - x_j^-, \quad x_j^+ \geq 0, x_j^- \geq 0.$$

Tale trasformazione non è univoca:

- caso 1: $x_j \geq 0$. Otteniamo $x_j^+ = x_j + \delta = |x_j| + \delta$ e $x_j^- = \delta$,
- caso 2: $x_j < 0$. Otteniamo $x_j^+ = \delta$ e $x_j^- = -x_j + \delta = |x_j| + \delta$,

Con $\delta \geq 0$. Se $\delta = 0$, una delle due componenti è nulla e l'altra è uguale a $|x_j|$. Notiamo inoltre che

$$x_j^+ + x_j^- = |x_j| + 2\delta.$$

Sostituendo nell'obiettivo, al posto di $|x_j|$, la somma di x_j^+ e x_j^-, abbiamo che all'ottimo (essendo $c_j \geq 0$) $x_j^+ = 0$ oppure $x_j^- = 0$ (o equivalentemente $\delta = 0$). Possiamo riscrivere il modello come segue:

$$\begin{aligned}&\min_{x,y} \sum_j c_j (x_j^+ + x_j^-) + \sum_k d_k y_k \\ &(x_j^+ - x_j^-, y) \in C \\ &x_j^+ \geq 0, x_j^+ \geq 0, \quad = 1, \dots, n\end{aligned}$$

Una formulazione alternativa si ottiene tenendo conto del fatto che

$$|x_j| = \max\{x_j, -x_j\}.$$

Sostituendo nel problema iniziale otteniamo:

$$\begin{aligned}&\min_{x,y} \sum_j c_j \max\{x_j, -x_j\} + \sum_k d_k y_k \\ &(x, y) \in C\end{aligned}$$

e introducendo le variabili ausiliarie z_j, riscriviamo come:

$$\begin{aligned}
&\min_{x,y,z} \sum_j c_j z_j + \sum_k d_k y_k \\
&z_j = \max\{x_j, -x_j\} \\
&(x, y) \in C.
\end{aligned}$$

Analogamente al caso Min-Max, abbiamo:

$$\begin{aligned}
&\min_{x,y,z} \sum_j c_j z_j + \sum_k d_k y_k \\
&-z_j \leq x_j \leq z_j \quad j = 1, \ldots, n \\
&(x, y) \in C.
\end{aligned}$$

2.1.3 Problemi frazionari

Consideriamo ora il problema di minimizzare il rapporto di funzioni affini su un poliedro:

$$\begin{aligned}
&\min_x \frac{c^T x + d}{e^T x + f} \\
\\
&Gx \leq h \\
&Ax = b
\end{aligned}$$

supponiamo che per ogni x ammissibile abbiamo $e^T x + f > 0$. Usando la trasformazione:

$$y = \frac{x}{e^T x + f}$$

e

$$z = \frac{1}{e^T x + f},$$

possiamo trasformare il problema scritto in precedenza, nel seguente problema di programmazione lineare:

$$\begin{aligned}
&\min_{y,z} c^T y + dz \\
\\
&Gy - hz \leq 0 \\
&Ay - bz = 0 \\
&e^T y + fz = 1 \\
&z \geq 0.
\end{aligned}$$

2.2 Problemi Equivalenti

In molti casi, come abbiamo visto nella precedente sezione, può essere utile trasformare un problema dato in un altro problema a esso equivalente (nel senso che una soluzione del problema originario è ottenibile a partire da una soluzione di quello trasformato). Riportiamo di seguito un risultato teorico che definisce un criterio di equivalenza tra problemi.

2.2.1 Risultato di Equivalenza

Consideriamo un problema di ottimizzazione generico:

$$\begin{aligned} &\min_x f(x) \\ &x \in X \end{aligned} \tag{2.2}$$

con $X \in R^n$ e $f:\ R^n \to R$, ed un problema

$$\begin{aligned} &\min_y g(x) \\ &y \in Y \end{aligned} \tag{2.3}$$

con $Y \in R^m$ e $g:\ R^m \to R$. Si dice che i due problemi sono equivalenti, se l'esistenza della soluzione di un problema implica ed è implicata dall'esistenza della soluzione di un altro problema ed esiste, inoltre, una corrispondenza tra le soluzioni dei due problemi.

Proposizione 2.2.1 *Consideriamo i problemi* (2.2) *e* (2.3)*. Supponiamo che esistano due trasformazioni $\beta : X \to Y$ e $\gamma : Y \to X$, tra gli insiemi dei due problemi, tali che*

(i) per ogni $x \in X$, l'elemento $\beta(x) \in Y$ soddisfa $g(\beta(x)) \leq f(x)$

(ii) per ogni $y \in Y$, l'elemento $\gamma(y) \in X$ soddisfa $f(\gamma(y)) \leq g(y)$.

Allora il problema (2.2) *ammette solluzione ottima se e solo se* (2.3) *ammette soluzione otttima.*

Prova. Supponiamo che il problema (2.2) abbia una soluzione ottima $x^\star \in X$. Considerando l'ipotesi (i) possiamo trovare un punto $y^\star = \beta(x^\star) \in Y$ tale che $g(y^\star) = g(\beta(x^\star)) \leq f(x^\star)$. Vogliamo dimostrare che tale punto è una soluzione ottima del problema (2.3). Supponiamo per assurdo che esista un punto $\tilde{y} \in Y$ tale che $g(\tilde{y}) < g(y^\star)$. per l'ipotesi (ii) il punto $\tilde{x} = \gamma(\tilde{y}) \in X$ è tale che $f(\tilde{x}) = f(\gamma(\tilde{y})) \leq g(\tilde{y})$. Tenendo conto delle precedenti disuguaglianze e del fatto che $x^\star$ è minimo di (2.2) per ipotesi, abbiamo la seguente contraddizione:

$$f(x^\star) \leq f(\tilde{x}) = f(\gamma(\tilde{y})) \leq g(\tilde{y}) < g(y^\star) = g(\beta(x^\star)) \leq f(x^\star).$$

L'implicazione inversa si mostra in maniera analoga. $\square$

2.2.2 L'Esempio del Problema Min-Max

Possiamo applicare il risultato appena descritto ai problemi Min-Max descritti in precedenza. Consideriamo il problema:

$$\min_{x \in X} \quad \max\{f_1(x), \dots, f_l(x)\}. \tag{2.4}$$

Introduciamo una variabile ausiliaria z e riscriviamo il problema come segue:

$$\begin{aligned} &\min_{x,z} z \\ &f_i(x) \leq z \quad i = 1, \dots, l \\ &x \in X. \end{aligned}$$

Possiamo definire due trasformazioni:

$$\beta(x) = \begin{pmatrix} \max\{f_1(x), \ldots, f_l(x)\} \\ x \end{pmatrix}$$

e

$$\gamma(y) = \gamma\left(\begin{pmatrix} z \\ x \end{pmatrix}\right) = x.$$

Verifichiamo facilmente che:

- $\beta(x) \in Y$ e che $\gamma(y) \in X$;
- $g(\beta(x)) = \max\{f_1(x), \ldots, f_l(x)\} = f(x)$;
- $f(\gamma(y)) = f(x) = \max\{f_1(x), \ldots, f_l(x)\} \leq z = g(y)$.

Per cui risultano soddisfatte le ipotesi della Proposizione 2.2.1.

2.3 Modelli di PNL: Stima dei parametri

Supponiamo di voler stimare i parametri di un modello matematico (lineare) connessi ad uno specifico problema fisico, avendo a disposizione un insieme finito di misure sperimentali. Sia

$$y = a^T x + b$$

il modello utilizzato, dove

- $x \in R^n$ rappresenta l'ingresso del modello;
- $y \in R$ rappresenta l'uscita del modello;
- $a \in R^n$ e $b \in R$ rappresenta l'insieme dei parametri da determinare.

Supponiamo, inoltre, di avere a disposizione un insieme di coppie ingresso uscita (campioni):

$$\{(x^1, y^1), \ldots, (x^m, y^l)\}.$$

Per ogni coppia definiamo un errore tra l'uscita misurata e l'uscita del modello:

$$E_i = y^i - (a^T x^i + b).$$

L'obiettivo è quello di determinare il vettore dei parametri che meglio rappresenta il fenomeno in analisi, ovvero quello che minimizza gli errori E_i, $i = 1, \ldots, l$.

2.3.1 Formulazione Min-Max

Formuliamo il problema di stima dei parametri come:

$$\min_{a,b} \max_{i} |y^i - a^T x^i - b|$$

e, utilizzando le trasformazioni viste in precedenza, otteniamo:

$$\begin{array}{l} \min\limits_{a,b,z} z \\ |y^i - a^T x^i - b| \leq z \quad \forall\ i = 1, \ldots, l \end{array}$$

ovvero

$$\begin{array}{l} \min\limits_{a,b,z} z \\ -z \leq y^i - a^T x^i - b \leq z \quad \forall\ i = 1, \ldots, l. \end{array}$$

2.3.2 Formulazione in norma ℓ_1

Una diversa formulazione del problema è :

$$\min_{a,b} \sum_{i=1}^{l} |y^i - a^T x^i - b|$$

e, utilizzando le trasformazioni viste in precedenza, otteniamo:

$$\begin{array}{l} \min\limits_{a,b,z} \sum\limits_{i=1}^{l} z_i \\ |y^i - a^T x^i - b| \leq z_i \quad \forall\ i = 1, \ldots, l \end{array}$$

ovvero

$$\begin{array}{l} \min\limits_{a,b,z} \sum\limits_{i=1}^{l} z_i \\ -z_i \leq y^i - a^T x^i - b \leq z_i \quad \forall\ i = 1, \ldots, l. \end{array}$$

Utilizzando la formulazione alternativa, possiamo scrivere:

$$\begin{array}{l} \min\limits_{a,b,v,u} \sum\limits_{i=1}^{l} v_i + u_i \\ v_i - u_i = y^i - a^T x^i - b \ \forall\ i = 1, \ldots, l \\ v \geq 0,\ u \geq 0. \end{array}$$

2.3.3 Rappresentazione Sparsa di Segnali

La rappresentazione sparsa di segnali reali è alla base di vari metodi di compressione (e.g. formati MPEG, JPEG). Dati

- un segnale reale $b \in R^m$,
- un dizionario di segnali standard $A = [a_1 \ldots a_n]$, con $A \in R^{m \times n}$ e $m \ll n$,

l'obiettivo è quello di rappresentare b come combinazione lineare del minor numero possibile di elementi del dizionario. Il problema può esser formalmente descritto come segue

$$\begin{aligned}&\min_x \|x\|_1\\&Ax = b,\end{aligned}$$

con $\|x\|_1 = \sum_{i=1}^n |x_i|$. Anche in questo caso, utilizzando una delle trasformazioni viste in precedenza, otteniamo:

$$\begin{aligned}&\min_{x,y} \sum_{i=1}^n y_i\\&Ax = b,\\&-y \le x \le y.\end{aligned}$$