

# 1 Lezione del 26-11-24

## 1.1 Introduzione alla programmazione non lineare

### Problema 1.1: di programmazione non lineare

In una nuova linea di produzione sono integrati 6 robot, ciascuno dei quali può rotare di 360 gradi attorno all'asse verticale. Le aree di lavoro di ciascun robot non devono sovrapporsi, in modo da evitare collisioni fra di essi. Inoltre, i robot devono essere collegati fra di loro attraverso cavi in fibra ottica, e poichè questi sono costosi, è necessario minimizzare la distanza da coprire.

I raggi di lavoro dei robot sono, poniamo, i seguenti:

Robot	Raggio
1	120
2	80
3	100
4	70
5	45
6	120

Il problema sarà quindi quello di individuare i 6 punti sul piano  $(x_1, y_1), (x_2, y_2), \dots, (x_6, y_6)$  che corrispondono alle posizioni degli assi verticali di ogni robot, imponendo il vincolo di distanza fra due robot adiacenti  $i$  e  $j$  di  $r_i + r_j$ . La funzione obiettivo sarà quindi la quantità di cavo in fibra ottica, cioè la distanza fra ogni coppia  $i, j$  di robot. Possiamo quindi disporre un modello:

$$\begin{cases} \max \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} + \sqrt{(x_2 - x_3)^2 + (y_2 - y_3)^2} + \dots + \sqrt{(x_5 - x_6)^2 + (y_5 - y_6)^2} \\ d_{ij} \geq r_i + r_j, \quad \forall i = 1, 2, \dots, 6, \quad \forall j = 1, 2, \dots, 6, \quad i < j \end{cases}$$

su 12 variabili e  $\frac{n(n-1)}{2} \Big|_{n=6} = 15$  vincoli. Nelle prossime lezioni vedremo come risolvere questo tipo di problemi, notando per ora che tutte le funzioni che abbiamo generato finora sono **quadratiche** tolte le radici (che possiamo fare conservando massimi e minimi).

In generale, nella programmazione non lineare, vorremo avere le nostre funzioni  $\in C^1$  in modo da poter derivare, cioè ricavare il gradiente e stabilire punti di massimo e di minimo, attraverso il teorema di Fermat:

### Teorema 1.1: di Fermat

Sia data una funzione  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ , con adeguate condizioni di continuità. Si ha che in ogni punto di massimo o minimo locale  $\bar{x}$  vale:

$$\nabla f(\bar{x}) = 0$$

Addirittura, fosse la funzione  $\in C^2$  potremmo ricavare l'Hessiana come ulteriore discriminante, ricordando che *Hessiana definita positiva* significa **punto di minimo**, e viceversa *Hessiana definita negativa* significa **punto di massimo** (con le opportune considerazioni di definizione positiva/negativa stretta e non).

Vediamo quindi di fare una classificazione di funzioni utili ai fini dell'ottimizzazione.

### 1.1.1 Funzioni quadratiche

Diamo quindi una forma generale di polinomi di secondo grado su cui fare ottimizzazione. Una funzione tipo potrebbe essere:

$$f(x) = \frac{1}{2}x^T Q x + c^T x$$

imponendo, vedremo poi come mai, la matrice  $Q$  definita positiva.

Di cui potrebbe essere un esempio:

$$f(x) = ax_1^2 + bx_1x_2 + cx_2^2 + dx_1 + ex_2$$

Vediamo come trovare la matrice  $Q$  e il vettore  $c$ . Nel caso in due variabili, che continuiamo per la convenzione della ricerca operativa a chiamare  $x_1$  e  $x_2$ , avremo per calcolo diretto, spezzando le componenti in  $Q$  e in  $C$ :

- $Q$ :

$$\begin{aligned} \frac{1}{2}x^T Q x &= \frac{1}{2} \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} q_{11}x_1 & q_{12}x_2 \\ q_{21}x_1 & q_{22}x_2 \end{pmatrix} \\ &= \frac{1}{2} (q_{11}x_1^2 + (q_{12} + q_{21})x_1x_2 + q_{22}x_2^2) \end{aligned}$$

sarà quindi:

$$\begin{cases} q_{11} = 2a \\ q_{12} = b \\ q_{21} = b \\ q_{22} = 2c \end{cases}$$

- $c$ : valgono le stesse regole viste sulle lineari, cioè nel nostro esempio:

$$\begin{cases} c_1 = d \\ c_2 = e \end{cases}$$

Abbiamo quindi stabilito che esiste una corrispondenza fra polinomi di secondo grado e matrici quadrate simmetriche di dimensione 2, dove la definizione positiva indica la presenza di un solo minimo globale della quadratica.

Si può poi dimostrare che in generale esiste una corrispondenza fra **funzioni polinomiali di secondo grado** e **matrici quadrate simmetriche** di qualsiasi dimensione.

Notiamo che, stabilita la  $Q$  e la  $c$ , il calcolo di gradiente risulta immediato in quanto:

$$\begin{cases} \nabla f = Qx + c \\ Hf = Q \end{cases}$$

da cui si spiega l' $\frac{1}{2}$  accanto a  $Q$ , che "normalizza" il modulo dell'Hessiana.

### 1.1.2 Funzioni convesse

Diamo innanzitutto una definizione:

**Definizione 1.1: Funzione convessa**

Una funzione convessa è qualsiasi funzione che rispetta:

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y)$$

con  $\forall \lambda \in [0, 1], \forall x, \forall y \in \mathbb{R}^n$ .

la definizione di funzione convessa stabilisce essenzialmente per una qualsiasi retta intercettante la funzione  $f$ , in qualsiasi punto compreso fra le intercette la funzione raggiunge un valore minore o uguale al valore della retta calcolato nello stesso punto.

Una proprietà importante delle funzioni convesse è che il gradiente è dato da una funzione **monotona crescente**, cioè  $f(x) < f(y) \forall x < y$ , e quindi che l'Hessiana è **semidefinita positiva** ( $\geq 0$  nel caso scalare).

Una conseguenza immediata di quest'ultima considerazione è che in una funzione convessa, ogni punto dove  $\nabla f(x) = 0$  è necessariamente **minimo**.

Notiamo alcune altre proprietà delle funzioni convesse:

- Una funzione convessa può non avere minimo: si pensi all'esponenziale, che tende a 0 senza raggiungerlo mai, e quindi ha solo un *limite inferiore*;
- Il minimo locale di una funzione convessa è anche minimo globale;
- Le funzioni convesse non hanno selle.

### 1.1.3 Funzioni coercive

Definiamo infine un'ultima classe, data da:

**Definizione 1.2: Funzione coerciva**

Una funzione coerciva è una funzione che rispetta il limite:

$$\lim_{|x| \rightarrow +\infty} f(x) = +\infty$$

Possiamo quindi essere sicuri che una funzione coerciva non ha massimo globale, ma al massimo solo massimi locali.