

Sum of squares.

Sono dati n punti in uno spazio Euclideo in d dimensioni. Si vuole trovare il modo ottimo di partizionarli in p sottoinsiemi. Il costo di una partizione è dato dalla somma dei costi di ogni sottoinsieme di punti. Il costo di ogni sottoinsieme è dato dalla somma dei quadrati delle distanze dei punti del sottoinsieme da un altro punto, rappresentativo del sottoinsieme, che deve essere posizionato opportunamente.

Formulare il problema, classificarlo e risolverlo con i dati dell'esempio riportato nel file `SUMOFSQUARES.TXT`.

Discutere l'ottimalità e l'unicità della soluzione trovata.

Esempio.

I punti da partizionare sono 16, lo spazio Euclideo considerato ha 4 dimensioni, il numero di sottoinsiemi è 3.

Punto	Coord 1	Coord 2	Coord 3	Coord 4
1	25	61	110	-57
2	32	-63	90	50
3	28	25	-14	-34
4	-41	-30	56	-29
5	70	-81	-58	30
6	-97	17	-71	-68
7	103	12	-90	67
8	12	-31	135	65
9	-28	-26	85	21
10	33	-78	84	-92
11	-51	44	-23	-94
12	40	-79	-27	63
13	-99	80	39	-70
14	96	-11	-33	38
15	5	6	-12	-91
16	-40	48	140	100

Table 1: Coordinate dei punti dati.

Soluzione.

Dati. Sia N l'insieme dei punti, D l'insieme delle dimensioni, P l'insieme dei sottoinsiemi. Sia x_i^k la k -esima coordinata del punto i -esimo, per ogni $i \in N$ e

ogni $k \in D$.

Variabili. Le variabili del problema devono descrivere sia il modo in cui vengono associati gli n punti ai p sottoinsiemi, sia le posizioni dei p punti da localizzare in ogni sottoinsieme. Occorrono quindi $n \times p$ variabili binarie di assegnamento $w_{ij} \ \forall i \in N, j \in P$ e $p \times d$ variabili continue e libere $X_j^k \ \forall j \in P, k \in D$ per le coordinate dei rappresentanti dei sottoinsiemi.

Vincoli. Gli unici vincoli del problema sono in vincoli di assegnamento, che impongono che ogni punto dato sia assegnato ad esattamente uno dei p sottoinsiemi.

$$\sum_{j \in P} w_{ij} = 1 \quad \forall i \in N.$$

Obiettivo. La funzione obiettivo è data da

$$\text{minimize } z = \sum_{i \in N, j \in P} w_{ij} \sum_{k \in D} (x_i^k - X_j^k)^2.$$

Classificazione. Dato che la funzione obiettivo è non-lineare nelle variabili continue e dato che la formulazione comprende sia variabili binarie che variabili continue, il modello risultante è di programmazione non-lineare misto-intera.

Per ogni scelta delle variabili di assegnamento, il sotto-problema di ottimizzazione non-lineare è convesso, poiché il problema si scompone in tanti sotto-problemi indipendenti di localizzazione ottima quanti sono i sottoinsiemi e ciascuno di essi è convesso.

Quindi la soluzione trovata è garantita essere un ottimo globale solo se il solutore utilizzato esegue un'enumerazione implicita dei possibili assegnamenti.

Soluzione calcolata da Knitro: $z^* = 101027$ con sottoinsiemi $\{5, 7, 12, 14\}$, $\{2, 4, 8, 9, 10, 16\}$ e $\{1, 3, 6, 11, 13, 15\}$. Soluzione calcolata da Baron (interrotto) e da Lindoglobal (interrotto): $z^* = 96137,9$ con sottoinsiemi $\{5, 7, 12, 14\}$, $\{1, 2, 4, 8, 9, 10, 16\}$ e $\{3, 6, 11, 13, 15\}$.