Appunti di Ricerca operativa A.A.2021-2022

Programmazione Non Lineare

Gabriele Frassi (**g.frassi2@studenti.unipi.it**)

26 aprile 2023

Sommario

La dispensa vuole porsi come supporto per il corso di *Ricerca operativa* della triennale di Ingegneria informatica dell'Università di Pisa (prof. Pappalardo). Il presente PDF include i miei appunti relativi alla *Programmazione Non Lineare* (PNL).

Alcuni ringraziamenti dovuti: Alex Parri per avermi sopportato nelle mie domande stupide, Alessandro Simonelli per il supporto nella stesura di parte delle definizioni e dei teoremi, Andrea Bedini per la segnalazione di alcuni errori individuati durante il suo periodo di preparazione all'esame.

Indice

I Unimap **3**

II Lezioni **7**

1 Lezioni di PNL **8**

1.1 Prerequisiti Analisi II . . . 8
1.1.1 Funzioni . . . 8
1.1.2 Derivate . . . 9
1.1.2.1 Derivate parziali . . . 9
1.1.2.2 Gradiente . . . 9
1.1.2.3 Matrice hessiana . . . 10
1.1.3 Usare gradiente ed hessiana per trovare max/min . . . 10
1.1.3.1 Teoremi per l'individuazione del minimo locale . . . 11
1.1.3.2 Teoremi per l'individuazione del massimo locale . . . 12
1.1.4 Teoremi per l'individuazione di massimi e minimi globali . . . 12
1.1.5 Restrizioni . . . 13
1.2 Funzione *quadprog* . . . 15
1.3 Funzione *fmincon* . . . 15
1.4 Introduzione alla programmazione non lineare (PNL) . . . 16
1.4.1 Problema: punti di min/max locali e globali . . . 16
1.4.2 Problema: numero di soluzioni da considerare . . . 16
1.4.3 Algoritmo di enumerazione totale (inefficiente) . . . 17
1.5 Classificazione delle funzioni . . . 17
1.5.1 Funzioni quadratiche . . . 17
1.5.1.1 Numero di soluzioni del sistema $\nabla f(x) = 0$ e autovalori . . . 18
1.5.2 Funzioni coercive . . . 19
1.5.3 Funzioni convesse . . . 20
1.5.3.1 Teorema sui punti stazionari in funzioni convesse . . . 20
1.6 Algoritmi iterativi per la risoluzione . . . 21
1.6.1 Confronto tra argomenti passati e algoritmi della PNL . . . 23
1.6.2 Regole di stop . . . 23
1.7 Metodo del gradiente (con ricerca esatta del passo) . . . 23
1.7.1 Spiegazione . . . 23
1.7.2 Teorema di convergenza . . . 24
1.7.3 Esempio di passo del metodo (1) . . . 24
1.7.4 Esempio di passo del metodo (2) . . . 25

1.8 Domini della PNL (o regione ammissibile) 26
1.8.1 Definizione di dominio 26
1.8.2 Confronto con le funzioni non vincolate 26
1.8.3 Proprietà del dominio che ci interessano 27
1.8.4 Esempi introduttuvi sulle proprietà dei domini 28
1.9 Condizione di ottimalità per problemi vincolati 30
1.9.1 Teorema (Lagrange-Karush-Kulm-Tucker, LKKT) 30
1.9.2 Spiegazione 31
1.9.3 Teorema complementare a LKKT (C.S.) 32
1.9.4 Primo esempio di studio 33
1.9.5 Secondo esempio di studio 36
1.10 Metodo delle restrizioni (distinzione tra selle e max/min) 38
1.10.1 Osservazione introduttiva 38
1.10.2 Esercizio precedente 38
1.11 Metodo di Frank-Wolfe 39
1.11.1 Spiegazione 39
1.11.2 Variante: problema di massimo 40
1.11.3 Teorema di convergenza 40
1.11.4 Criteri di stop 41
1.11.5 Esempio 1 42
1.11.6 Esempio 2 43
1.12 Metodo del gradiente proiettato 43
1.12.1 Spiegazione 43
1.12.2 Variante: problema di massimo 44
1.12.3 Teorema di convergenza 44
1.12.4 Esempio 1 45
1.13 Schema concettuale per la classificazione 46

III Appendici **49**

A Definizioni **50**
A.1 Programmazione Non Lineare 50

B Teoremi **53**
B.1 Programmazione Non Lineare 53

Parte I

Unimap

1. **Lun 28/02/2022 08:30-10:30 (2:0 h)** lezione: Problemi di Ricerca Operativa. Variabili decisionali, funzione obiettivo e vincoli. Modello matematico. La Programmazione Lineare (PL) in formato primale standard. Un problema di produzione ottima. Soluzione ammissibile, soluzione ottima. (MASSIMO PAPPALARDO)
2. **Mar 01/03/2022 08:30-10:30 (2:0 h)** lezione: Risoluzione geometrica di un problema di PL in 2 variabili. Curve di isocosto. Distinzione tra PL e PLI (Programmazione Lineare Intera). Definizione di poliedro. Trasformazioni canoniche per forme standard. (MASSIMO PAPPALARDO)
3. **Mer 02/03/2022 08:30-10:30 (2:0 h)** lezione: Combinazioni convesse e combinazioni coniche. Involucro convesso e involucro conico. Teorema di rappresentazione dei poliedri. Il problema dell'assegnamento di costo minimo. Caso cooperativo e caso non cooperativo. (MASSIMO PAPPALARDO)
4. **Ven 04/03/2022 08:30-10:30 (2:0 h)** lezione: Teorema fondamentale della PL. Esempi ed esercizi. Forma duale standard. Le conversioni ai formati standard. Esempi ed esercizi. (MASSIMO PAPPALARDO)
5. **Lun 07/03/2022 08:30-10:30 (2:0 h)** lezione: Verici di un poliedro in formato primale standard. Soluzioni di base ammissibili e non ammissibili. Il caso degenere. La function "linprog" e i suoi parametri. (MASSIMO PAPPALARDO)
6. **Mar 08/03/2022 08:30-10:30 (2:0 h)** lezione: Il problema duale standard. La teoria della dualità. Il teorema della dualità. I vertici del poliedro duale standard. Soluzioni di base. Il caso degenere. Esempi ed esercizi. (MASSIMO PAPPALARDO)
7. **Mer 09/03/2022 08:30-10:30 (2:0 h)** lezione: Il problema del trasporto ed il suo modello matematico. Il duale del duale. Test di ottimalità per un vertice del duale. Esempi ed esercizi. I vertici del poliedro dell'assegnamento e del poliedro del trasporto. (MASSIMO PAPPALARDO)
8. **Ven 11/03/2022 08:30-10:30 (2:0 h)** lezione: L'algoritmo del simplesso primale. Correttezza e convergenza in un numero finito di passi. Esempi svolti. Interpretazione geometrica del simplesso primale. (MASSIMO PAPPALARDO)
9. **Lun 14/03/2022 08:30-10:30 (2:0 h)** lezione: Algoritmo del simplesso duale. Regole anticiclo di Bland. Esercizi svolti in aula. (MASSIMO PAPPALARDO)
10. **Mar 15/03/2022 08:30-10:30 (2:0 h)** lezione: Il problema duae ausiliario. Esempio svolto. Il caso degenere. Interpretazione geometrica del simplesso duale. (MASSIMO PAPPALARDO)
11. **Mer 16/03/2022 08:30-10:30 (2:0 h)** lezione: Problema del caricamento ("zaino"). Modello matematico: caso intero e caso binario. Il metodo dei rendimenti per le vlutazioni superiori. Algoritmo dei rendimenti per le valutazioni infeirori. Esempio svolto. (MASSIMO PAPPALARDO)
12. **Ven 18/03/2022 08:30-10:30 (2:0 h)** lezione: Relazioni tra PL e PLI: valutazioni superiori ed inferiori. Il prblema del T.S:P: asimmetrico. Relazioni tra assegnamenti e cicli hamiltoniani. I vncoli di connessione per l'eliminazione dei cicli disgiunti. (MASSIMO PAPPALARDO)
13. **Lun 21/03/2022 08:30-10:30 (2:0 h)** lezione: Problema del TSP simmetrico. I vincoli di grado. I k-alberi. La valutazione inferiore come k-albero di costo minimo. L'algoritmo del nodo più vicino per la valutaione superiore. Esempi ed esercizi. (MASSIMO PAPPALARDO)
14. **Mar 22/03/2022 08:30-10:30 (2:0 h)** lezione: Teorema di equivalenza tra PL e PLI. Disuguaglianze valide e piani di taglio. I piani di taglio di Gomory. (MASSIMO PAPPALARDO)

15. **Mer 23/03/2022 08:30-10:30 (2:0 h)** lezione: Esercitazione di ricapitolazione di tutti gli argomenti svolti. (MASSIMO PAPPALARDO)
16. **Ven 25/03/2022 08:30-10:30 (2:0 h)** lezione: Il problema del "bin-packing". Il modello matematico e i vincoli di semiassegnamento. Valutazione inferiore come problema di PL. Algoritmi "greedy" come valutazioni superiori: First-Fit-Decreasing (FFT), Next-Fit-Decreasing (NFT) e Best-Fit-Decreasing (BFT). (MASSIMO PAPPALARDO)
17. **Lun 28/03/2022 08:30-10:30 (2:0 h)** lezione: Il "Branch and Bound". Regole di taglio per problemi di minimo. L'aggiornamento della soluzione ammissibile corrente. Esercizi sul TSP simmetrico. (MASSIMO PAPPALARDO)
18. **Mar 29/03/2022 08:30-10:30 (2:0 h)** lezione: Il "Branch and Bound" per problemi di massimo. Modifiche delle regole di taglio. Esercizi sullo zaino binario. (MASSIMO PAPPALARDO)
19. **Mer 30/03/2022 08:30-10:30 (2:0 h)** lezione: Problemi di copertura. Riduzione della matrice di copertura. Righe "dominate". Il modello matematico. Valutazioni inferiori e superiori. Gestione di vincoli aggiuntivi nella PLI. (MASSIMO PAPPALARDO)
20. **Ven 01/04/2022 08:30-10:30 (2:0 h)** lezione: Il problema di massima copertura. Il modello matematico ed i comandi intlinprog. Algoritmi di Chvatal per soluzioni ammissibili di problemi di copertura e di massima copertura. (MASSIMO PAPPALARDO)
21. **Lun 04/04/2022 08:30-10:30 (2:0 h)** lezione: Il problema del flusso di costo minimo su reti bilanciate: modello matematico. La matrice di incidenza della rete e le equazioni di bilancio. Gli alberi di copertura come basi. I flussi di base. Esercizi. (MASSIMO PAPPALARDO)
22. **Mar 05/04/2022 08:30-10:30 (2:0 h)** lezione: Problema dei potenziali su reti non capacitate. Potenziali di base. Teorema di Bellman. Esercizi. (MASSIMO PAPPALARDO)
23. **Mer 06/04/2022 08:30-10:30 (2:0 h)** lezione: Il simplesso su reti. Arco entrante e ciclo. Orientamento del ciclo. Arco uscente. Il caso degenere. Il caso illimitato. Esercizi. (MASSIMO PAPPALARDO)
24. **Ven 08/04/2022 08:30-10:30 (2:0 h)** lezione: Il modello matematico del problema del flusso di costo minimo su reti capacitate. La tecnica della tripartizione degli archi. Caratterizzazioni delle basi. Unimodularità. (MASSIMO PAPPALARDO)
25. **Lun 11/04/2022 10:30-12:30 (2:0 h)** lezione: Il calcolo del flusso di base su reti capacitate. Flussi ammissibili e flussi degeneri. Esercizi. (MASSIMO PAPPALARDO)
26. **Mar 12/04/2022 08:30-10:30 (2:0 h)** lezione: Il problema dei potenziali su reti capacitate. Il calcolo del potenziale di base. Il teorema di Bellman su reti capacitate. I potenziali degeneri. Esercizi. Il problema dell'assegnamento di costo minimo come problema di flusso su reti. (MASSIMO PAPPALARDO)
27. **Mer 13/04/2022 08:30-10:30 (2:0 h)** lezione: Il simplesso per problemi di flusso di costo minimo su reti capacitate. Il cambio di base. Esercizi. (MASSIMO PAPPALARDO)
28. **Ven 22/04/2022 08:30-10:30 (2:0 h)** lezione: Prerequisiti di Analisi Matematica II per la parte di corso relativa alla Programmazione non Lineare (PNL). Gradiente, hessiana, teorema di Fermat. Condizioni necessari e/o sufficienti per minimi locali liberi attraverso il "segno" della matrice hessiana. Autovalori di matrici simmetriche. (MASSIMO PAPPALARDO)
29. **Mar 26/04/2022 08:30-10:30 (2:0 h)** lezione: Il problrema del flusso massimo. Il modello matematico di PL. La tecnica dell'arco fittizio per ricondurlo ad un problema di flusso di costo minimo su reti capacitate. Esercizi. (MASSIMO PAPPALARDO)

30. **Mer 27/04/2022 08:30-10:30 (2:0 h)** lezione: Il problema del taglio di capacità minima. Il teorema Max-Flow-Min-Cut. L'algoritmo di Ford-Fulkerson. L'algoritmo di Edmonds-Karp. Esercizi svolti. (MASSIMO PAPPALARDO)
31. **Ven 29/04/2022 08:30-10:30 (2:0 h)** lezione: L'algoritmo di Dijkstra per la ricerca dell'albero dei cammini orientti di costo minimo con costi positivi. Esercizi svolti in aula. (MASSIMO PAPPALARDO)
32. **Lun 02/05/2022 08:30-10:30 (2:0 h)** lezione: Problemi di ricapitolazione generale su PL, PLI e PLR. Problemi di trasporto come problemi di flusso su reti; problemi di cammini minimi con vincoli di budget; verifiche di ottimalità di soluzioni ammissibili di problemi di flusso massimo; soluzioni ottime non intere di problemi di cammini minimi. (MASSIMO PAPPALARDO)
33. **Mar 03/05/2022 08:30-10:30 (2:0 h)** lezione: Funzioni quadratiche e loro rappresentazione, funzioni coercive ed esistenza dei minimi globali, funzioni convesse e proprietà sui minimi locali e globali. Restrizioni a semirette. Equazioni parametriche delle semirette in $\mathbb{R}^n$. Esempi. (MASSIMO PAPPALARDO)
34. **Mer 04/05/2022 08:30-10:30 (2:0 h)** lezione: Direzioni di discesa. Condizione sufficiente per stabilire se la direzione sia di discesa. Metodo del gradiente con ricerca esatta. Teorema di convergenza. Esempi ed esercizi. (MASSIMO PAPPALARDO)
35. **Ven 06/05/2022 08:30-10:30 (2:0 h)** lezione: Regioni ammissibili della PNL come curve di livello di funzioni di classe $\mathbb{C}^2$. Domini chiusi, domini limitati, domini convessi (condizione sufficiente per la convessità), domini regolri (condizioni sufficienti per a regolarità). Esempi ed esercizi. (MASSIMO PAPPALARDO)
36. **Lun 09/05/2022 08:30-10:30 (2:0 h)** lezione: Il teorema di Lagrange-Karush-Kuhn-Tucker (LKKT) per problemi di minimo o massimo vincolati. Esempi ed esercizi. (MASSIMO PAPPALARDO)
37. **Mar 10/05/2022 08:30-10:30 (2:0 h)** lezione: Condizioni sufficienti del primo ordine per problemi convessi o concavi. Interpretazione geometrica di minimi o massimi locali. Esempi ed esercizi. (MASSIMO PAPPALARDO)
38. **Mer 11/05/2022 08:30-10:00 (2:0 h)** lezione: Il metodo di Frank-Wolfe per la minimizzazione di funzioni su poliedri limitati. Esempi ed esercizi. (MASSIMO PAPPALARDO)
39. **Ven 13/05/2022 08:30-10:30 (2:0 h)** lezione: Studio delle restrizioni per determinare se un punto stazionario è minimo locale o sella. Equazioni parametriche delle curve. Esempi ed esercizi. (MASSIMO PAPPALARDO)
40. **Lun 16/05/2022 08:30-10:30 (2:0 h)** lezione: Teorema di convergenza del metodo si Frank-Wolfe. Esercizio svolto. La lagrangiana. Studio di minimi su domini illimitati. Esercizio svolto. (MASSIMO PAPPALARDO)
41. **Mar 17/05/2022 08:30-10:30 (2:0 h)** lezione: Il metodo del gradiente proiettato. La matrice di proiezione. Criteri di stop. Esercizi svolti. (MASSIMO PAPPALARDO)
42. **Mer 18/05/2022 08:30-10:30 (2:0 h)** lezione: Teorema di convrgenza del metodo del gradiente proiettato. Cambiamenti nel caso dei problemi di massimo. La funzione "quadprog" di Matlab. Esempi ed esercizi. (MASSIMO PAPPALARDO)

Parte II

Lezioni

Capitolo 1

Lezioni di PNL

1.1 Prerequisiti Analisi II

Necessario introdurre una serie di strumenti per la PNL (*Programmazione Non Lineare*).

1.1.1 Funzioni

Le funzioni che tratteremo nella PNL sono del tipo $f : \mathbb{R}^n \to \mathbb{R}$. Notazioni possibili:

$$\begin{aligned} f(x) & \qquad\qquad \text{Dove } x \in \mathbb{R}^n \\ f(x_1, \dots, x_n) & \\ f & \end{aligned}$$

Esempio di funzione

$$f(x_1, x_2) = x_1^2 - x_2 x_1$$

dove, ad esempio, $x = (1, 1) \longrightarrow f(x) = 0$

Solito obiettivo In queste funzioni vogliamo trovare, come sempre, massimi e minimi. Ricordiamo che se sappiamo trovare il minimo sappiamo trovare il massimo (e viceversa)! Ribadiamo quanto già detto precedentemente

$$\begin{cases} \min\limits_{x \in A} f(x) = -\max\limits_{x \in A}[-f(x)] \\ \overline{x} \in \text{argmin} f(x) \Longleftrightarrow \overline{x} \in \text{argmax}[-f(x)] \end{cases}$$

Ricordandosi che argmin è l'insieme dei punti di minimo, mentre argmax è l'insieme dei punti di massimo. La considerazione fondamentale è che il punto di minimo di f è il punto di massimo di $-f$!

Attenzione Fino ad ora abbiamo considerato solo il caso particolare in cui le funzioni f sono lineari (Esempio: $3x_1 + 5x_2$), adesso generalizziamo introducendo nuove questioni.

1.1.2 Derivate

Per affrontare la PNL ci servono *derivate parziali* e le *derivate direzionali* per funzioni in $f : \mathbb{R}^n \to \mathbb{R}$. Ci servono anche i relativi concetti di *gradiente* e *matrice Hessiana*.

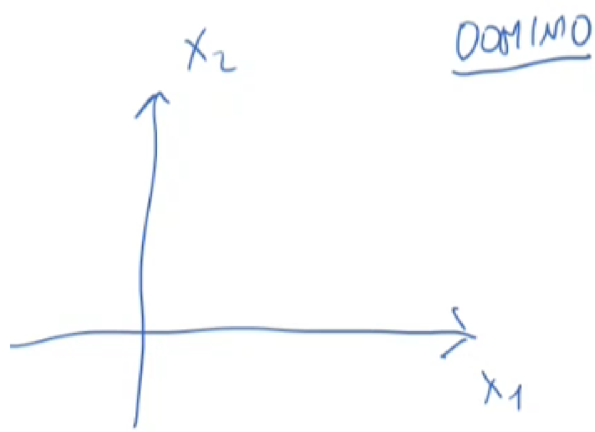
E i limiti? La cosa non ci interesserà moltissimo, visto che la stragrande maggioranza dei limiti in Analisi II non sono risolvibili.

1.1.2.1 Derivate parziali

Banalmente

$$\begin{aligned} f(x_1, x_2) &= x_1^2 + 5x_1x_2 \\ \frac{\partial f}{\partial x_1} &= 2x_1 + 5x_2 \\ \frac{\partial f}{\partial x_2} &= 5x_1 \end{aligned}$$

Cos'è la derivata parziale rispetto a x_1, graficamente parlando?



La funzione che si ottiene da f restringendomi all'asse x_1, ipotizzando che x_2 sia un valore costante. Diventa una funzione a singola variabile, e questa per noi è un'ottima notizia!

1.1.2.2 Gradiente

Il gradiente è una notazione con cui raccogliamo tutte le derivate parziali possibili relative alla funzione $f : \mathbb{R}^n \to \mathbb{R}$

$$\nabla f(x) = \left(\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n} \right)$$

ovviamente $\nabla f(x) \in \mathbb{R}^n$. Riprendendo l'esempio precedente

$$f(x_1, x_2) = x_1^2 + 5x_1x_2 \longrightarrow \nabla f(x) = (2x_1 + 5x_2, 5x_1)$$

1.1.2.3 Matrice hessiana

La matrice Hessiana $Hf(x)$ consiste in una matrice avente per componenti tutte le derivate seconde. Presenta n^2 componenti, dove n consiste nella dimensione del dominio. Abbiamo $Hf(x) \in \mathbb{R}^{n \times n}$. Vediamo il seguente esempio con $n = 2$

$$Hf(x) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(x) & \frac{\partial^2 f}{\partial x_1 x_2}(x) \\ \frac{\partial^2 f}{\partial x_2 \partial x_1}(x) & \frac{\partial^2 f}{\partial x_2^2}(x) \end{pmatrix}$$

1.1.3 Usare gradiente ed hessiana per trovare max/min

Fino ad ora non abbiamo mai fatto distinzione tra min/max locali e globali: questo perchè nelle funzioni lineari non esiste il concetto di min/max locale! Adesso, nella PNL, dobbiamo considerare pure i max/min locali. Introduciamo una serie di teoremi necessari per l'individuazione dei punti di minimo e massimo locale.

***Teorema di AN2 sulla simmetria delle matrici hessiane*.**

Dato $f : \mathbb{R}^n \to \mathbb{R}, f \in C^2$, $f \in C^\infty$, la matrice hessiana $Hf(x)$ è simmetrica!

***Teorema di AL su matrici simmetriche e autovalori*.**

Le matrici simmetriche hanno autovalori reali

Definizione di *Punto stazionario*.

Data una funzione $f : \mathbb{R}^n \to \mathbb{R}$ definisco *punto stazionario* il valore $x \in \mathbb{R}^n$ tale che $\nabla f(x) = 0$

Definizione di *Vettori linearmente indipendenti*.

Dati n vettori $x_1, \ldots, x_n$ definiamo questi linearmente indipendenti se e solo se

$$\alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n = 0$$

con $\alpha_1 = \alpha_2 = \cdots = \alpha_n = 0$. Qualunque combinazioni di vettori può essere posta uguale a zero ponendo $\alpha_1 = \cdots = \alpha_n = 0$, ma in presenza di linearità indipendente è l'unico modo per farlo (se i vettori sono linearmente dipendenti allora posso ottenere una somma nulla con alcuni coefficienti non nulli).

Definizione di *Matrice definita e semidefinita positiva e negativa*.

Definiamo quanto segue.

- Con **matrice definita positiva** si intende una matrice $A \in \mathbb{R}^{n\times n}$ tale che il seguente prodotto scalare risulta positivo $< A \cdot x, x > > 0 \ \forall x \in \mathbb{R}^n$
- Con **matrice semidefinita positiva** si intende una matrice $A \in \mathbb{R}^{n\times n}$ tale che il seguente prodotto scalare risulta positivo o nullo $< A \cdot x, x > \geq 0 \ \forall x \in \mathbb{R}^n$
- Con **matrice definita negativa** si intende una matrice $A \in \mathbb{R}^{n\times n}$ tale che il seguente prodotto scalare risulta negativo $< A \cdot x, x > < 0 \ \forall x \in \mathbb{R}^n$
- Con **matrice semidefinita negativa** si intende una matrice $A \in \mathbb{R}^{n\times n}$ tale che il seguente prodotto scalare risulta negativo o nullo $< A \cdot x, x > \leq 0 \ \forall x \in \mathbb{R}^n$

Per definire la matrice troviamo gli autovalori calcolando i valori λ nel determinante $\det(A - \lambda I) = 0$. Prendiamo ad esempio la matrice seguente, poniamo la matrice $A - \lambda I$ e calcoliamo quanto detto

$$\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \Longrightarrow \begin{pmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{pmatrix} \Longrightarrow (2-\lambda)(2-\lambda) - 1 = 0 \Longrightarrow \begin{matrix} \lambda^2 - 4\lambda + 3 = 0 \\ \lambda_1 = 3, \lambda_2 = 1 \end{matrix}$$

1.1.3.1 Teoremi per l'individuazione del minimo locale

***Teorema di Fermat per l'Analisi II*.**

Dato $f : \mathbb{R}^n \to \mathbb{R}, f \in C^2$, $f \in C^\infty$, se $\overline{x}$ è minimo locale di f allora $\overline{x}$ è un p.to stazionario. Attenzione: è CN, ma non CS:

- se ho un p.to di minimo $\overline{x}$ è automatico che questo sia p.to stazionario;
- se ho un p.to stazionario $\overline{x}$ non è automatico che questo sia p.to di minimo.

***Teorema 2*.**

Dato $f : \mathbb{R}^n \to \mathbb{R}, f \in C^2$, $f \in C^\infty$, se $\overline{x}$ è minimo locale di f e $\overline{x}$ p.to stazionario allora

$$Hf(\overline{x}) \geq 0$$

la Hessiana è *semidefinita positiva*. Attenzione: è CN, ma non CS:

- se ho p.to di minimo $\overline{x}$ è automatico che $Hf(\overline{x}) \geq 0$;
- se $Hf(\overline{x}) \geq 0$ non è automatico che $\overline{x}$ sia p.to di minimo.

***Teorema 3*.**
Dato $f : \mathbb{R}^n \to \mathbb{R}, f \in C^2$, se $\nabla f(\overline{x}) = 0$ e $Hf(\overline{x}) > 0$ (hessiana nel p.to *definita positiva*) allora $\overline{x}$ è minimo locale! Questa è CS!

1.1.3.2 Teoremi per l'individuazione del massimo locale

***Teorema di Fermat per l'Analisi II*.**
Dato $f : \mathbb{R}^n \to \mathbb{R}, f \in C^2$, $f \in C^\infty$, se $\overline{x}$ è massimo locale di f allora $\overline{x}$ è un p.to stazionario. Attenzione: è CN, ma non CS:

- se ho un p.to di massimo $\overline{x}$ è automatico che questo sia p.to stazionario;
- se ho un p.to stazionario $\overline{x}$ non è automatico che questo sia p.to di massimo.

***Teorema 2*.**
Dato $f : \mathbb{R}^n \to \mathbb{R}, f \in C^2$, $f \in C^\infty$, se $\overline{x}$ è massimo locale di f e $\nabla f(\overline{x}) = 0$ allora

$$Hf(\overline{x}) \leq 0$$

la Hessiana è *semidefinita negativa*. Attenzione: è CN, ma non CS:

- se ho p.to di massimo $\overline{x}$ è automatico che $Hf(\overline{x}) \leq 0$;
- se $Hf(\overline{x}) \leq 0$ non è automatico che $\overline{x}$ sia p.to di massimo.

***Teorema 3*.**
Dato $f : \mathbb{R}^n \to \mathbb{R}, f \in C^2$, se $\nabla f(\overline{x}) = 0$ e $Hf(\overline{x}) < 0$ (hessiana nel punto *definita negativa*) allora $\overline{x}$ è massimo locale! Questa è CS!

1.1.4 Teoremi per l'individuazione di massimi e minimi globali

Per quanto riguarda l'individuazione di massimi e minimi globali

- i primi due teoremi sono validi (se un min/max è globale è anche locale)
- non abbiamo il terzo teorema.

Le cose si fanno più complicate perchè chiaramente siamo interessati al massimo o al minimo globale.

1.1.5 Restrizioni

Recuperiamo il concetto di *restrizione*, che ci risulterà utile parlando di funzioni coercive e di selle.

Definizione di *Restrizione della funzione*.

Data una funzione $f : \mathbb{R}^n \to \mathbb{R}$ definiamo *restrizione della funzione* una funzione $\gamma : \mathbb{R} \to \mathbb{R}$ ottenuta restringendo il dominio della funzione. Nel nostro caso ciò che ci interessa è la *restrizione a una semiretta*, dove scriviamo l'equazione della semiretta in formato parametrico e successivamente sostituiamo nella funzione f. Consideriamo il seguente esempio

$$f(x_1, x_2) = x_1^2 - x_2^2$$
$$\begin{cases} x_1 = 0 + t \\ x_2 = 3 + t \end{cases} \qquad \implies \gamma(t) = t^2 - (3+t)^2 = -9 - 6t$$

Le restrizioni delle funzioni avvengono rispetto a delle curve che dobbiamo *parametrizzare*. Facciamo degli esempi di parametrizzazione!

- **Semiretta avente origine in un punto $\overline{x}$ e direzione $\overline{d}$.**

  Normalmente rappresentiamo una semiretta con la formula

  $$\overline{x} + t\overline{d}$$

  Supponiamo che $\overline{x} = (3, 2)$ e $\overline{d} = (2, 5)$. Otteniamo la seguente parametrizzazione

  $$\begin{cases} x_1 = 3 + 2t \\ x_2 = 2 + 5t \end{cases}$$

  dove $t \geq 0$

- **Semiretta lungo l'ordinata.**

  Consideriamo semirette dirette lungo l'ordinata e aventi origine nell'origine del piano cartesiano. Osserviamo subito che $x_1 = 0$, poichè siamo lungo l'ordinata. Segue

  $$\begin{cases} x_1 = 0 \\ x_2 = t \end{cases}$$

  dove $t \ldots$? Dipende da dove è diretta la semiretta!

  - Se è diretta verso l'alto abbiamo $t \geq 0$.
  - Se è diretta verso il basso abbiamo $t \leq 0$.

  Si osservi che è possibile porre anche la seguente parametrizzazione

  $$\begin{cases} x_1 = 0 \\ x_2 = -t \end{cases}$$

  dove

- $t \geq 0$ se la semiretta è diretta verso il basso;
- $t \leq 0$ se la semiretta è diretta verso l'alto.

- **Circonferenza unitaria**.

  Consideriamo una circonferenza avente raggio 1, con centro nell'origine del piano cartesiano.

  $$x_1^2 + x_2^2 = 1$$

  In questo caso abbiamo due parametrizzazioni possibili:

  - con sin e cos

    $$\begin{cases} x_1 = \cos t \\ x_2 = \sin t \end{cases}$$

    e $t \in [0, 2\pi]$. Possiamo verificare sostituendo

    $$\cos^2 t + \sin^2 t = 1$$

    che sappiamo essere valida.
  - senza sin cos.

    Pongo $x_1 = t$ e sostituisco ottenendo x_2

    $$t^2 + x_2^2 = 1 \longrightarrow x_2^2 = 1 - t^2 \longrightarrow x_2 = \sqrt{1 - t^2}$$

    quindi

    $$\begin{cases} x_1 = t \\ x_2 = \sqrt{1 - t^2} \end{cases}$$

    dove $t \in [-1, 1]$ (si studi il dominio della funzione x_2 per avere conferma). Si osservi che è equivalente parametrizzare ponendo $x_2 = t$, in quel caso otterremo

    $$\begin{cases} x_1 = \sqrt{1 - t^2} \\ x_2 = t \end{cases}$$

- **Parabola**.

  Prendiamo una semplice parabola avente vertice nell'origine e rivolta verso l'alto

  $$x_2 = x_1^2$$

  poniamo $x_1 = t$, ottenendo $x_2 = t^2$

  $$\begin{cases} x_1 = t \\ x_2 = t^2 \end{cases}$$

  Contrariamente alla circonferenza non è equivalente parametrizzare ponendo $x_2 = t$, cioè

  $$\begin{cases} x_1 = \sqrt{t} \\ x_2 = t \end{cases}$$

  dove $t \in [0, +\infty]$: stiamo parametrizzando un arco di parabola e non la parabola nel suo complesso!

1.2 Funzione *quadprog*

Questa funzione di Matlab è utilizzata per risolvere problemi di funzioni quadratiche.

$$\begin{cases} \min \frac{1}{2}xHx + c \cdot x \\ A \cdot x \leq b \\ A_{eq} \cdot x = b_{eq} \\ LB \leq x \leq UB \end{cases}$$

Il comando è il seguente:

```
>> quadprog(H, c, A, b, Aeq, beq, LB, UB);
```

H invece è la matrice Hessiana. Consideriamo il seguente esempio

$$4x_1^2 - x_1x_2 + 5x_2$$

Metteremo

$$H = 2A = 2\begin{pmatrix} 4 & -\frac{1}{2} \\ -\frac{1}{2} & 0 \end{pmatrix} = \begin{pmatrix} 8 & -1 \\ -1 & 0 \end{pmatrix} \qquad\qquad c = \begin{pmatrix} 0 & 5 \end{pmatrix}$$

1.3 Funzione *fmincon*

Questo formato è il formato utilizzato in Matlab per risolvere il seguente problema

$$\begin{cases} \min f(x) \\ A \cdot x \leq b \\ A_{eq} \cdot x = b_{eq} \\ LB \leq x \leq Ub \end{cases}$$

Il comando è il seguente

```
>> fmincon(fun, x0, A, b, Aeq, beq, LB, UB);
```

$x0$ è il punto iniziale (obv $x0 \in P$). f è la funzione, che deve essere scritta con sintassi compatibile con Matlab (Cercate online cit. - tanto non serve per lo scritto.).

INTRODUZIONE ALLA PROGRAMMAZIONE NON LINEARE

TIPOLOGIE DI PROBLEMI

ABBIAMO DUE TIPOLOGIE DI PROBLEMI:

- $f: \mathbb{R}^n \to \mathbb{R}$, PNL NON VINCOLATA
- $f: D \to \mathbb{R}$, $D \subset \mathbb{R}^n$, PNL VINCOLATA (D INSIEME CHIUSO) CONTIENE LA SUA FRONTIERA

AFFRONTEREMO PRIMA I PROBLEMI NON VINCOLATI, E SUCCESSIVAMENTE QUELLI VINCOLATI.

PROBLEMA: PUNTI DI MIN/MAX LOCALI E GLOBALI

COME SEMPRE È NOSTRO INTERESSE MINIMIZZARE / MASSIMIZZARE FUNZIONI OBIETTIVO. ABBIAMO RECUPERATO ALCUNI CONCETTI DI ANALISI II, IN PRIMIS QUELLO DI PUNTO STAZIONARIO. UN PUNTO STAZIONARIO PUÓ ESSERE:

- MINIMO LOCALE
- MINIMO GLOBALE
- MASSIMO LOCALE
- MASSIMO GLOBALE
- SELLA (IL FLESSO IN AN1)

CIÓ CHE EFFETTIVAMENTE CI INTERESSA

PER POTER APPLICARE GLI ALGORITMI DI PNL È NECESSARIO SAPER CLASSIFICARE I PUNTI STAZIONARI. A TAL PROPOSITO VIENE IN NOSTRO SOCCORSO LA HESSIANA E IL CALCOLO DEGLI AUTOVALORI

> ***Teorema 3.***
>
> Dato $f: \mathbb{R}^n \to \mathbb{R}, f \in C^2$, se $\nabla f(\overline{x}) = 0$ e $Hf(\overline{x}) > 0$ (hessiana nel p.to *definita positiva*) allora $\overline{x}$ è minimo locale! Questa è CS!

> ***Teorema 3.***
>
> Dato $f: \mathbb{R}^n \to \mathbb{R}, f \in C^2$, se $\nabla f(\overline{x}) = 0$ e $Hf(\overline{x}) < 0$ (hessiana nel punto *definita negativa*) allora $\overline{x}$ è massimo locale! Questa è CS!

TUTTAVIA LA DISTINZIONE TRA PUNTI LOCALI E GLOBALI COMPLICA LE COSE. AVREMO ULTERIORI COMPLICAZIONI IN PRESENZA DI VINCOLI: IL Th. DI FERMAT NON BASTA PIÚ.

PROBLEMA: NUMERO DI SOLUZIONI DA CONSIDERARE

PRENDIAMO LA CN DEI PUNTI STAZIONARI $\nabla f(x) = 0$

$\nabla f(x) = 0$ PUÓ ESSERE VISTO COME UN SISTEMA DI EQUAZIONI A DUE INCOGNITE IN $\mathbb{R}^2$

ES. $f(x) = 4x_1^3 - 5x_2 x_1 + 4x_2^2$

$$\nabla f(x) = \begin{pmatrix} 12x_1^2 - 5x_2 \\ -5x_1 + 8x_2 \end{pmatrix} = 0 \iff \begin{cases} 12x_1^2 - 5x_2 = 0 \\ -5x_1 + 8x_2 = 0 \end{cases}$$

IL NUM. DI SOLUZIONI NON È IMMEDIATO SE IL SISTEMA NON È LINEARE. NOI LAVORIAMO IN $\mathbb{R}^2$, MA NELLA REALTÀ SI LAVORA CON MOLTE PIÚ COMPONENTI

ALGORITMO DI ENUMERAZIONE TOTALE (INEFFICIENTE)

CONSIDERIAMO IL SEGUENTE ALGORITMO DI ENUMERAZIONE TOTALE:

1) CALCOLO IL GRADIENTE
2) PRENDO IL SISTEMA $\nabla f(x) = 0$ E TROVO TUTTE LE SOLUZIONI
3) TESTO L'HESSIANA SU TUTTE LE SOLUZIONI (COSA NON PROPRIO A COSTO ZERO) — CALCOLO DEGLI AUTOVALORI
4) PRENDO TRA I RISULTATI MINIMI (O MASSIMI) LOCALI QUELLO DA CUI SI OTTIENE VALORE MINIMO (O MASSIMO) DELLA FUNZIONE OBIETTIVO

PROMEMORIA: SIAMO NELLA PNL NON VINCOLATA!

CLASSIFICAZIONE DELLE FUNZIONI

SI INDIVIDUANO ALCUNE TIPOLOGIE DI FUNZIONI DOVE I PROBLEMI PRECEDENTI RISULTANO "SEMPLIFICATI". ABBIAMO:

- FUNZIONI QUADRATICHE;
- FUNZIONI COERCIVE;
- FUNZIONI CONVESSE.

FUNZIONI QUADRATICHE

UNA FUNZIONE $f: \mathbb{R}^m \longrightarrow \mathbb{R}$ È DETTA QUADRATICA SE È UN POLINOMIO DI SECONDO GRADO CHE PRESENTA LA SEGUENTE STRUTTURA

$$\underbrace{\langle x, Ax \rangle}_{\text{PARTE DI SECONDO GRADO}} + \underbrace{\langle c, x \rangle}_{\text{PARTE LINEARE}}$$

$n=2$ $\quad f: \mathbb{R}^2 \to \mathbb{R}$

(CI INTERESSANO POLINOMI DI 2° GRADO IN x_1, x_2)

ES. GIUSTO: $\quad 3x_1^2 + 4x_1x_2 + x_2^2 + 5x_1 - 6x_2$

ES. SBAGLIATO 1: $\quad 4x_1^3 - 5x_1x_2$ $\quad$ PRIMO TERMINE GRADO 3

ES. SBAGLIATO 2: $\quad x_1^2x_2 - 3x_2^2$ $\quad$ TERMINE MISTO GRADO 3

ES: $$Ax = \begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 + 2x_2 \\ 3x_1 + 2x_2 \end{pmatrix}$$

$$x^T(Ax) = (x_1 \; x_2) \begin{pmatrix} x_1 + 2x_2 \\ 3x_1 + 2x_2 \end{pmatrix} = x_1^2 + 2x_1x_2 + 3x_1x_2 + 2x_2^2 =$$
$$= x_1^2 + 5x_1x_2 + 2x_2^2$$

OSSERVAZIONI NON CASUALI

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix} \qquad x^T(Ax) = 1x_1^2 + 5x_1x_2 + 2x_2^2$$

$$5 = 2+3$$

I NUMERI IN BLU SONO GLI STESSI, IL 5 SI OTTIENE SOMMANDO GLI ELEMENTI EVIDENZIATI IN ROSSO.

ES2: PARTIAMO DALLA FUNZIONE QUADRATICA $f(x) = 4x_1^2 - 2x_1x_2 - 5x_2$

QUAL È LA MATRICE A?

- LA DIAGONALE È IMMEDIATA
- LA CONTRODIAGONALE NON È UNIVOCA, L'IMPORTANTE È CHE LA SOMMA DEI TERMINI SIA UGUALE AL COEFFICIENTE VICINO A x_1x_2

$$A = \begin{pmatrix} 4 & -1 \\ -1 & 5 \end{pmatrix} \dots \quad A = \begin{pmatrix} 4 & 2 \\ 0 & 5 \end{pmatrix} \dots \quad A = \begin{pmatrix} 4 & 0 \\ 2 & 5 \end{pmatrix}$$ E COSÍ VIA!!

PRENDEREMO COME RIFERIMENTO LA PRIMA, QUELLA SIMMETRICA, VISTO LA SIMMETRIA DELL'HESSIANA

$$\nabla f(x) = \begin{pmatrix} 8x_1 - 2x_2 \\ -2x_1 - 10x_2 \end{pmatrix} \Longrightarrow Hf = \begin{pmatrix} 8 & -2 \\ -2 & -10 \end{pmatrix} = 2\begin{pmatrix} 4 & -1 \\ -1 & -5 \end{pmatrix} = 2A$$

ALTRO ESEMPIO: $Q(x) = 5x_1^2 - x_1x_2 - x_2^2$

CALCOLIAMO LA HESSIANA

$$A = \begin{pmatrix} 5 & -\frac{1}{2} \\ -\frac{1}{2} & -1 \end{pmatrix} \longleftrightarrow HQ = \begin{pmatrix} 10 & -1 \\ -1 & -2 \end{pmatrix} = 2A$$

NUM. DI SOLUZIONI DEL SISTEMA $\nabla f(x) = 0$ E AUTOVALORI.

NEL CASO DI FUNZIONI QUADRATICHE IL SECONDO PROBLEMA DETTO SI SEMPLIFICA PESANTEMENTE: NEL GRADIENTE TROVIAMO FUNZIONI DI PRIMO GRADO E QUINDI

$\nabla f(x) = 0$ È UN SISTEMA LINEARE!!!

SEGUE LA HESSIANA COSTANTE E UN POLINOMIO DI SECONDO GRADO PER IL CALCOLO DEGLI AUTOVALORI!

FUNZIONI COERCIVE

UNA FUNZIONE SI DICE COERCIVA SE $\lim_{\|x\| \to +\infty} f(x) = +\infty$ — $\|x\|$ MODULO DI UN VETTORE

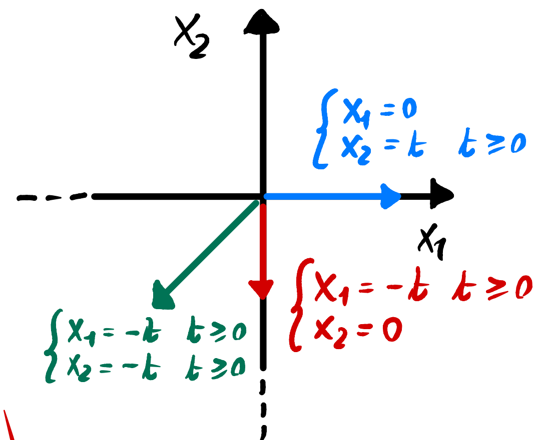
LA COSA CI PIACE POICHÉ
1) f COERCIVA $\Rightarrow \exists$ MIN. ASSOLUTO
2) $-f$ COERCIVA $\Rightarrow \exists$ MAX. ASSOLUTO (f ANTI-COERCIVA)

(f CONTINUA)

COME VEDO SE UNA FUNZIONE È COERCIVA? DEVE ANDARE ALL'INFINITO IN TUTTE LE DIREZIONI POSSIBILI

PRENDIAMO LA SEGUENTE FUNZIONE: $f(x) = x_1^3 + 5x_1x_2 + x_2^2$

NELLA FIGURA ABBIAMO RAPPRESENTATO DEGLI ESEMPI DI DIREZIONE.



PRENDIAMO I VALORI E SOSTITUIAMOLI IN $f(x)$:

$f(x) = t^2 \qquad \lim_{t \to +\infty} f(x) = +\infty$

$f(x) = -t^3$ ✗

POSSIAMO FERMARCI, FUNZIONE NON COERCIVA (SI VA A $-\infty$)

NEL CASO DELLE FUNZIONI QUADRATICHE SI HA UNA FUNZIONE COERCIVA SE LA MATRICE A È DEFINITA POSITIVA.

ES: $f(x) = 2x_1^2 + 3x_1x_2 + 4x_2^2 - x_1 + 5x_2$

$$A = \begin{pmatrix} 2 & \frac{3}{2} \\ \frac{3}{2} & 4 \end{pmatrix} \qquad A - \lambda I = \begin{pmatrix} 2-\lambda & \frac{3}{2} \\ \frac{3}{2} & 4-\lambda \end{pmatrix}$$

$$\det(A - \lambda I) = (2-\lambda)(4-\lambda) - \frac{9}{4} = 0$$

$$8 - 2\lambda - 4\lambda + \lambda^2 - \frac{9}{4} = 0$$

$$\lambda^2 - 6\lambda + \frac{23}{4} = 0$$

$$\Delta = 36 - 4(1)\left(\frac{23}{4}\right) = 13$$

$$\lambda_1 = \frac{6 + \sqrt{13}}{2} \qquad \lambda_2 = \frac{6 - \sqrt{13}}{2}$$

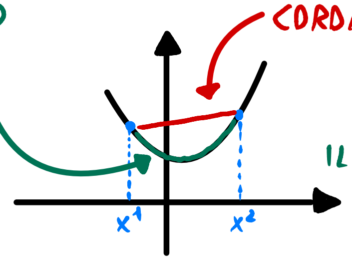
ENTRAMBI POSITIVI!

FUNZIONI CONVESSE

UNA FUNZIONE SI DICE CONVESSA SE

$$\underbrace{f(\lambda x^1+(1-\lambda)x^2)}_{\text{GRAFICO}} \leq \underbrace{\lambda f(x^1)+(1-\lambda)f(x^2)}_{\text{CORDA}} \quad \forall \lambda \in [0,1],\ \forall x^1,x^2 \in \mathbb{R}^n$$

ESEMPIO:



IL GRAFICO STA SEMPRE SOTTO LA CORDA

COME VERIFICO SE $f(x)$ É CONVESSA?

FUNZIONE CONVESSA SE E SOLO SE HESSIANA SEMIDEFINITA POSITIVA $Hf(x) \geq 0 \ \forall x$

DEVO CALCOLARMI TUTTE LE HESSIANE!!! MI SALVO CON LE FUNZIONI QUADRATICHE (HESSIANA COSTANTE)

PERCHÉ CI PIACCIONO LE FUNZIONI CONVESSE? PER QUESTO TEOREMA:

TEOREMA SUI P.TI STAZIONARI IN FUNZIONI CONVESSE. I P.TI STAZIONARI DI FUNZIONI CONVESSE, SE ESISTONO, SONO MINIMI ASSOLUTI.

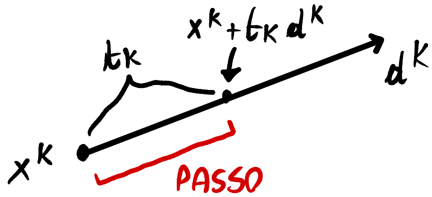
ALGORITMI ITERATIVI PER LA RISOLUZIONE

L'ALGORITMO DI ENUMERAZIONE GIÀ VISTO È BUONO IN $\mathbb{R}^2$ E CON FUNZIONI QUADRATICHE. IN ALTRI CASI CONVIENE RIVOLGERSI AD ALGORITMI ITERATIVI.

DATO UN PUNTO INIZIALE x^0 SI COSTRUISCE ITERATIVAMENTE LA SEGUENTE SUCCESSIONE:

$$x^{k+1} = x^k + t_k d^k$$

OGNI VOLTA PRENDO IL P.TO PRECEDENTE E LO PONGO COME ORIGINE DI UNA SEMIRETTA AVENTE COME DIREZIONE IL VERSORE d^k



NEI VARI ALGORITMI DISCUTEREMO:
- LA DIREZIONE d^k, E
- QUANTO SPOSTARCI LUNGO QUELLA DIREZIONE (t_k, "STEPSIZE" - DIMENSIONE DEL PASSO)

ABBIAMO INTRODOTTO IL CONCETTO DI RESTRIZIONE: $\varphi(t) \triangleq f(x^k + t\, d^k)$
(DEF. NELLE PREMESSE ALLA PNL) $\varphi : \mathbb{R}^+ \longrightarrow \mathbb{R}$

IL 90% DEGLI ALGORITMI CHE VEDREMO SONO DETTI "METODI DI DISCESA", DOVE LE MIGLIORI DIREZIONI d^k SONO QUELLE TALI CHE

$$f(x^{k+1}) < f(x^k)$$

MATEMATICAMENTE PARLANDO POSSIAMO VEDERE SE φ SCENDE A PARTIRE DALLA DERIVATA $\varphi'(0)$, IMPONENDO $\varphi'(0) < 0$

CALCOLO LA DERIVATA DI UNA COMPOSIZIONE: $\underbrace{\varphi'(t) = \nabla f(x^k + t d^k)\, d^k}_{\text{DIFF. DELLA FUNZIONE COMPOSTA}}$
(DA ANALISI II)

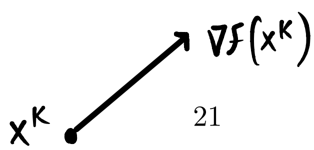
QUINDI:

$$\varphi'(0) = \nabla f(x^k) \cdot d^k < 0$$

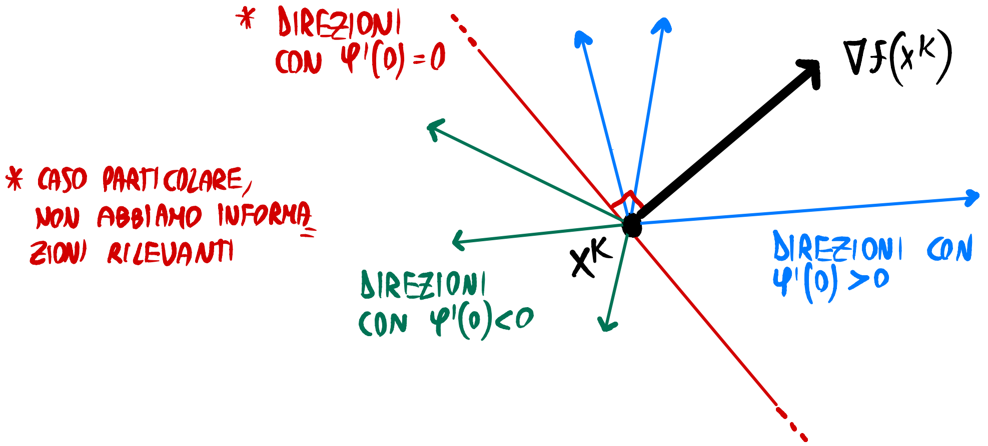
SCELGO DIREZIONE TALE DA AVERE PRODOTTO SCALARE COL GRADIENTE NEGATIVO

INTERPRETIAMO GRAFICAMENTE:

SUPPONIAMO DI AVERE x^k E DI CALCOLARE IL GRADIENTE $\nabla f(x^k)$: OTTENIAMO UNA DIREZIONE.



SAPPIAMO CHE IL PRODOTTO SCALARE È $\varphi'(0) = \|\nabla f(x^k)\| \cdot \|d^k\| \cos\theta^k$
QUINDI IL SEGNO DI $\varphi'(0)$ DIPENDE DA $\cos\theta^k$ E DALL'ANGOLO θ^k



IMMAGINARSI LA DIREZIONE $\nabla f(x^k)$ COME L'ASSE X DI UN PIANO CARTESIANO, QUINDI I QUATTRO QUADRANTI:

- NEL I° E NEL IV° SIAMO IN SALITA
- NEL II° E NEL III° SIAMO IN DISCESA

ESEMPIO: $f(x_1, x_2) = x_1^2 - x_1 x_2 + 3x_2^2 \qquad x^k = (1,1) \qquad d^k = (1,2)$

LA FUNZIONE SCENDE O SALE?

$\nabla f(x_1, x_2) = (2x_1 - x_2, -x_1 + 6x_2) \implies \nabla f(1,1) = (1,5)$

$\implies \varphi'(0) = (1,5)\cdot(1,2) = 11 > 0$ LA FUNZIONE CRESCE

NON VOGLIO STARE A PROVARE PRODOTTI SCALARI: VOGLIO UN d^k CHE SONO CERTO A PRIORI MI DIA PRODOTTO SCALARE NEGATIVO

$\nabla f(x^k) \cdot (-\nabla f(x^k)) = -\|\nabla f(x^k)\|^2$

RIPRENDIAMO L'ESEMPIO

$\nabla f(1,1) = (1,5) \implies (1,5)\cdot(-1,-5) = -26$ ✓

RISULTATO SEMPRE NEGATIVO SALVO GRADIENTE NULLO (A QUEL PUNTO SIAMO IN PRESENZA DI UN PUNTO STAZIONARIO, CHE PONIAMO DA PARTE).

CONFRONTO CON GLI ARGOMENTI PASSATI E GLI ALGORITMI DELLA PNL

- A COSA CONVERGONO I METODI? (CORRETTEZZA)
  PRIMA: MINIMO GLOBALE

  ORA: PUNTI STAZIONARI (MIN. LOCALE, MIN. GLOBALE, SELLA)
- QUANTI PASSI SONO NECESSARI? (TERMINAZIONE)
  PRIMA: NUMERO FINITO DI STEP

  ORA: NUMERO INFINITO DI STEP (!!!)

REGOLE DI STOP

CI SERVONO DELLE REGOLE PER GESTIRE IL NUMERO INFINITO DI STEP.

1) DETERMINO UN NUM. DI STEP MASSIMO (ES: MI FERMO DOPO 10^5 PASSI)
2) IMPOSTO UN TEMPO MASSIMO
3) $|f(x^{k+1}) - f(x^k)| < \Delta$ (ES: $\Delta = 10^{-4}$, SE DOPO TOT PASSI LA DIFF. È $< 10^{-4}$ ALLORA MI FERMO)
4) $|\nabla f(x^k)| < \Delta$ (ES: $\Delta = 10^{-4}$, COME PRIMA... SE PER TOT PASSI NON MI AVVICINO ALLO ZERO MI FERMO)

METODO DEL GRADIENTE (CON RICERCA ESATTA DEL PASSO)

SPIEGAZIONE

METODO CLASSICO PER OTTIMIZZARE UNA FUNZIONE NON LINEARE, DI n VARIABILI E ALMENO CLASSE C^1.

È UN ESEMPIO DI METODO ITERATIVO, DETTO ANCHE "METODO DELLA DISCESA PIÙ RIPIDA"

$x^{k+1} = x^k + t\, d^k$ DOVE $d^k = -\nabla f(x^k)$

PASSO $\rightarrow$ $t_k = \underset{t \geq 0}{\arg\min}\, \varphi(t)$] MINIMO DELLA φ, SCENDO IL PIÙ POSSIBILE LUNGO LA DIREZIONE

$\varphi(t)$ È UNA FUNZIONE A UNA VARIABILE! NOI SAPPIAMO LAVORARE AGILE CON QUESTE FUNZIONI!

FUNZIONE QUADRATICA $\longrightarrow$ $\varphi(t)$ È UNA PARABOLA

FUNZIONE CONVESSA $\longrightarrow$ P.TO STAZIONARIO SICURAMENTE MINIMO ASSOLUTO (CERTEZZA DELLA SUA ESISTENZA SE f COERCIVA)

TEOREMA DI CONVERGENZA (MET. DEL GRADIENTE)

SIA f FUNZIONE COERCIVA, LA SUCCESSIONE DEL GRADIENTE CON RICERCA ESATTA $\{x^k\}$

- TERMINA CON UN NUM. FINITO DI PASSI IN UN P.TO STAZIONARIO, O
- I SUOI PUNTI DI ACCUMULAZIONE SONO P.TI STAZIONARI.

IN REALTÁ CI SERVE POCO IL TEOREMA: APPLICO EURISTICAMENTE I CRITERI DI STOP.

ESEMPIO DI PASSO DEL METODO (1)

$$f(x_1, x_2) = 2x_1^2 - 2x_1x_2 + 3x_2^2 + x_1$$

RICERCA DEL MINIMO

CONSIDERO COME PUNTO INIZIALE $x^K = (1,0)$

1) CALCOLO IL GRADIENTE

$$\nabla f(x_1, x_2) = (4x_1 - 2x_2 + 1, -2x_1 + 6x_2) \qquad \nabla f(1,0) = (4-0+1, -2+0) = (5,-2)$$

2) TROVO d^K

$$d^K = -\nabla f(1,0) = (-5,2)$$

3) CALCOLO $\varphi(t)$ RICORDANDO CHE $\varphi(t) \triangleq f(x^K + t d^K)$

$$\varphi(t) = f\big((1,0) + t(-5,2)\big) = f(1-5t, 2t) =$$
$$= 2(1-5t)^2 - 2(2t)(1-5t) + 12t^2 + 1 - 5t =$$
$$= 2(1+25t^2-10t) - 4t + 20t^2 + 12t^2 + 1 - 5t =$$
$$= 82t^2 - 29t + 3$$

4) TROVO $\underset{t>0}{\operatorname{argmin}}\, \varphi(t)$

$\varphi(t)$ É UNA PARABOLA RIVOLTA VERSO L'ALTO CON VERTICE $-\frac{b}{2a} = \frac{29}{164}$

QUESTO É argmin

5) TROVO x^{K+1}

$$x^{K+1} = (1,0) + \frac{29}{164}(-5,2) = \left(1 - \frac{145}{164}, \frac{68}{164}\right) = \left(\frac{19}{164}, \frac{58}{164}\right)$$

ESEMPIO DI PASSO DEL METODO (2)

$f(x_1, x_2) = -x_1^2 + x_1 x_2 - x_2^2$

RICERCA DEL MASSIMO

CONSIDERO COME PUNTO INIZIALE $x^K = (1,0)$

1) CALCOLO IL GRADIENTE

$$\nabla f(x_1, x_2) = (-2x_1 + x_2,\ x_1 - 2x_2) \qquad \nabla f(1,0) = (-2, 1)$$

2) TROVO d^K

$$d^K = -\nabla f(1,0) = (1,-2)$$

3) CALCOLO $\varphi(t)$ RICORDANDO CHE $\varphi(t) \triangleq f(x^K + t\, d^K)$

$$\begin{aligned}\varphi(t) = f\big((1,0) + t(-2,1)\big) &= f(1-2t,\ t) = \\ &= -(1-2t)^2 + t(1-2t) - t^2 \\ &= -1 + 4t - 4t^2 + t - 2t^2 - t^2 = \\ &= -7t^2 + 5t - 1\end{aligned}$$

4) TROVO $\operatorname*{argmax}_{t>0} \varphi(t)$

$\varphi(t)$ È UNA PARABOLA RIVOLTA VERSO IL BASSO CON VERTICE $-\frac{b}{2a} = \frac{5}{14}$

QUESTO È argmax

5) TROVO x^{K+1}

$$x^{K+1} = (1,0) + \frac{5}{14}(-2,1) = \left(\frac{4}{14}, \frac{5}{14}\right)$$

DOMINI DELLA PNL (o REGIONE AMMISSIBILE)

DEFINIZIONE DI DOMINIO

INIZIAMO A INTRODURRE VINCOLI TRATTANDO UN DOMINIO $D \subseteq \mathbb{R}^n$

$$f : D \subseteq \mathbb{R}^n \to \mathbb{R}$$

DOVE f È LA FUNZIONE OBIETTIVO. DEFINIAMO MATEMATICAMENTE D

$$D = \{ x \in \mathbb{R}^n : g_1(x) \leq 0, \dots, g_m(x) \leq 0, h_1(x) = 0, \dots, h_p(x) = 0 \}$$

DOVE m È IL NUMERO DI DISUGUAGLIANZE, p IL NUMERO DI UGUAGLIANZE.

$$g : \mathbb{R}^n \to \mathbb{R}^m \qquad g = (g_1, \dots, g_m)$$

$$h : \mathbb{R}^n \to \mathbb{R}^p \qquad h = (h_1, \dots, h_p)$$

NELLA DEFINIZIONE RIENTRANO ANCHE LE REGIONI AMMISSIBILI VISTE NEGLI ARGOMENTI PRECEDENTI.

ES: $\begin{cases} \max \ 3x_1 + 5x_2 \\ 4x_1 + 6x_2 \leq 7 \\ 3x_1 - 2x_2 = 8 \\ 4x_1 + 2x_2 = 9 \\ x_1 \leq 0 \\ x_1^2 + x_2^2 \leq 5 \end{cases}$ ← FUNZIONE NON LINEARE !!!

$$\left.\begin{array}{l} g_1(x_1, x_2) = 4x_1 + 6x_2 - 7 \\ g_2(x_1, x_2) = x_1 \\ g_3(x_1, x_2) = x_1^2 + x_2^2 - 5 \end{array}\right] m = 3$$

$$\left.\begin{array}{l} h_1(x_1, x_2) = 3x_1 - 2x_2 - 8 \\ h_2(x_1, x_2) = 4x_1 + 2x_2 - 9 \end{array}\right] p = 2$$

ES2: $\begin{cases} \min \ x_1^2 - x_2^2 \\ 5x_1 + 6x_2 \leq 8 \\ 4x_1 + 3x_2 \leq 9 \\ 3x_1 + 2x_2 \leq 10 \end{cases}$ ← FUNZIONE NON LINEARE !!!

$$\left.\begin{array}{l} g_1(x_1, x_2) = 5x_1 + 6x_2 - 8 \\ g_2(x_1, x_2) = 4x_1 + 3x_2 - 9 \\ g_3(x_1, x_2) = 3x_1 + 2x_2 - 10 \end{array}\right] \begin{array}{l} m = 3 \\ p = 0 \end{array}$$

CONFRONTO CON LE FUNZIONI NON VINCOLATE

FINO ALLA PAGINA PRECEDENTE ABBIAMO DISCUSSO SUL COME OTTIMIZZARE FUNZIONI NON LINEARI PRIVE DI VINCOLI.

DELLE PRECEDENTI DISCUSSIONI RECUPERIAMO:
- LA CLASSIFICAZIONE DELLE FUNZIONI (COERCIVE, CONVESSE, QUADRATICHE)
- LE C.N. E LE C.S. (I "TRE TEOREMI", TARGET E TEST)

C.N. C.S.

SUI TEOREMI OSSERVIAMO CHE:
- NELLA REGIONE INTERNA SONO SEMPRE VALIDI
- LUNGO LA FRONTIERA NO (IN PRESENZA DI VINCOLI NON È DETTO $\nabla f(x) = 0$)

CI SERVONO DEGLI ACCORGIMENTI PER LAVORARE SULLA FRONTIERA.

PROPRIETÀ DEL DOMINIO CHE CI INTERESSANO

INSIEME CHIUSO O APERTO?

NON CONSIDERARE LA FRONTIERA (INSIEME APERTO) COMPLICA LE COSE. FORTUNATAMENTE LAVOREREMO SEMPRE CON DOMINI CHIUSI.

RICORDARSI LA DEF. DI FRONTIERA DA ANI

SE CHIUSO E LIMITATO VALE WEIERSTRASS

INSIEME LIMITATO O NO?

AL 99% DEI CASI LA REGIONE È LIMITATA. NON CI SERVE SAPERE SE f È COERCIVA IN QUESTI CASI (SOLO IN PRESENZA DI REGIONI ILLIMITATE).

COME VEDO SE UN P.TO È INTERNO O SULLA FRONTIERA?

SOSTITUISCO NEI VINCOLI, E VERIFICO CHE SIA RISPETTATO IL MINORE STRETTO. SE QUALCHE VINCOLO È RISPETTATO CON L'UGUALE SIAMO SULLA FRONTIERA.

INSIEME CONVESSO O NO?

VARIE CONDIZIONI DA CUI ATTINGERE. A NOI INTERESSANO LE SEGUENTI:

1) g_i CONVESSA $\forall i$
2) h_j LINEARE $\forall j$

$\underset{(c.s.)}{\Longrightarrow}$ INSIEME CONVESSO

INSIEME REGOLARE O NO?

VARIE CLASSI DI INSIEMI REGOLARI. CI INTERESSANO LE SEGUENTI:

1) POLIEDRI (g, h LINEARI)

2) $\overbrace{g \text{ CONVESSO, } h \text{ LINEARE,}}^{\text{INSIEME CONVESSO}}$ $\exists \bar{x} : g(\bar{x}) < 0$ (CONDIZIONE DI SLATER)

ESISTE ALMENO UN PUNTO INTERNO

3) $\nabla g_i(\bar{x}), \nabla h_j(\bar{x})$ $\forall i$ $\forall j$ ATTIVI IN $\bar{x}$ SONO LINEARMENTE INDIPENDENTI $\forall \bar{x}$ AMMISSIBILE

VINCOLI SODDISFATTI CON L'UGUALE

(CONDIZIONE DI MANGASARIAN)

STEP DA SEGUIRE:

a) PRENDO UN PUNTO
b) VEDO QUALI SONO I VINCOLI ATTIVI
c) VERIFICO IN QUESTI PUNTI SE I GRADIENTI SONO LINEARMENTE INDIPENDENTI

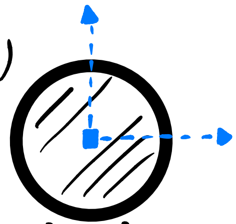
$\downarrow$

TROVARE UNA SOTTOMATRICE $m \times m$ INVERTIBILE ($\det \neq 0$) (RANGO m)

ESEMPI INTRODUTTIVI SULLE PROPRIETÁ DEI DOMINI

CONSIDERIAMO I SEGUENTI DOMINI: COSA POSSIAMO DIRE?

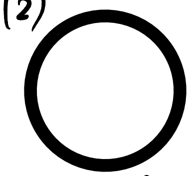
(1)



$x_1^2 + x_2^2 \leq 1$

ENTRAMBI CHIUSI E LIMITATI

(2)



$x_1^2 + x_2^2 = 1$

(1) $g_1 = x_1^2 + x_2^2 - 1 \quad g_1 \leq 0$

$$H_g = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \longrightarrow (2-\lambda)(2-\lambda) = 0$$

$$4 - 2\lambda - 2\lambda + \lambda^2 = 0$$

$$\lambda^2 - 4\lambda + 4 = 0$$

$$\Delta = 16 - 4(1)(4) = 0$$

$$\frac{4}{2} \longrightarrow \lambda_1 = \lambda_2 = +2$$

g CONVESSO $\Downarrow$

g CONVESSO, h NON C'É — DOMINIO CONVESSO

VALE LA CONDIZIONE DI SLATER: P.TO INTERNO (0,0) — DOMINIO REGOLARE PER SLATER

(2) $h_1 = x_1^2 + x_2^2 - 1 \quad h_1 = 0$

h NON LINEARE

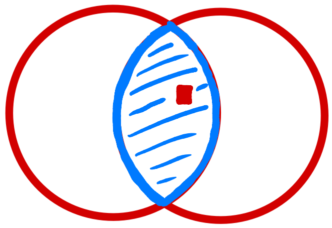
— DOMINIO NON CONVESSO

- SICURAMENTE NON É POLIEDRO
- SE NON É CONVESSO NON POSSO APPLICARE SLATER
- MANGASARIAN?
  - a) PRENDO UN PUNTO DELLA CIRCONFERENZA
  - b) UNICO VINCOLO SEMPRE ATTIVO
  - c) $\nabla h(x) = (2x_1, 2x_2)$ SEMPRE LINEARMENTE INDIPENDENTE DA SE STESSO

((0,0) NON É PARTE DELLA CIRCONFERENZA)

— DOMINIO REGOLARE PER MANGASARIAN

ULTERIORE ESEMPIO CON DUE VINCOLI g_1, g_2



NELLA REGIONE IN ~~AZZ~~ AZZURRO VALGONO ENTRAMBI I VINCOLI.

- É CONVESSA? GRAFICAMENTE É CHIARO. IN ASSENZA DI GRAFICO CALCOLO LE HESSIANE. ENTRAMBE RISULTERANNO SEMIDEFINITE POSITIVE, QUINDI IL DOMINIO É CONVESSO.
- É REGOLARE?

1) SICURAMENTE NON ABBIAMO UN POLIEDRO (g_1, g_2 NON LINEARI)
2) IL DOMINIO É CONVESSO, E SI INDIVIDUA AGILE UN PUNTO INTERNO. DOMINIO REGOLARE! (POSSIAMO FERMARCI QUA)
3) PRENDIAMO IL P.TO INTERNO: ABBIAMO $g_1 < 0$, $g_2 < 0$

QUINDI NESSUN VINCOLO ATTIVO $\Longrightarrow$ MANGASARIAN AUTOMATICAMENTE SODDISFATTO

CONDIZIONE DI OTTIMALITÁ PER PROBLEMI VINCOLATI

ABBIAMO DETTO CHE IN PROBLEMI DI PNL VINCOLATI É IMPROPRIO AFFERMARE CHE IN PRESENZA DI MIN/MAX $\bar{x}$ SI HA $\nabla f(\bar{x}) = 0$ (TH. DI FERMAT)

CI SERVE UNA C.N. ALTERNATIVA PER POTER GESTIRE I PUNTI DI MIN/MAX CHE SI TROVANO LUNGO LA FRONTIERA.

SI CONSIDERI A TAL PROPOSITO IL SEGUENTE TEOREMA

TEOREMA (LAGRANGE – KARUSH – KUHN – TUCKER, LKKT)

ENUNCIATO. SIA (P) IL SEGUENTE PROBLEMA

$$(P) = \begin{cases} \min/\max f(x) \\ g_1(x) \le 0 \\ g_2(x) \le 0 \\ \vdots \\ g_m(x) \le 0 \\ h_1(x) = 0 \\ \vdots \\ h_p(x) = 0 \end{cases}$$

$f, g_1, \dots, g_m, h_1, \dots, h_p : \mathbb{R}^n \to \mathbb{R}$

TUTTE DI CLASSE C^2

$D = \{ x \in \mathbb{R}^n : g(x) \le 0,\ h(x) = 0 \}$

$\hookrightarrow$ INSIEME REGOLARE

CONDIZIONI NECESSARIE DI OTTIMALITÁ

i) SIA $\bar{x}$ MINIMO LOCALE PER (P). ALLORA $\exists \bar{\lambda} = (\bar{\lambda}_1, \dots, \bar{\lambda}_m) \in \mathbb{R}^m_+$ $\exists \bar{\mu} = (\bar{\mu}_1, \dots, \bar{\mu}_p)$, TALI CHE

$$\begin{cases} \nabla f(\bar{x}) + \sum_{i=1}^{m} \bar{\lambda}_i \nabla g_i(\bar{x}) + \sum_{j=1}^{p} \bar{\mu}_j \nabla h_j(\bar{x}) = 0 \\ \langle \bar{\lambda}, g(\bar{x}) \rangle = 0 \\ h(\bar{x}) = 0 \end{cases}$$

[SISTEMA LKKT]

[$g(x) \le 0$ NON SERVE POICHÉ $\bar{x}$ AMMISSIBILE]

ii) SIA $\bar{x}$ MASSIMO LOCALE PER (P). ALLORA $\exists \bar{\lambda} = (\bar{\lambda}_1, \dots, \bar{\lambda}_m) \in \mathbb{R}^m_-$ $\exists \bar{\mu} = (\bar{\mu}_1, \dots, \bar{\mu}_p)$, TALI CHE

$$\begin{cases} \nabla f(\bar{x}) + \sum_{i=1}^{m} \bar{\lambda}_i \nabla g_i(\bar{x}) + \sum_{j=1}^{p} \bar{\mu}_j \nabla h_j(\bar{x}) = 0 \\ \langle \bar{\lambda}, g(\bar{x}) \rangle = 0 \\ h(\bar{x}) = 0 \end{cases}$$

[SISTEMA LKKT]

[$g(x) \le 0$ NON SERVE POICHÉ $\bar{x}$ AMMISSIBILE]

DOVE $\nabla_X L = \nabla f(\bar{x}) + \sum_{i=1}^{m} \bar{\lambda}_i \nabla g_i(\bar{x}) + \sum_{j=1}^{p} \bar{\mu}_j \nabla h_j(\bar{x})$ É IL GRADIENTE, COMPONENTE RISPETTO AD X, DELLA FUNZIONE LAGRANGIANA

$$L(\underset{\mathbb{R}^n}{x}, \underset{\mathbb{R}^m}{\lambda}, \underset{\mathbb{R}^p}{\mu}) = f(x) + \lambda_1 g_1(x) + \dots + \lambda_m g_m(x) + \mu_1 h_1(x) + \dots + \mu_p h_p(x)$$

PARALLELISMO

$\nabla f = 0$ FERMAT

$\nabla_X L = 0$ LAGRANGIANA, LKKT

SPIEGAZIONE

I MINIMI/MASSIMI LOCALI SONO SOLUZIONI DEL SISTEMA LKKT (C.N.) LE VARIABILI IN CUI SI RISOLVE IL PROBLEMA SONO $(\bar{x}, \bar{\lambda}, \bar{\mu})$: PRENDO TUTTE LE SOLUZIONI POSSIBILI, TRA QUESTE TROVEREMO IL MINIMO/MASSIMO IN $\bar{x}$ (OVVIAMENTE SE ESISTE IL MINIMO/MASSIMO).

- DETTAGLIO: IL SISTEMA NON É DETTO SIA LINEARE.

  POSSIBILE AVERE UN SISTEMA PURAMENTE LINEARE SOLO CON UN PROBLEMA DI PL. PREFERITA COMBO $\underbrace{\text{QUADRATICA + POLIEDRO}}_{\text{quadprog}}$.

- QUANTE EQUAZIONI STIAMO CONSIDERANDO?

  1) PER QUANTO RIGUARDA $h(\bar{x}) = 0$ ABBIAMO p EQUAZIONI.

  2) PER QUANTO RIGUARDA $\langle \lambda, g(\bar{x}) \rangle = 0$

     - I λ SONO TUTTI POSITIVI
     - I $g_K(\bar{x})$ SONO TUTTI ≤ 0

     $\Longrightarrow$ TUTTI GLI ADDENDI SONO NEGATIVI

     SE LA SOMMA É NULLA ALLORA TUTTI GLI ADDENDI SONO NULLI.

     $\Downarrow$

     $$\begin{cases} \bar{\lambda}_1 g_1(\bar{x}) = 0 \\ \vdots \\ \bar{\lambda}_m g_m(\bar{x}) = 0 \end{cases}$$

     ABBIAMO m EQUAZIONI

  3) PER QUANTO RIGUARDA $\underset{\mathbb{R}^n}{\nabla f(\bar{x})} + \sum_{i=1}^{m} \underset{\mathbb{R}^n}{\bar{\lambda}_i \nabla g_i(\bar{x})} + \sum_{j=1}^{p} \underset{\mathbb{R}^n}{\bar{\mu}_j \nabla h_j(\bar{x})} = 0$

     ABBIAMO n EQUAZIONI

ABBIAMO, IN CONCLUSIONE, UN SISTEMA DI $m+n+p$ EQUAZIONI IN $m+n+p$ INCOGNITE (UN SISTEMONE, CIT.).

- IL SISTEMA È LETTERALMENTE UN PANIERE: IN BASE AL SEGNO DEI MOLTIPLICATORI λ PRENDO I MINIMI O I MASSIMI. PRENDIAMO COME ESEMPIO UN PROBLEMA CON $n=2$, 2 DISUGUAGLIANZE E 1 UGUAGLIANZA.

PROPOSTE DI SOLUZIONI $(x_1, x_2, \lambda_1, \lambda_2, \mu_1)$

$(2,0,\underbrace{3,1},-5)$ MINIMO LOCALE?

$(3,1,\underbrace{5,-2},4)$ SEGNI DISCORDI, SCARTO

$(4,0,\underbrace{-3,-2},0)$ MASSIMO LOCALE?

INCUBO: E SE AVESSI UNA SELLA? SU MIN E MAX POTREBBERO ESSERCI DUBBI (RICORDARSI CHE È UNA C.N., NON UNA C.N.S.)

SUPPONIAMO CHE LA REGIONE AMMISSIBILE SIA LIMITATA: A QUEL PUNTO SIAMO A POSTO, WEIERSTRASS CI DA CERTEZZE SULL'ESISTENZA DI MAX/MIN GLOBALI. LE SOLUZIONI INDICATE SONO EFFETTIVAMENTE MIN/MAX LOCALI. (SENZA RIMANIAMO NEL DUBBIO)

↓ ADDIRITTURA GLOBALI

TEOREMA COMPLEMENTARE A LKKT (C.S.)

SIA f CONVESSA E D CONVESSO, DOVE $D=\{x : g(x)\leq 0,\ h(x)=0\}$ È REGOLARE, INOLTRE $f, g, h \in C^2$. SIA $(\bar{x}, \bar{\lambda}, \bar{\mu})$ UNA SOLUZIONE DEL SISTEMA LKKT.
ALLORA, SE $\bar{\lambda} \geq 0$ POSSIAMO DIRE CHE $\bar{x}$ È MINIMO GLOBALE.

↑ MOLTIPLICATORI TUTTI POSITIVI

SIA f CONCAVA E D CONVESSO (IBIDEM SULLE CARATTERISTICHE DI D), SIA $(\bar{x}, \bar{\lambda}, \bar{\mu})$ UNA SOLUZIONE DEL SISTEMA LKKT.
ALLORA, SE $\lambda \leq 0$ POSSIAMO DIRE CHE $\bar{x}$ È MASSIMO GLOBALE.

↑ MOLTIPLICATORI TUTTI NEGATIVI

NELLO SVOLGIMENTO DEGLI ESERCIZI SI UTILIZZI LO SCHEMA IN FONDO AL CAPITOLO DELLA DISPENSA

PRIMO ESEMPIO DI STUDIO

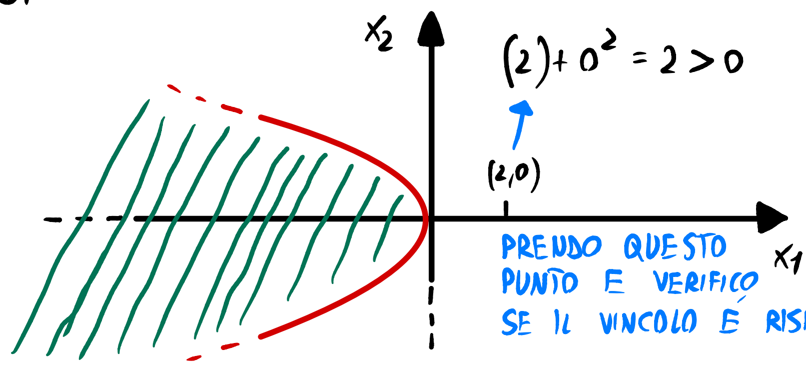
CONSIDERIAMO IL SEGUENTE PROBLEMA

$$\begin{cases} f = x_1^2 + 2x_1 + x_2^2 + 2x_2 \\ g_1 = x_1 + x_2^2 \\ g_2 = x_1^2 - 4 \end{cases}$$

IL DOMINIO D È REGOLARE? RISOLVIAMO GRAFICAMENTE

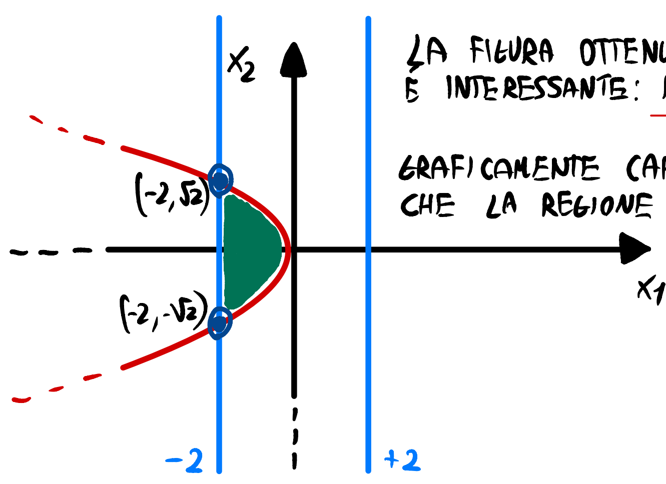
DISEGNO I VINCOLI DI DISUGUAGLIANZA K-ESIMO PONENDO $g_k(x) = 0$ (SI DISEGNA LA FRONTIERA), SUCCESSIVAMENTE INDIVIDUO IL SEMIPIANO DA CONSIDERARE VERIFICANDO SE CON ALCUNI PUNTI IL VINCOLO È RISPETTATO.

g_1 È UNA PARABOLA "ROVESCIATA": $x_1 + x_2^2 = 0 \longrightarrow x_1 = -x_2^2$



VINCOLO NON RISPETTATO, LA REGIONE DA CONSIDERARE È L'ALTRO SEMIPIANO.

g_2 : $x^2 - 4 \leq 0 \longrightarrow x^2 \leq 4 \longrightarrow -2 \leq x \leq 2$



LA FIGURA OTTENUTA È INTERESSANTE: REGIONE LIMITATA!

$\exists$ MIN. ASSOLUTO

$\exists$ MAX. ASSOLUTO

GRAFICAMENTE CAPIAMO ANCHE CHE LA REGIONE È CONVESSA.

PROVIAMO A RIBADIRE QUANTO VISTO SENZA GRAFICI

1) POLIEDRI (g, h LINEARI)

g, h NON SONO LINEARI $\implies$ NON HO UN POLIEDRO

INSIEME CONVESSO

2) g CONVESSO, h LINEARE, $\exists \bar{x}: g(\bar{x})<0$ (CONDIZIONE DI SLATER)

ESISTE ALMENO UN PUNTO INTERNO

MEGLIO SAPER DISEGNARE

h NON PRESENTE $\longrightarrow$ NON È UN PROBLEMA!

PARTE INTERNA DI UNA PARABOLA PALESEMENTE CONVESSA
PRESENTE P.TO INTERNO: SLATER SODDISFATTO!

3) $\nabla g_i(\bar{x}), \nabla h_j(\bar{x})$ $\forall i$ $\forall j$ ATTIVI IN $\bar{x}$ SONO LINEARMENTE INDIPENDENTI $\forall \bar{x}$ AMMISSIBILE

VINCOLI SODDISFATTI CON L'UGUALE

(CONDIZIONE DI MANGASARIAN)

IL DOMINIO È SENZ'ALTRO REGOLARE, MA SUPPONIAMO DI NON ACCORGERCI CHE LA FIGURA È CONVESSA, ANDANDO DIRETTI A MANGASARIAN

LA CONDIZIONE RICHIEDE DI CONSIDERARE OGNI P.TO AMMISSIBILE $\bar{x}$, DOBBIAMO ASSICURARCI CHE NON CI SIANO PUNTI PER CUI I GRADIENTI SONO LINEARMENTE DIPENDENTI

- I P.TI INTERNI NON CI INTERESSANO: NON CI SONO P.TI ATTIVI
- ANDIAMO SULLA FRONTIERA:

UN VINCOLO DA SOLO È LIN. INDIP. SOLO CON P.TO $\neq (0,0)$

- LUNGO LA FRONTIERA PARABOLICA HO $g_1 = 0$.
  CALCOLO IL GRADIENTE: $\nabla g_1 = (1, 2x_2) \neq 0$ SEMPRE
- LUNGO LA "BASE" HO $\nabla g_2 = (2x_1, 0)$ HO $(0,0)$ CON $x_1 = 0$ MA IN QUEL P.TO g_2 NON È VINCOLO ATTIVO.

- RIMANGONO I PUNTI DOVE g_1 E g_2 SONO ENTRAMBI ATTIVI.
  RISOLVO IL SISTEMA

$$\begin{cases} g_1 = 0 \\ g_2 = 0 \end{cases} \Longrightarrow \begin{cases} x_1 + x_2^2 = 0 \\ x_1^2 - 4 = 0 \end{cases} \Longrightarrow \begin{array}{l} P_1 = (-2, \sqrt{2}) \\ P_2 = (-2, \sqrt{2}) \end{array}$$

SOSTITUISCO P_1 IN $\nabla g(\bar{x})$ E VERIFICO LINEARITÀ INDIPENDENTE

$$\nabla g(P_1) = \begin{pmatrix} 1 & 2\sqrt{2} \\ -4 & 0 \end{pmatrix} \qquad \det \neq 0 \qquad \text{OK!}$$

LE JACOBIANE

SOSTITUISCO P_2 IN $\nabla g(\bar{x})$ E VERIFICO LINEARITÀ INDIPENDENTE

$$\nabla g(P_2) = \begin{pmatrix} 1 & -2\sqrt{2} \\ -4 & 0 \end{pmatrix} \qquad \det \neq 0 \qquad \text{OK!}$$

SE IL DOMINIO D È REGOLARE ALLORA SIAMO CERTI CHE LKKT CONTERRÀ QUANTO STIAMO CERCANDO

SCRIVIAMO IL SISTEMA LKKT

$$\begin{cases} (2\bar{x}_1+2, 2\bar{x}_2+2) + \bar{\lambda}_1 (1, 2\bar{x}_2) + \bar{\lambda}_2 (2\bar{x}_1, 0) = 0 \\ \bar{\lambda}_1 (x_1 + x_2^2) = 0 \\ \bar{\lambda}_2 (x_1^2 - 4) = 0 \end{cases}$$

CALCOLO DEI GRADIENTI

$$\nabla f = (2x_1+2, 2x_2+2)$$
$$\nabla g_1 = (1, 2x_2)$$
$$\nabla g_2 = (2x_1, 0)$$

CIOÈ

$$\begin{cases} 2x_1 + 2 + \lambda_1 + 2\lambda_2 x_1 = 0 \\ 2x_2 + 2 + \lambda_1 2x_2 = 0 \\ \lambda_1 (x_1 + x_2^2) = 0 \\ \lambda_2 (x_1^2 - 4) = 0 \end{cases}$$

$$\begin{cases} \nabla f(\bar{x}) + \sum_{i=1}^{m} \bar{\lambda}_i \nabla g_i(\bar{x}) + \sum_{j=1}^{p} \bar{\mu}_j \nabla h_j(\bar{x}) = 0 \\ \langle \bar{\lambda}, g(\bar{x}) \rangle = 0 \\ h(\bar{x}) = 0 \end{cases}$$

COSA DOBBIAMO FARE NELLO SCRITTO?

A SEGUITO DI CONTROVERSIE SULLE PROVE PASSATE IL PROF HA DECISO DI INTRODURRE UNA SEMPLIFICAZIONE: FORNIRE PARZIALMENTE LE SOLUZIONI. SI VEDA IL SEGUENTE ESEMPIO (DATI FORNITI IN ROSSO)

x	λ
$(-1,-1)$	$(0,0)$
$(-2,-1)$	$(0,-\frac{1}{2})$
$(-2,-\sqrt{2})$	$(-,-)$
$(-2,\sqrt{2})$	$(-,-)$

$$\begin{cases} 2x_1 + 2 = 0 \\ 2x_2 + 2 = 0 \\ 0 = 0 \\ 0 = 0 \end{cases} \qquad \begin{cases} x_1 = -1 \\ x_2 = -1 \end{cases}$$

$$\begin{cases} 2x_1 + 2 - x_1 = 0 \\ 2x_2 + 2 = 0 \end{cases} \qquad \begin{cases} x_1 = -2 \\ x_2 = -1 \end{cases}$$

$$\begin{cases} -4 + 2 + \lambda_1 - 4\lambda_2 = 0 \\ -2\sqrt{2} + 2 + \lambda_1(-2\sqrt{2}) = 0 \\ \lambda_1 (0) = 0 \\ \lambda_2 (0) = 0 \end{cases} \quad \times$$

$$\begin{cases} \lambda_1 - 4\lambda_2 = 2 \\ -2\sqrt{2}\lambda_1 = 2\sqrt{2} - 2 \end{cases}$$

$$\lambda_1 = \frac{2\sqrt{2}-2}{-2\sqrt{2}}$$

$$4\lambda_2 = \lambda_1 - 2 \Rightarrow \lambda_2 = \frac{\lambda_1 - 2}{4}$$

$$\begin{cases} 2x_1 + 2 + \lambda_1 + 2\lambda_2 x_1 = 0 \\ 2x_2 + 2 + \lambda_1 (2x_2) = 0 \end{cases}$$

$$\begin{cases} -4 + 2 + \lambda_1 + 2\lambda_2(-2) = 0 \\ 2\sqrt{2} + 2 + \lambda_1 (2\sqrt{2}) = 0 \end{cases}$$

$$\lambda_1 = -\frac{2+2\sqrt{2}}{2\sqrt{2}} \qquad \lambda_2 = \frac{\lambda_1 - 2}{4}$$

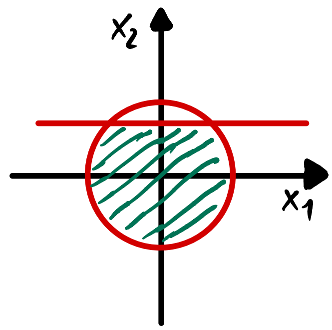
SECONDO ESEMPIO DI STUDIO

CONSIDERIAMO IL SEGUENTE PROBLEMA

$$\begin{cases} f = x_2 - x_1^2 \\ g_1 = x_1^2 + x_2^2 - 4 \leq 0 \\ g_2 = x_2 - 1 \leq 0 \end{cases}$$

IL DOMINIO D É REGOLARE?

GRAFICAMENTE ABBIAMO UNA CIRCONFERENZA "TAGLIATA"



1) NON É UN POLIEDRO

2) h NON PRESENTE
g SICURAMENTE CONVESSO $\Rightarrow$ D REGOLARE
$\exists$ UN P.TO INTERNO (ESEMPIO $(0,0)$)

3) E SE L'ESERCIZIO RICHIEDESSE MANGASARIAN?

$$\nabla g_1 = (2x_1, 2x_2) \qquad \nabla g_2 = (0, 1)$$

JACOBIANA $\longrightarrow$ $\nabla g = \begin{pmatrix} 2x_1 & 2x_2 \\ 0 & 1 \end{pmatrix}$

D É SICURAMENTE LIMITATO

$\exists$ MIN ASSOLUTO
$\exists$ MAX ASSOLUTO

- P.TI INTERNI: NESSUN VINCOLO ATTIVO
- LUNGO LA FRONTIERA, UN SOLO VINCOLO ATTIVO:

VETTORI LINEARMENTE INDIPENDENTI DA SOLI

$\nabla g_1 = (0,0)$ CON P.TO $(0,0)$
NESSUN PROBLEMA, g_1 NON ATTIVO

$\nabla g_2 \neq (0,0)$ SEMPRE

- LUNGO LA FRONTIERA, VINCOLI ENTRAMBI ATTIVI

$$\begin{cases} x_1^2 + x_2^2 - 4 = 0 \\ x_2 - 1 = 0 \end{cases} \qquad \begin{matrix} P_1 = (\sqrt{3}, 1) \\ P_2 = (-\sqrt{3}, 1) \end{matrix}$$

SOSTITUIAMO P_1 E P_2 IN $\nabla g(\bar{x})$

VETTORI LINEARMENTE INDIPENDENTI INSIEME

$$\nabla \bar{g}(P_1) = \begin{pmatrix} 2\sqrt{3} & 2 \\ 0 & 1 \end{pmatrix} \qquad det \neq 0$$

$$\nabla \bar{g}(P_2) = \begin{pmatrix} -2\sqrt{3} & 2 \\ 0 & 1 \end{pmatrix} \qquad det \neq 0$$

SE IL DOMINIO D É REGOLARE ALLORA SIAMO CERTI CHE LKKT CONTERRÁ QUANTO STIAMO CERCANDO.

SCRIVIAMO IL SISTEMA LKKT

$$\begin{cases} \nabla f(\bar{x}) + \sum_{i=1}^{m} \bar{\lambda}_i \nabla g_i(\bar{x}) + \sum_{j=1}^{p} \bar{\mu}_j \nabla h_j(\bar{x}) = 0 \\ \langle \bar{\lambda}, g(\bar{x}) \rangle = 0 \\ h(\bar{x}) = 0 \end{cases}$$

$$\begin{cases} (-2\bar{x}_1, 1) + \bar{\lambda}_1 (2\bar{x}_1, 2x_2) + \bar{\lambda}_2 (0,1) = 0 \\ \bar{\lambda}_1 (x_1^2 + x_2^2 - 4) = 0 \\ \bar{\lambda}_2 (x_2 - 1) = 0 \end{cases}$$

CALCOLO DEI GRADIENTI

$\nabla f = (-2x_1, 1)$

$\nabla g_1 = (2x_1, 2x_2)$

$\nabla g_2 = (0, 1)$

CIOÉ

$$\begin{cases} -2\bar{x}_1 + \bar{\lambda}_1 2\bar{x}_1 = 0 \\ 1 + \bar{\lambda}_1 2\bar{x}_2 + \bar{\lambda}_2 = 0 \\ \bar{\lambda}_1 (\bar{x}_1^2 + \bar{x}_2^2 - 4) = 0 \\ \bar{\lambda}_2 (\bar{x}_2 - 1) = 0 \end{cases}$$

COMPILIAMO LA SEGUENTE TABELLA

f.o.	x	λ	mL	mG	ML	MG	S	
$-\frac{17}{4}$	$\left(\frac{\sqrt{15}}{2}, -\frac{1}{2}\right)$	$(1,0)$	SÍ	SÍ	NO	NO	NO	(1)
$-\frac{17}{4}$	$\left(-\frac{\sqrt{15}}{2}, -\frac{1}{2}\right)$	$(1,0)$	SÍ	SÍ	NO	NO	NO	(1)
-2	$(0, -2)$	$(1/4, 0)$	(*)	NO	NO	NO	(*)	(3)
-2	$(\sqrt{3}, 1)$	$(1, -3)$	NO	NO	NO	NO	SÍ	(2)
-2	$(-\sqrt{3}, 1)$	$(1, -3)$	NO	NO	NO	NO	SÍ	(2)
1	$(0, 1)$	$(0, -1)$	NO	NO	SÍ	SÍ	NO	(1)

SOSTITUENDO $\bar{x}$ NEL SISTEMA LKKT OTTENIAMO

1) IL DOMINIO É CHIUSO E LIMITATO $\Longrightarrow$ $\exists$ MIN ASSOLUTO, $\exists$ MAX ASSOLUTO

DAI VALORI DI F.O. DEDUCO CHE $(0,1)$ É P.TO DI MAX GLOBALE !!!

" " " " " " $\left(\frac{\sqrt{15}}{2}, -\frac{1}{2}\right)$ E $\left(-\frac{\sqrt{15}}{2}, -\frac{1}{2}\right)$ SONO P.TI DI MINIMO GLOBALE !!!

SE SONO GLOBALI SONO ANCHE LOCALI.

2) I MOLTIPLICATORI CON SEGNI DISCORDI SEGNALANO P.TI DI SELLA

3) SICURAMENTE $(0, -2)$ NON É P.TO DI MIN O MAX GLOBALE. SICURAMENTE NON É MAX LOCALE VISTO IL SEGNO DEI MOLTIPLICATORI. CI RIMANGONO I (*)

MI INTERESSA SAPERE SE f É COERCIVA? NO, DOMINIO D LIMITATO.

f É CONVESSA? CALCOLIAMO LA HESSIANA, RICORDANDO CHE $\nabla f = \begin{pmatrix} -2x_1 \\ 1 \end{pmatrix}$

$H f = \begin{pmatrix} -2 & 0 \\ 0 & 0 \end{pmatrix}$ SEMIDEFINITA NEGATIVA $\Rightarrow$ f CONCAVA

g_1 É CONVESSA? CIRCONFERENZA, SÍ!

g_2 É CONVESSA? RETTA, SIA CONVESSA CHE CONCAVA.
HESSIANA SEMIDEFINITA POSITIVA, MA ANCHE
SEMIDEFINITA NEGATIVA

METODO DELLE RESTRIZIONI (DISTINZIONE TRA SELLE E MAX/MIN)

OSSERVAZIONE INTRODUTTIVA

SE $\bar{x}$ É MINIMO LOCALE DI f SU $\mathbb{R}^n \supset D = \{g(x) \leq 0,\ h(x) = 0\}$

ALLORA $\bar{t}$ É MINIMO LOCALE DI φ SULLA RESTRIZIONE

ESERCIZIO PRECEDENTE

RIPRENDIAMO L'ULTIMO ESERCIZIO NELLA PARTE LASCIATA IN SOSPESO

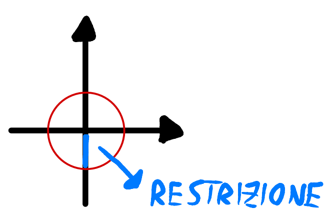
$f = x_2 - x_1^2$

	-2	$(0,-2)$	$(1/4, 0)$	(*)	NO	NO	NO	(*)	(3)

L'IDEA É DI APPLICARE DELLE RESTRIZIONI ALLA FUNZIONE

1) SCELGO LA RESTRIZIONE

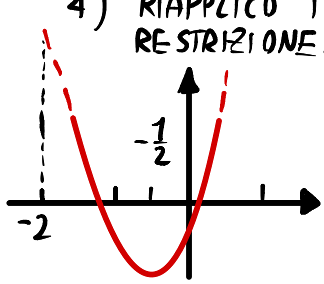
$\begin{cases} x_1 = 0 \\ x_2 = t \end{cases} \quad t \in [-2, 0]$



2) TROVO $\bar{t}$ $\Longrightarrow$ $\bar{t} = -2$ (PONGO $\bar{t}$ IN MODO DA AVERE $(0,2)$)

3) CALCOLO $\varphi(\bar{t})$ $\Longrightarrow$ $\varphi(\bar{t}) = t - 0^2 = t$ $\longleftarrow$ MINIMO -2

4) RIAPPLICO I PUNTI PRECEDENTI CON UNA NUOVA RESTRIZIONE. PRENDIAMO $x_1^2 + x_2^2 = 4$
$\rightarrow x_1^2 = 4 - x_2^2$



SOSTITUISCO IN $f = x_2 - (4 - x_2^2) = x_2^2 + x_2 - 4 = t^2 + t - 4$

MA IN QUESTO CASO CON $\bar{t} = -2$ $\Longrightarrow$ **SELLA**
ABBIAMO IL MASSIMO!!!

METODO DI FRANK-WOLFE (PER MINIMO GLOBALE SU POLIEDRI LIMITATI)

GLI ALGORITMI DI RISOLUZIONE NELLA PNL SI BASANO SU QUANTO GIÀ DETTO NELLA PNL NON VINCOLATA: ABBIAMO UNA SUCCESSIONE ITERATIVA

$$x^{k+1} = x^k + t_k d^k$$

a) SCELTA DI DIREZIONE E PASSO UN CERTO NUM. DI VOLTE

b) TEOREMA DI CONVERGENZA.

TEOREMA IDEALE:

LA SUCCESSIONE $\{x^k\}$ (CON t_k E d^k SCELTI) CONVERGE IN UN NUMERO FINITO DI PASSI AL MINIMO GLOBALE.

POSSIBILE? → PUNTI DI ACCUMULAZIONE

"AREA" IN CUI SI ACCUMULANO INFINITI PUNTI DI UNA SUCCESSIONE.

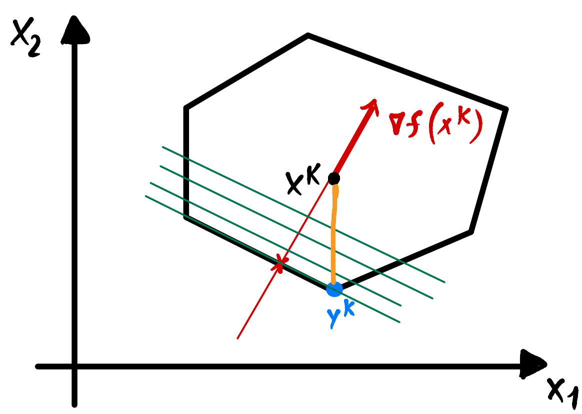
POSSIBILE? → "AL LIMITE"

$$\lim_{k \to +\infty} x^k = x^*$$

POSSIBILE? → METODO ITERATIVO PER SOLUZIONI SISTEMA LKKT (PROBLEMI DIFFICILI, NON QUELLI VISTI DA NOI)

RICORDARE I CRITERI DI STOP

L'ALGORITMO FRANK-WOLF SI UTILIZZA PER TROVARE IL MINIMO GLOBALE DI FUNZIONI NON LINEARI AVENTI COME DOMINIO UN POLIEDRO LIMITATO.



1) CONSIDERIAMO UN PUNTO DI PARTENZA x^k

2) PRENDIAMO IL GRADIENTE $\nabla f(x^k)$ E MUOVIAMOCI IN DIREZIONE OPPOSTA (METODO DI MASSIMA DISCESA)

PROBLEMA: NELLA PNL VINCOLATA POTREI USCIRE DAL POLIEDRO

FACCIAMO QUINDI "DIVERSAMENTE": INTRODUCIAMO IL "PROBLEMA LINEARIZZATO DI x^k"

$$PL(x^k) = \begin{cases} \min \nabla f(x^k) x \\ Ax \leq b \end{cases}$$

← MINIMO, ANDIAMO COMUNQUE IN DIR. OPPOSTA

LA SOLUZIONE OTTIMA DEL PROBLEMA, CHE CHIAMIAMO y^k, È UN VERTICE DEL POLIEDRO (SI VEDA FIGURA, LA COSA È IMMEDIATA SE SI TRACCIANO LE LINEE DI ISOCOSTO).

CONSIDERO IL SEGMENTO $[x^k, y^k]$ CHE È INTERNO AL POLIEDRO

NOI CI MUOVIAMO LUNGO QUEL SEGMENTO, E NON È DETTO CHE CI FERMEREMO AL VERTICE y^K.

SEGUE $\boxed{t_K \in \underset{t \in [0,1]}{\operatorname{argmin}} \ \underbrace{f(x^K + t(y^K - x^K))}_{\varphi(t)}}$ $\varphi(t)$ RESTRIZIONE

DA UNA PARTE CI AIUTA MOLTISSIMO IL SIMPLESSO, DALL'ALTRA PARTE LA DECRESCITA È MINORE, RISPETTO ALL'ALGORITMO NON VINCOLATO.

VARIANTE: PROBLEMA DI MASSIMO

NEL CASO VOLESSI RISOLVERE UN PROBLEMA DI MASSIMO, PONIAMO IL PASSO COME SEGUE

$\boxed{t_K \in \underset{t \in [0,1]}{\operatorname{argmax}} \ \underbrace{f(x^K + t(y^K - x^K))}_{\varphi(t)}}$ argmax E NON argmin!

TEOREMA DI CONVERGENZA

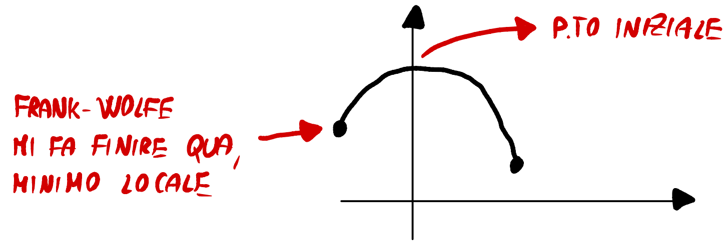
LA SUCCESSIONE $\{x^K\}$ COSTRUITA DALL'ALGORITMO DI FRANK-WOLFE "CONVERGE" AD UN P.TO STAZIONARIO.

$\exists \lim\limits_{k \to +\infty} x^K = x^*$ (SOLUZIONE LKKT)

MEMO: METODO DI FRANK-WOLFE È METODO DI DISCESA

COSA SUCCEDE CON f CONVESSA? CONVERGENZA A MINIMO GLOBALE (NIENTE MINIMI LOCALI, NIENTE SELLE $\longrightarrow$ CON D CHIUSO E LIMITATO) SICURAMENTE.

E CON f CONCAVA? IL METODO DI DISCESA PUÒ FARMI CONVERGERE A MINIMI LOCALI



E SE CERCASSI IL MASSIMO? FRANK-WOLFE AIUTA CON f CONCAVO.

COME TROVO IL P.TO INIZIALE? IN PRIMIS DEVE ESSERE AMMISSIBILE ...

DUALE AUSILIARIO!

CRITERI DI STOP

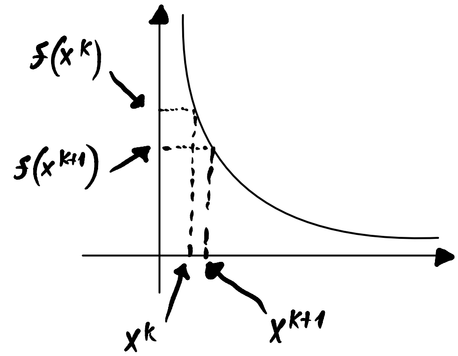
1) $y^K = x^K$? $\longrightarrow$ x^K P.TO STAZIONARIO

2) $t^K = 0$? $\longrightarrow$ x^K P.TO STAZIONARIO

E QUELLI GENERALI...

- NUMERO DI PASSI
- TEMPO
- DISTANZE $\longrightarrow$ $\| x^{k+1} - x^k \| < \varepsilon$; $| f(x^{k+1}) - f(x^k) | < \varepsilon$

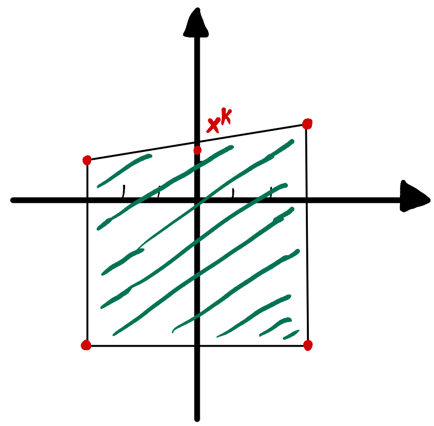
NON HO UNA DISTANZA MIGLIORE



ESEMPIO 1

SI CONSIDERI IL SEGUENTE PROBLEMA

$$\begin{cases} f = 2x_1^2 - 2x_1x_2 - 6x_1 + 5x_2 \\ V = \{(-3,-4), (-3,1), (3,-4), (3,2)\} \\ \text{MAX} \\ x^K = \left(0, \frac{3}{2}\right) \end{cases}$$



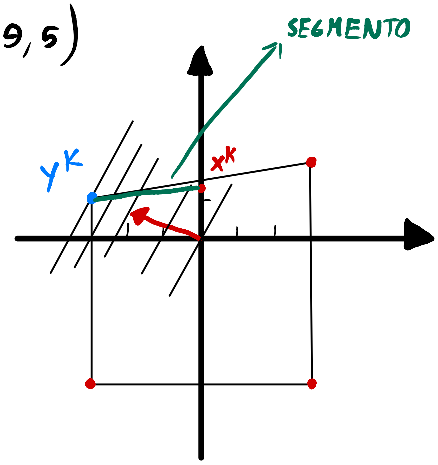
1) CALCOLIAMO IL GRADIENTE NEL P.TO $\left(0, \frac{3}{2}\right)$

$\nabla f(x) = (4x_1 - 2x_2 - 6, -2x_1 + 5)$ $\qquad \nabla f\left(0, \frac{3}{2}\right) = (-9, 5)$

2) SCRIVIAMO IL PROBLEMA LINEARIZZATO

$$PL(x^k) = \begin{cases} \max & -9x_1 + 5x_2 \\ x \in P \end{cases}$$

TROVIAMO COSÌ $y^K = (-3, 1)$ E IL SEGMENTO $[y^K, x^K]$



3) RESTRINGIAMO LA FUNZIONE AL SEGMENTO

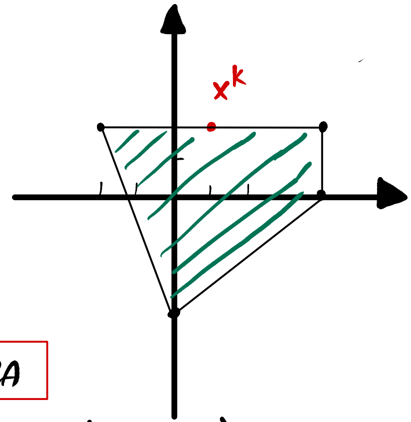
$$\begin{aligned} \varphi(t) &= f\left(\left(0, \tfrac{3}{2}\right) + t\left((-3,1) - \left(0, \tfrac{3}{2}\right)\right)\right) \\ &= f\left(\left(0, \tfrac{3}{2}\right) + \left(-3t, -\tfrac{t}{2}\right)\right) = f\left(-3t, \tfrac{3}{2} - \tfrac{t}{2}\right) \\ &= 2(-3t)^2 - 2(-3t)\left(\tfrac{3}{2} - \tfrac{t}{2}\right) - 6(-3t) + 5\left(\tfrac{3}{2} - \tfrac{t}{2}\right) = 15t^2 + \tfrac{49}{2}t \end{aligned}$$

POICHÉ IL MAX NON È IL VERTICE OTTENIAMO $t_K = 1$ (P.TO DI MASSIMO IN $[0,1]$)

$\Longrightarrow x^{K+1} = \left(0, +\frac{3}{2}\right) + t\left((-3,1) - \left(0 + \frac{3}{2}\right)\right) = (-3, 1)$

ESEMPIO 2

SI CONSIDERI IL SEGUENTE PROBLEMA



$$\begin{cases} f = -2x_1^2 - 6x_1x_2 - 10x_1 - 3x_2 \\ V = \{(-2, 2), (4, 2), (4, 0), (0, -3)\} \\ \text{MIN} \\ x^K = (1, 2) \end{cases}$$

$\boxed{\text{RICORDARSI CHE } Hf = 2A}$

FUNZIONE CONVESSA?

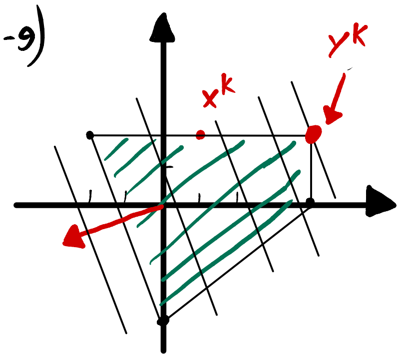
CALCOLIAMO LA HESSIANA $Hf = \begin{pmatrix} -4 & -6 \\ -6 & 0 \end{pmatrix}$

$(-4-\lambda)(-\lambda) - 36 = 0$

$\lambda^2 + 4\lambda - 36 = 0 \longrightarrow \lambda = \dfrac{-4 \pm \sqrt{16+4\cdot 4}}{2}$ NÉ CONCAVA NÉ CONVESSA

1) CALCOLIAMO IL GRADIENTE NEL P.TO $(1,2)$

$\nabla f(x) = (-4x_1 - 6x_2 - 10, -6x_1 - 3) \qquad \nabla f(1,2) = (-26, -9)$



2) SCRIVIAMO IL PROBLEMA LINEARIZZATO

$PL(x^k) = \begin{cases} \min -26x_1 - 9x_2 \\ x \in P \end{cases}$

TROVIAMO COSÌ $y^K = (4, 2)$ E IL SEGMENTO $[y^K, x^K]$

3) RESTRINGIAMO LA FUNZIONE AL SEGMENTO

$$\varphi(t) = f((1,2) + t((4,2) - (1,2)))$$

$$= f(1+3t, 2) = -2(1+3t)^2 - 6(1+3t)(2) - 10(1+3t) - 3\cdot 2 =$$

$$= -2(1+9t^2-6t) - 12(1+3t) - 10(1+3t) - 6 =$$

$$= -2(1+9t^2-6t) - 22(1+3t) - 6 =$$

$$= -18t^2 - 78t + K \rightarrow$$ LA COSTANTE NON CI INTERESSA

IL MINIMO DELLA FUNZIONE TROVATA STA IN $t_K = 1$

$$\Longrightarrow x^{K+1} = (1,2) + t((4,2) - (1,2)) = (4,2)$$

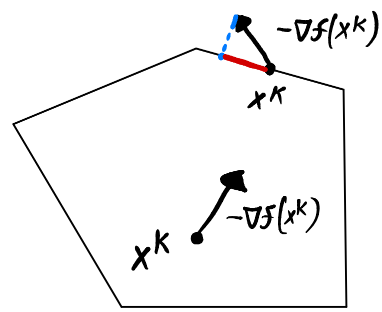
METODO DEL GRADIENTE PROIETTATO

SPIEGAZIONE

CONSIDERIAMO IL PROBLEMA DI MIN $\begin{cases} \min f(x) \\ Ax \leq b \end{cases}$

CONSIDERIAMO x^K E SEGUIAMO LA DIREZIONE $-\nabla f(x^k)$.

- SE x^k É P.TO INTERNO AL POLIEDRO NON CI SONO PROBLEMI
- SE x^k É SULLA FRONTIERA SONO DOLORI: LA DIREZIONE $-\nabla f(x^k)$ POTREBBE PORTARMI FUORI DAL POLIEDRO.



RISOLVIAMO ATTRAVERSO LA PROIEZIONE ORTOGONALE DEL GRADIENTE. MI CHIEDO IN PRIMIS SE É UNA DIREZIONE DI DISCESA.

RICORDARSI IL CALCOLO, GIÁ VISTO

$$f(x^K + t d^K) \triangleq \varphi(t)$$

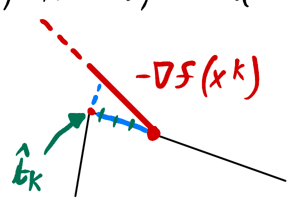
$$\varphi'(t) = \nabla f(x^K + t d^K) d^K$$

$$\varphi'(0) = \nabla f(x^K) d^K \longrightarrow$$

> 0 DIREZIONE DI SALITA
$= 0$ NON ABBIAMO INFORMAZIONI
< 0 DIREZIONE DI DISCESA

CON $d^K = -\nabla f(x^K)$ PER ESSERE CERTI DI AVERE UNA DIREZIONE DI DISCESA.

MA, ADESSO, DI QUANTO POSSO MUOVERMI LUNGO LA PROIEZIONE?



$$t_K \in \underset{t \in [0, \hat{t}_K]}{\operatorname{argmin}} \overbrace{f(x^K + t d^K)}^{\varphi(t)}$$

$\hat{t}_K$ MASSIMO SPOSTAMENTO POSSIBILE

1) CONSIDERO UN PUNTO INIZIALE x^0, CIOÉ HO INIZIALMENTE $K=0$

2) INDIVIDUO I VINCOLI ATTIVI $(i : A_i x^K = b_i)$ E COSTRUISCO CON ESSI LA MATRICE M ($M_i = A_i$, $\forall i$ VINCOLO ATTIVO)

3) COSTRUISCO LA MATRICE DI PROIEZIONE $H = \begin{cases} I & \text{NON HO V.A.} \\ I - M^T (MM^T)^{-1} M & \text{HO V.A.} \end{cases}$
E LA DIREZIONE $d^K = -H \nabla f(x^K)$

4) SE $d^K \neq 0$ RISOLVO IL PROBLEMA $\begin{cases} \max t \\ A(x^K + t d^K) \leq b \end{cases}$ OTTENENDO $\hat{t}_K$

TROVO $t_K \in \underset{t \in [0, \hat{t}_K]}{\operatorname{argmin}} f(x^K + t d^K)$ E CALCOLO $x^{K+1} = x^K + t_K d^K$

PONGO $K = K+1$ E TORNO AL PASSO (2)

5) SE $d^K = 0$ ALLORA VA GESTITO IL CASO PARTICOLARE.

CALCOLO $\lambda = -(MM^T)^{-1} M \nabla f(x^K)$

SE

$\lambda \geq 0$ ABBIAMO FINITO (x^K SODDISFA LKKT)

ALTRIMENTI

$\lambda_J = \min\{\lambda_i : i \text{ VINCOLO ATTIVO}\}$

TOLGO DALLA MATRICE M LA RIGA A_J

TORNO AL PASSO (3)

VARIANTE: PROBLEMA DI MASSIMO

NEL CASO VOLESSI RISOLVERE UN PROBLEMA DI MASSIMO, PONIAMO

$$d^K = H \nabla f(x^K)$$

$$t_K \in \underset{t \in [0,1]}{\operatorname{argmax}} \underbrace{f(x^K + t d^K)}_{\varphi(t)}$$

TOLTO IL SEGNO NEGATIVO

argmax E NON argmin!

RIMANE UGUALE $\hat{t}_K \longrightarrow \begin{cases} \max t \\ A(x^K + t d^K) \leq b \end{cases}$

TEOREMA DI CONVERGENZA

LA SUCCESSIONE $\{x^K\}$ COSTRUITA DALL'ALGORITMO DEL GRADIENTE PROIETTATO "CONVERGE" AD UN P.TO STAZIONARIO.

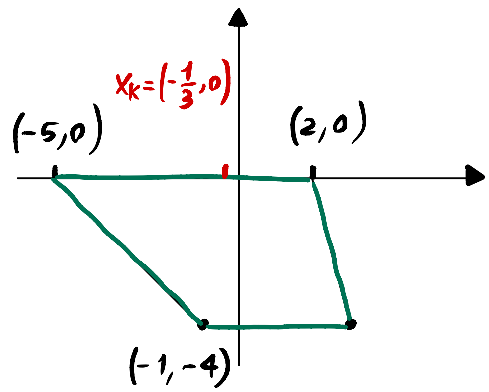
$\exists \lim_{k \to +\infty} x^K = x^*$ (SOLUZIONE LKKT)

COSA SUCCEDE CON f CONVESSA? CONVERGENZA A MINIMO GLOBALE (NIENTE MINIMI LOCALI, NIENTE SELLE $\longrightarrow$ CON D CHIUSO E LIMITATO) SICURAMENTE.

CON f QUADRATICA E CONVESSA $x^K \longrightarrow x^*$ IN UN NUM. FINITO DI PASSI.

ESEMPIO 1

$$\begin{cases} f = 2x_1^2 + 8x_1x_2 - 4x_2^2 + 5x_1 - 6x_2 \\ V = \{(3,-4),\ (-5,0),\ (-1,-4),\ (2,0)\} \\ \text{MIN} \\ x^K = \left(-\frac{1}{3}, 0\right) \end{cases}$$



1) INDIVIDUO I VINCOLI ATTIVI IN x^K E COSTRUISCO LA MATRICE M

IN QUESTO CASO L'UNICO VINCOLO ATTIVO É $x_2 \le 0 \implies M = (0, 1)$

2) CALCOLIAMO LA MATRICE DI PROIEZIONE H

$$H = \underbrace{I}_{\begin{pmatrix}1&0\\0&1\end{pmatrix}} - \underbrace{M^T}_{M^T=\begin{pmatrix}0\\1\end{pmatrix}}\underbrace{(MM^T)^{-1}}_{(0\ 1)\begin{pmatrix}0\\1\end{pmatrix}=1}M = \begin{pmatrix}1&0\\0&1\end{pmatrix} - \begin{pmatrix}0\\1\end{pmatrix}(0\ 1) = \begin{pmatrix}1&0\\0&1\end{pmatrix} - \begin{pmatrix}0&0\\0&1\end{pmatrix} = \begin{pmatrix}1&0\\0&0\end{pmatrix}$$

3) CALCOLIAMO IL VETTORE DIREZIONE $d^K = -H\nabla f(x^K)$

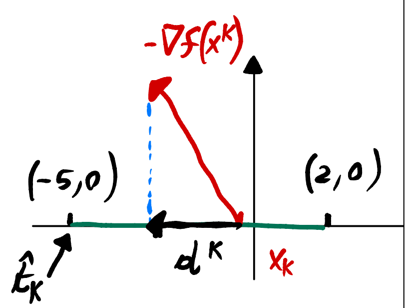
$$\nabla f(x) = (4x_1 + 8x_2 + 5,\ 8x_1 - 8x_2 - 6)$$

$$\nabla f(x^K) = \nabla f\left(-\frac{1}{3}, 0\right) = \left(-\frac{4}{3} + 5,\ -\frac{8}{3} - 6\right) = \left(\frac{11}{3},\ -\frac{26}{3}\right)$$

$$d^K = -\begin{pmatrix}1&0\\0&0\end{pmatrix}\left(\frac{11}{3},\ -\frac{26}{3}\right) = \begin{pmatrix}-\frac{11}{3}\\0\end{pmatrix}$$

4) CALCOLIAMO $\hat{t}_K$

$$-\frac{1}{3} + \hat{t}_K\left(-\frac{11}{3}\right) = -5 \qquad 1 + 11\hat{t}_K = 15 \qquad \hat{t}_K = \frac{14}{11}$$



5) CALCOLIAMO IL PASSO t_K CON RESTRIZIONE AL SEGMENTO $[0, \hat{t}_K]$

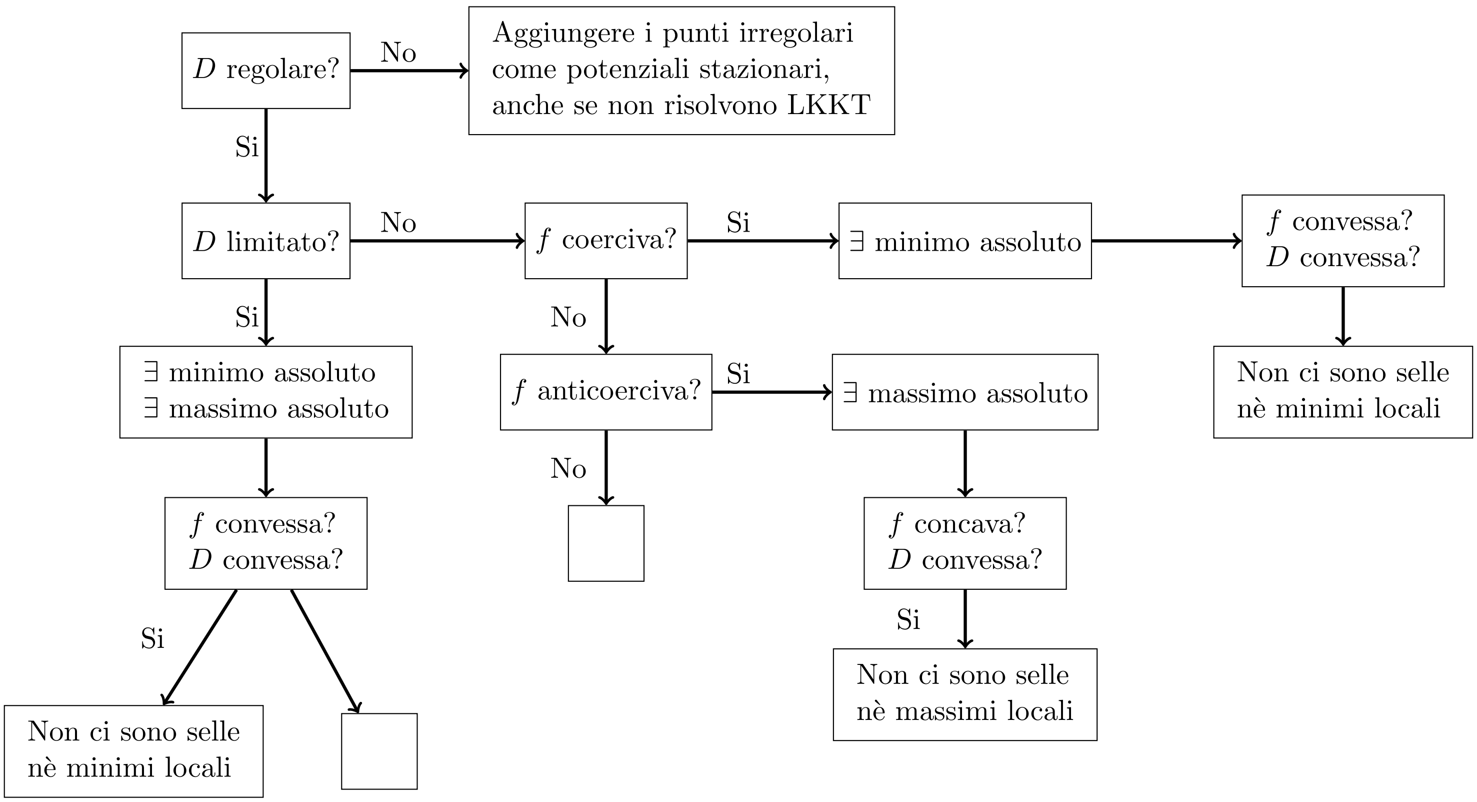
$$\varphi(t) = f\left(\left(-\frac{1}{3}, 0\right) + t\left(-\frac{11}{3}, 0\right)\right) = f\left(-\frac{1}{3} - \frac{11}{3}t,\ 0\right) =$$

$$= 2\left(\frac{1}{9} + \frac{121}{9}t^2 + \frac{22}{9}t\right) + 5\left(-\frac{1}{3} - \frac{11}{3}t\right)$$

$$= \frac{242}{9}t^2 - \frac{121}{9}t + K \qquad \varphi'(t_K) = \frac{484}{9}t_K - \frac{121}{9} = 0 \implies t_K = \frac{\frac{121}{9}}{\frac{484}{9}} = \frac{1}{4}$$

6) TROVO $x^{K+1} = \left(-\frac{1}{3} - \frac{11}{3}\cdot\frac{1}{4},\ 0\right) = \left(-\frac{5}{4}, 0\right)$

1.13 Schema concettuale per la classificazione



Ricordarsi che min/max globale è anche min/max locale.

- **Funzione f convessa?**
  Calcolo la Hessiana per ogni punto x: se ogni hessiana è semidefinita positiva allora abbiamo una funzione convessa. Nel caso di funzioni quadratiche la Hessiana da vedere è una sola, poichè costante.
- **Funzione f concava?**
  Calcolo la Hessiana per ogni punto x: se ogni hessiana è semidefinita negativa allora abbiamo una funzione concava. Nel caso di funzioni quadratiche la Hessiana da vedere è una sola, poichè costante.
- **Dominio D convesso?**
  Un dominio è convesso se g_i è convessa $\forall i$, ed h_j lineare $\forall j$. Se uno ha il disegno verifica ad occhio.
- **Dominio D poliedro?**
  Verifico che g_i sia lineare $\forall i$, e che h_j sia lineare $\forall j$.
- **Dominio D regolare?**
  Sfruttiamo i seguenti criteri (ci fermiamo al primo soddisfatto):
  1. il dominio è un poliedro;
  2. soddisfatta la *condizione di Slater* (il dominio è convesso e si ha un punto interno, $\exists \overline{x} : g(\overline{x}) < 0$)
  3. soddisfatta la *condizione di Mangasarian* (verifico che nei punti in cui si hanno vincoli attivi i gradienti sono linearmente indipendenti)

- Considero i punti in cui ho vincoli attivi. Se non esistono punti del genere posso fermarmi e dichiarare la regolarità del dominio.
- Nel caso di punti con un unico vincolo attivo verifico l'indipendenza lineare assicurandomi di non poter ottenere il punto $(0, 0)$ per mezzo di questi punti.
- Nel caso di punti con più vincoli attivi componiamo la matrice Jacobiana (l'unione dei gradienti), sostituiamo con i punti che stiamo analizzando e calcoliamo il determinante. Se il determinante è diverso da zero ho sicuramente linearità indipendente.

- **Funzione f limitata?**

  Si vede ad occhio. Se lo è siamo certi che esiste un minimo assoluto e un massimo assoluto in virtù del teorema di Weierstrass.

- **Funzione f coerciva? Funzione f anti-coerciva?**

  Ci serve verificare questo solo se siamo in presenza di un dominio non limitato. Verifichiamo la coercività effettuando una serie di restrizioni della funzione: se si va sempre a $+\infty$ allora abbiamo una funzione coerciva. Verifico l'anti-coercività controllando se la funzione f cambiata di segno sia coerciva.

  - Se la funzione f è coerciva siamo certi dell'esistenza del minimo assoluto, in un contesto di dominio non limitato.
  - Se la funzione f è anti-coerciva siamo certi dell'esistenza del massimo assoluto, in un contesto di dominio non limitato.

- **Condizioni che rendono valido l'uso di LKKT**.

  Si osservi che:

  - la regolarità del dominio D è necessaria ai fini del teorema LKKT (C.N)
    1. dato un punto di minimo $\overline{x}$ esistono i relativi moltiplicatori $\lambda \in \mathbb{R}^n_+$ e i coefficienti μ che ris. LKKT
    2. dato un punto di massimo $\overline{x}$ esistono i relativi moltiplicatori $\lambda \in \mathbb{R}^n_-$ e i coefficienti μ che ris. LKKT
  - f convessa e D convessa sono necessarie ai fini del teorema LKKT (C.S.): in presenza di coefficienti $\lambda \geq 0$ possiamo dire che il relativo $\overline{x}$ è minimo globale
  - f concava e D convessa sono necessarie ai fini del teorema LKKT (C.S.): in presenza di coefficienti $\lambda \leq 0$ possiamo dire che il relativo $\overline{x}$ è massimo globale

  Si tenga a mente che f convessa e D convessa, nella PNL non vincolata (D non limitato), mi danno certezza della non esistenza di massimi locali e massimi globali.

- **Compilazione del sistema LKKT**.
  - La regolarità di D ci garantisce l'usabilità del sistema LKKT.
  - Scrivo una tabella con punti $\overline{x}$, relativi moltiplicatori λ e valori della funzione obiettivo.
  - Se il dominio D è limitato sono certa dell'esistenza di un minimo assoluto e di un massimo assoluto. Consideriamo i valori massimi e minimi delle funzioni obiettivo:
    * se ho il valore minimo posso segnalare la soluzione $\overline{x}$ come minimo globale e minimo locale
    * se ho il valore massimo posso segnalare la soluzione $\overline{x}$ come massimo globale e massimo locale
  - Successivamente passo alle soluzioni $\overline{x}$ che mi offrono valori intermedi.
    * posso già escludere a priori che le soluzioni rimanenti siano minimi globali e massimi globali;
    * le soluzioni $\overline{x}$ con moltiplicatori discordi sono selle;
    * un punto con moltiplicatori tutti positivi non è sicuramente un massimo locale;
    * un punto con moltiplicatori tutti negativi non è sicuramente un minimo locale.

- Rimangono casi con moltiplicatori concordi: se f convessa e D convesso possiamo applicare il teorema LKKT (C.S.), altrimenti emerge il dubbio a proposito delle selle.
  * Se ho f convessa e D convessa, e un punto ha moltiplicatori tutti positivi, allora parlo sicuramente di un minimo locale.
  * Se ho f convessa e D convessa, e un punto ha moltiplicatori tutti negativi, allora parlo sicuramente di un massimo locale.
  * Se non posso dire che f è convessa e D convessa, e un punto ha moltiplicatori tutti positivi, quel punto può essere minimo locale o sella.
  * Se non posso dire che f è convessa e D convessa, e un punto ha moltiplicatori tutti negativi, quel punto può essere massimo locale o sella.

  In caso di dubbio sulle selle si applica il metodo delle restrizioni: supponiamo di avere il dubbio tra minimo e massimo, l'idea è che se applico una restrizione il relativo parametro t sarà minimo nella restrizione. Se applicando più restrizioni ho casi in cui il parametro t è massimo allora ho una sella.

Parte III

Appendici

Appendice A

Definizioni

A.1 Programmazione Non Lineare

Definizione di *Punto stazionario*.

Data una funzione $f : \mathbb{R}^n \rightarrow \mathbb{R}$ definisco *punto stazionario* il valore $x \in \mathbb{R}^n$ tale che $\nabla f(x) = 0$

Definizione di *Matrice definita e semidefinita positiva e negativa*.

Definiamo quanto segue.

- Con **matrice definita positiva** si intende una matrice $A \in \mathbb{R}^{n \times n}$ tale che il seguente prodotto scalare risulta positivo $< A \cdot x, x > > 0 \; \forall x \in \mathbb{R}^n$
- Con **matrice semidefinita positiva** si intende una matrice $A \in \mathbb{R}^{n \times n}$ tale che il seguente prodotto scalare risulta positivo o nullo $< A \cdot x, x > \geq 0 \; \forall x \in \mathbb{R}^n$
- Con **matrice definita negativa** si intende una matrice $A \in \mathbb{R}^{n \times n}$ tale che il seguente prodotto scalare risulta negativo $< A \cdot x, x > < 0 \; \forall x \in \mathbb{R}^n$
- Con **matrice semidefinita negativa** si intende una matrice $A \in \mathbb{R}^{n \times n}$ tale che il seguente prodotto scalare risulta negativo o nullo $< A \cdot x, x > \leq 0 \; \forall x \in \mathbb{R}^n$

Per definire la matrice troviamo gli autovalori calcolando i valori λ nel determinante $\det(A - \lambda I) = 0$. Prendiamo ad esempio la matrice seguente, poniamo la matrice $A - \lambda I$ e calcoliamo quanto detto

$$\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \Longrightarrow \begin{pmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{pmatrix} \Longrightarrow (2-\lambda)(2-\lambda) - 1 = 0 \Longrightarrow \begin{array}{l} \lambda^2 - 4\lambda + 3 = 0 \\ \lambda_1 = 3, \lambda_2 = 1 \end{array}$$

Definizione di *Restrizione della funzione*.

Data una funzione $f : \mathbb{R}^n \to \mathbb{R}$ definiamo *restrizione della funzione* una funzione $\gamma : \mathbb{R} \to \mathbb{R}$ ottenuta restringendo il dominio della funzione. Nel nostro caso ciò che ci interessa è la *restrizione a una semiretta*, dove scriviamo l'equazione della semiretta in formato parametrico e successivamente sostituiamo nella funzione f. Consideriamo il seguente esempio

$$f(x_1, x_2) = x_1^2 - x_2^2$$
$$\begin{cases} x_1 = 0 + t \\ x_2 = 3 + t \end{cases} \quad \Longrightarrow \gamma(t) = t^2 - (3+t)^2 = -9 - 6t$$

Definizione di *Funzione quadratica*.

Una funzione $f : \mathbb{R}^n \to \mathbb{R}$ è detta *quadratica* se consiste in un polinomio di secondo grado avente la seguente struttura

$$< x, Ax > + < c, x >$$

dove si distingue una componente lineare da una non lineare. Queste funzioni ci piacciono poichè il gradiente è un sistema lineare, e la hessiana costante!

- **Esempio**. $3x_1^2 + 4x_1x_2 + x_2^2 + 5x_1 - 6x_2$
- **Esempio sbagliato 1**. $4x_1^3 - 5x_1x_2$ (primo termine di grado 3)
- **Esempio sbagliato 2**. $x_1^2x_2 - 3x_2^2$ (termine misto di grado 3)

Definizione di *Funzione coerciva*.

Una funzione $f : \mathbb{R}^n \to \mathbb{R}$ è detta *coerciva* se

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

cioè se la funzione va a $+\infty$ in tutte le direzioni. Queste funzioni ci piacciono perchè ci danno certezza dell'esistenza di un minimo assoluto. Verifico se una funzione è coerciva testando una serie di restrizioni della funzione.

Definizione di *Funzione anti-coerciva*.

Una funzione $f : \mathbb{R}^n \to \mathbb{R}$ è detta *anti-coerciva* se la funzione f cambiata di segno è coerciva

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

Queste funzioni ci piacciono perchè ci danno certezza dell'esistenza di un massimo assoluto.

Definizione di *Funzione convessa*.

Una funzione $f : \mathbb{R}^n \to \mathbb{R}$ è detta *convessa* se, data una qualunque coppia di punti $x_1, x_2 \in \mathbb{R}^n$ possiamo definire una corda che sottende la curva

$$f\left(\lambda x^1 + (1-\lambda)x^2\right) \leq \lambda f(x^1) + (1-\lambda)f(x^2), \forall \lambda \in [0,1], \forall x^1, x^2 \in \mathbb{R}^n$$

Una funzione f è convessa se la hessiana è semidefinita positiva $\forall x$. Ci piacciono molto le funzioni convesse e quadratiche poichè la hessiana è costante, e questo permette una verifica veloce. In più si veda il *teorema su punti stazionari in funzioni convesse.*

Appendice B

Teoremi

B.1 Programmazione Non Lineare

Teorema di AN2 sulla simmetria delle matrici hessiane.
Dato $f : \mathbb{R}^n \to \mathbb{R}, f \in C^2$, $f \in C^\infty$, la matrice hessiana $Hf(x)$ è simmetrica!

Teorema di AL su matrici simmetriche e autovalori.
Le matrici simmetriche hanno autovalori reali

Teorema di Fermat per l'Analisi II (min locale).
Dato $f : \mathbb{R}^n \to \mathbb{R}, f \in C^2$, $f \in C^\infty$, se $\overline{x}$ è minimo locale di f allora $\overline{x}$ è un p.to stazionario.
Attenzione: è CN, ma non CS:

- se ho un p.to di minimo $\overline{x}$ è automatico che questo sia p.to stazionario;
- se ho un p.to stazionario $\overline{x}$ non è automatico che questo sia p.to di minimo.

Teorema 2 (min locale).
Dato $f : \mathbb{R}^n \to \mathbb{R}, f \in C^2$, $f \in C^\infty$, se $\overline{x}$ è minimo locale di f e $\overline{x}$ p.to stazionario allora

$$Hf(\overline{x}) \geq 0$$

la Hessiana è *semidefinita positiva*. Attenzione: è CN, ma non CS:

- se ho p.to di minimo $\overline{x}$ è automatico che $Hf(\overline{x}) \geq 0$;
- se $Hf(\overline{x}) \geq 0$ non è automatico che $\overline{x}$ sia p.to di minimo.

Teorema 3 (min locale).

Dato $f : \mathbb{R}^n \to \mathbb{R}, f \in C^2$, se $\nabla f(\overline{x}) = 0$ e $Hf(\overline{x}) > 0$ (hessiana nel p.to *definita positiva*) allora $\overline{x}$ è minimo locale! Questa è CS!

Teorema di Fermat per l'Analisi II (max locale).

Dato $f : \mathbb{R}^n \to \mathbb{R}, f \in C^2$, $f \in C^\infty$, se $\overline{x}$ è massimo locale di f allora $\overline{x}$ è un p.to stazionario. Attenzione: è CN, ma non CS:

- se ho un p.to di massimo $\overline{x}$ è automatico che questo sia p.to stazionario;
- se ho un p.to stazionario $\overline{x}$ non è automatico che questo sia p.to di massimo.

Teorema 2 (max locale).

Dato $f : \mathbb{R}^n \to \mathbb{R}, f \in C^2$, $f \in C^\infty$, se $\overline{x}$ è massimo locale di f e $\nabla f(\overline{x}) = 0$ allora

$$Hf(\overline{x}) \leq 0$$

la Hessiana è *semidefinita negativa*. Attenzione: è CN, ma non CS:

- se ho p.to di massimo $\overline{x}$ è automatico che $Hf(\overline{x}) \leq 0$;
- se $Hf(\overline{x}) \leq 0$ non è automatico che $\overline{x}$ sia p.to di massimo.

Teorema 3 (max locale).

Dato $f : \mathbb{R}^n \to \mathbb{R}, f \in C^2$, se $\nabla f(\overline{x}) = 0$ e $Hf(\overline{x}) < 0$ (hessiana nel punto *definita negativa*) allora $\overline{x}$ è massimo locale! Questa è CS!

Teorema sui punti stazionari in funzioni convesse.

I punti stazionari di funzioni convesse, se esistono, sono minimi assoluti.

Teorema di convergenza del metodo del gradiente.

Sia f una funzione coerciva, la successione del gradiente con ricerca esatta $\{x^k\}$

- termina con un numero finito di passi in un punto stazionario, o
- i suoi punti di accumulazione sono punti stazionari.

Teorema LKKT (Lagrange-Karush-Kulm-Tucker, C.N.).

Sia P il seguente problema

$$(P) = \begin{cases} \min / \max f(x) \\ g_1(x) \leq 0 \\ \vdots \\ g_m(x) \leq 0 \\ h_1(x) = 0 \\ h_p(x) = 0 \end{cases}$$

siano $f, g_1, \dots, g_m, h_1, \dots, h_p : \mathbb{R}^n \to \mathbb{R}$, siano $f, g_1, \dots, g_m, h_1, \dots, h_p \in \mathbb{C}^2$. Sia D un insieme regolare. Affermiamo che (C.N., ma non C.S.):

- sia $\overline{x}$ minimo locale per (P). Allora $\exists \lambda \in \mathbb{R}^n_+$ ed $\exists \mu$ tali che

$$\begin{cases} \nabla f(x) + \sum_{i=1}^{m} \lambda_i \nabla g_i(\overline{x}) + \sum_{j=1}^{p} \mu_j \nabla h_j(\overline{x}) = 0 \\ < \lambda, g(\overline{x}) >= 0 \\ h(\overline{x}) = 0 \end{cases}$$

- sia $\overline{x}$ massimo locale per (P). Allora $\exists \lambda \in \mathbb{R}^n_-$ ed $\exists \mu$ tali che

$$\begin{cases} \nabla f(x) + \sum_{i=1}^{m} \lambda_i \nabla g_i(\overline{x}) + \sum_{j=1}^{p} \mu_j \nabla h_j(\overline{x}) = 0 \\ < \lambda, g(\overline{x}) >= 0 \\ h(\overline{x}) = 0 \end{cases}$$

Teorema LKKT (Lagrange-Karush-Kulm-Tucker, C.S.).

- Sia f convessa e D convesso, oltre che regolare. Inoltre $f, g, h \in C^2$. Sia $(\overline{x}, \overline{\lambda}, \overline{\mu})$ una soluzione del sistema LKKT. Se $\lambda \geq 0$ allora abbiamo $\overline{x}$ minimo globale.
- Sia f concava e D convesso, oltre che regolare. Inoltre $f, g, h \in C^2$. Sia $(\overline{x}, \overline{\lambda}, \overline{\mu})$ una soluzione del sistema LKKT. Se $\lambda \leq 0$ allora abbiamo $\overline{x}$ massimo globale.

Teorema di convergenza del metodo di Frank-Wolfe.

Una successione $\{x_k\}$ costruita con l'algoritmo di Frank-Wolfe converge ad un punto stazionario.

Teorema di convergenza del metodo del gradiente proiettato.

Una successione $\{x_k\}$ costruita con l'algoritmo del gradiente proiettato converge ad un punto stazionario.