Teoria essenziale del primo esonero di Programmazione Matematica

Canale 1 · Prof. Marco Sciandrone

Il minimo indispensabile per l'esonero dell'11 novembre 2026, con i numeri del libro di G. Liuzzi e M. Sciandrone. Il resto del programma, comprese le dimostrazioni dubbie, è nella *Teoria del primo esonero* completa.

Dimostrazione da fare: va saputa. **Solo enunciato**: basta l'enunciato.

1 Lo spazio $\mathbb{R}^n$

Norme e distanze in $\mathbb{R}^n$

Definizione 1.12 – Norma

Una **norma** su $\mathbb{R}^n$ è una funzione $\|\cdot\| : \mathbb{R}^n \to \mathbb{R}_+$, con $\mathbb{R}_+ = \{t \in \mathbb{R} : t \geq 0\}$, tale che, per ogni $x, y \in \mathbb{R}^n$ e $\alpha \in \mathbb{R}$:

(1) $\|x\| \geq 0$;
(2) $\|x\| = 0$ se e solo se $x = 0$;
(3) $\|\alpha x\| = |\alpha| \, \|x\|$;
(4) $\|x + y\| \leq \|x\| + \|y\|$ (disuguaglianza triangolare).

Proposizione 1.16 – Disuguaglianza di Cauchy-Schwarz

Per ogni $x, y \in \mathbb{R}^n$, con $x^\mathsf{T} y = \sum_{i=1}^n x_i y_i$ il prodotto scalare,

$$|x^\mathsf{T} y| \leq \|x\|_2 \, \|y\|_2,$$

e vale l'uguaglianza se e solo se i due vettori sono paralleli, cioè uno è multiplo dell'altro.

Solo enunciato

Intorni, insiemi aperti, insiemi chiusi

Definizione 1.28-1.29 – Insieme aperto e insieme chiuso

Sia $E \subseteq \mathbb{R}^n$. L'**intorno sferico** di $\bar{x} \in \mathbb{R}^n$ di raggio $\rho > 0$ è l'insieme dei punti che distano da $\bar{x}$ meno di ρ, $\mathcal{B}_\rho(\bar{x}) = \{x \in \mathbb{R}^n : \|x - \bar{x}\| < \rho\}$. Un punto $\bar{x} \in E$ è **interno** a E se esiste un intorno sferico di $\bar{x}$ tutto contenuto in E, cioè $\exists\, \rho > 0 : \mathcal{B}_\rho(\bar{x}) \subseteq E$. Il **complementare** di E è $\mathrm{co}(E) = \{x \in \mathbb{R}^n : x \notin E\}$. L'insieme E è

- **aperto** se ogni suo punto è interno;
- **chiuso** se il suo complementare $\mathrm{co}(E)$ è aperto.

Punti di accumulazione e isolati

Definizione 1.34 – Punto di accumulazione

Sia $E \subseteq \mathbb{R}^n$. Un punto $\bar{x} \in \mathbb{R}^n$ è **di accumulazione** per E se ogni suo intorno sferico contiene un punto di E diverso da $\bar{x}$, cioè

$$\big(\mathcal{B}_\rho(\bar{x}) \cap E\big) \setminus \{\bar{x}\} \neq \emptyset \qquad \forall \rho > 0.$$

L'insieme dei punti di accumulazione di E si chiama **derivato** di E e si indica con $\mathscr{D}(E)$.

Proposizione 1.36 – Caratterizzazione degli insiemi chiusi

Sia $E \subseteq \mathbb{R}^n$. Le tre affermazioni seguenti sono equivalenti:

(1) E è chiuso;
(2) $\partial E \subseteq E$;
(3) $\mathscr{D}(E) \subseteq E$.

Dimostrazione da fare

Dimostrazione

Si dimostra il cerchio (1) $\Rightarrow$ (2) $\Rightarrow$ (3) $\Rightarrow$ (1), perché tre implicazioni bastano per l'equivalenza.

(1) **(1) $\Rightarrow$ (2).** Dimostrare $\partial E \subseteq E$ equivale a dimostrare che un punto fuori da E non è di frontiera. Se $\bar{x} \notin E$, allora $\bar{x} \in \mathrm{co}(E)$, che è aperto perché E è chiuso: esiste $\mathcal{B}_\rho(\bar{x}) \subseteq \mathrm{co}(E)$. Questo intorno non contiene punti di E, quindi $\bar{x} \notin \partial E$.

(2) **(2) $\Rightarrow$ (3).** Per assurdo, sia $\bar{x} \in \mathscr{D}(E)$ con $\bar{x} \notin E$. Non è interno, perché non sta in E, e non è isolato, perché è di accumulazione. È quindi di frontiera: ogni suo intorno sferico contiene un punto di E, perché $\bar{x}$ è di accumulazione, e contiene anche $\bar{x} \in \mathrm{co}(E)$. Allora $\bar{x} \in \partial E \subseteq E$, contro $\bar{x} \notin E$.

(3) **(3) $\Rightarrow$ (1).** Si dimostra che $\mathrm{co}(E)$ è aperto. Sia $\bar{x} \in \mathrm{co}(E)$. Poiché $\mathscr{D}(E) \subseteq E$, $\bar{x}$ non è di accumulazione: un intorno $\mathcal{B}_\rho(\bar{x})$ non contiene punti di E diversi da $\bar{x}$, e poiché $\bar{x} \notin E$ non ne contiene nessuno. Quindi $\mathcal{B}_\rho(\bar{x}) \subseteq \mathrm{co}(E)$.

□

Definizione 1.37-1.38 – Insieme limitato e insieme compatto

Un insieme $E \subseteq \mathbb{R}^n$ è **limitato** se la norma dei suoi elementi non supera una costante:

$$\exists\, M > 0 : \quad \|x\| \leq M \qquad \forall x \in E.$$

In $\mathbb{R}^n$ un insieme si dice **compatto** se è chiuso e limitato.

2 Funzioni in $\mathbb{R}^n$

Limiti di funzioni vettoriali

Definizione 2.4, 2.6 – Intorni sferici e intorni

Si pone $\mathbb{R}^* = \mathbb{R} \cup \{-\infty, +\infty\}$ e $\mathbb{R}^n_* = \mathbb{R}^n \cup \{\infty\}$, aggiungendo a $\mathbb{R}^n$ un solo elemento, il punto all'infinito. Ogni elemento di $\mathbb{R}^*$ e di $\mathbb{R}^n_*$ ha i suoi **intorni sferici**:

di $x_0 \in \mathbb{R}^n$	$\mathcal{B}_\rho(x_0) = \{x \in \mathbb{R}^n : \|x - x_0\| < \rho\}$,	$\rho > 0$;
di $+\infty$	$\{x \in \mathbb{R} : x > A\} \cup \{+\infty\}$,	$A \in \mathbb{R}$;
di $-\infty$	$\{x \in \mathbb{R} : x < A\} \cup \{-\infty\}$,	$A \in \mathbb{R}$;
di ∞	$\{x \in \mathbb{R}^n : \|x\| > A\} \cup \{\infty\}$,	$A > 0$.

Nello spazio esteso, un **intorno** di un elemento è un insieme aperto che contiene un suo intorno sferico.

Definizione 2.5, 2.7, 2.9 – Limite

Siano $f : D \subseteq \mathbb{R}^n \to \mathbb{R}^m$, $x_0 \in \mathbb{R}^n_*$ un punto di accumulazione di D e $\ell \in \mathbb{R}^m_*$. Si scrive $\lim_{x \to x_0} f(x) = \ell$ se per ogni intorno $\mathcal{U}(\ell)$ esiste un intorno $\mathcal{N}(x_0)$ tale che

$$x \in \big(\mathcal{N}(x_0) \cap D\big) \setminus \{x_0\} \implies f(x) \in \mathcal{U}(\ell).$$

Per $n = 1$ al posto di $\mathbb{R}^n_*$ si usa $\mathbb{R}^*$, e per $m = 1$ al posto di $\mathbb{R}^m_*$.

Definizione – Limite di una funzione di più variabili

Siano $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$ e $x_0 \in \mathbb{R}^n$ un punto di accumulazione di D. Si scrive $\lim_{x \to x_0} f(x) = \ell$ se per ogni $\varepsilon > 0$ esiste $\delta > 0$ tale che

$$x \in D,\ 0 < \|x - x_0\| < \delta \implies |f(x) - \ell| < \varepsilon;$$

si scrive $\lim_{x \to x_0} f(x) = +\infty$ (rispettivamente $-\infty$) se per ogni $M \in \mathbb{R}$ esiste $\delta > 0$ tale che

$$x \in D,\ 0 < \|x - x_0\| < \delta \implies f(x) > M \quad \text{(rispettivamente } f(x) < M\text{)}.$$

Successioni in $\mathbb{R}^n$

Definizione – Sottosuccessione

Sia $K = \{k_1, k_2, k_3, \dots\}$ un insieme infinito di indici del dominio di $\{x_k\}$, con $k_1 < k_2 < k_3 < \cdots$. La **sottosuccessione** di $\{x_k\}$ di indici K è la successione $\{y_i\}$ con $y_i = x_{k_i}$, $i = 1, 2, 3, \dots$; si indica anche con $\{x_k\}_K$.

Proposizione 2.14 – Successioni limitate e componenti

Con la norma euclidea, una successione $\{x_k\} \subset \mathbb{R}^n$ è limitata se e solo se ogni successione $\{(x_k)_i\}$ delle componenti è limitata, per $i = 1, \dots, n$.

Solo enunciato

Limiti di successioni in $\mathbb{R}^n$

Definizione 2.17 – Limite di una successione

Sia $\{x_k\}$ una successione in $\mathbb{R}^n$. Si scrive $\lim_{k\to\infty} x_k = \bar{x} \in \mathbb{R}^n$ se per ogni $\varepsilon > 0$ esiste un indice k_ε tale che

$$k \geq k_\varepsilon \implies \|x_k - \bar{x}\| < \varepsilon.$$

Teorema 2.24 – Bolzano-Weierstrass in $\mathbb{R}$

Ogni successione reale limitata ammette almeno una sottosuccessione convergente.

Solo enunciato

Teorema 2.25 – Bolzano-Weierstrass per successioni

Ogni successione $\{x_k\}$ in $\mathbb{R}^n$ limitata ammette almeno una sottosuccessione convergente: esistono un insieme infinito di indici K e $\bar{x} \in \mathbb{R}^n$ tali che

$$\lim_{\substack{k\in K \\ k\to\infty}} x_k = \bar{x}.$$

Dimostrazione da fare

Dimostrazione

Le successioni delle componenti sono limitate (Proposizione 2.14).

(1) **Prima componente.** $\{(x_k)_1\}$ è reale e limitata, quindi per Bolzano-Weierstrass in $\mathbb{R}$ (Teorema 2.24) esiste K_1 con $(x_k)_1 \to \bar{x}_1$ per $k \in K_1$.

(2) **Seconda componente.** Lungo K_1, $\{(x_k)_2\}$ è ancora limitata, quindi esiste $K_2 \subseteq K_1$ con $(x_k)_2 \to \bar{x}_2$ per $k \in K_2$. Lungo K_2 anche la prima componente tende a $\bar{x}_1$, perché K_2 dà una sottosuccessione di quella su K_1.

(3) **Le altre.** Si ripete fino alla componente n: $K_n \subseteq K_{n-1} \subseteq \cdots \subseteq K_1$, e lungo K_n ogni componente i tende a $\bar{x}_i$.

(4) **Conclusione.** Sia $\bar{x} = (\bar{x}_1, \dots, \bar{x}_n)^\mathsf{T}$. Per ogni $y \in \mathbb{R}^n$ si ha $\|y\| \leq |y_1| + \cdots + |y_n|$. Dato $\varepsilon > 0$, per ogni i vale $|(x_k)_i - \bar{x}_i| < \varepsilon/n$ per $k \in K_n$ abbastanza grande. Con $y = x_k - \bar{x}$ ne segue $\|x_k - \bar{x}\| < \varepsilon$. Quindi $x_k \to \bar{x}$ per $k \in K_n$.

□

Corollario 2.26 – Successioni in un compatto

Siano $E \subseteq \mathbb{R}^n$ un insieme compatto, cioè chiuso e limitato, e $\{x_k\}$ una successione con $x_k \in E$ per ogni k. Allora esistono un insieme infinito di indici K e $\bar{x} \in E$ tali che

$$\lim_{\substack{k \in K \\ k \to \infty}} x_k = \bar{x}.$$

Solo enunciato

Teorema 2.27 – Bolzano-Weierstrass per insiemi (anche Teorema 1.5)

Ogni insieme $E \subseteq \mathbb{R}^n$ limitato e con infiniti punti ha almeno un punto di accumulazione.

Dimostrazione da fare

Dimostrazione

(1) E ha infiniti punti, quindi si sceglie una successione $\{x_k\}$ di punti di E distinti fra loro.

(2) E è limitato, quindi anche $\{x_k\}$ è limitata. Per Bolzano-Weierstrass per successioni (Teorema 2.25) esistono K e $\bar{x}$ con $x_k \to \bar{x}$ per $k \in K$.

(3) Dato $\rho > 0$, per la definizione di limite esiste k_ρ tale che $x_k \in \mathcal{B}_\rho(\bar{x})$ per ogni $k \in K$ con $k \geq k_\rho$. Questi termini sono infiniti e distinti, quindi al più uno coincide con $\bar{x}$. Così l'intorno contiene un punto di E diverso da $\bar{x}$, e $\bar{x}$ è di accumulazione per E.

□

Continuità

Definizione 2.30 – Funzione continua

Siano $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$ e $x_0 \in D$. La funzione f è **continua** in x_0 se

- x_0 è un punto isolato di D, oppure
- x_0 è un punto di accumulazione di D e

$$\lim_{x \to x_0} f(x) = f(x_0).$$

È continua su un insieme $E \subseteq D$ se è continua in ogni punto di E, e si scrive $f \in C^0(E)$.

Lemma 2.33 – Lipschitz implica continua

Sia $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$ Lipschitz continua in D con costante $L > 0$, cioè

$$|f(x) - f(y)| \leq L\|x - y\| \qquad \forall x, y \in D.$$

Allora f è continua in D, cioè è continua in ogni $x_0 \in D$. Dato $\varepsilon > 0$, il raggio $\delta = \varepsilon/L$ va bene in ogni punto di D.

Dimostrazione da fare

Dimostrazione

1. Siano $x_0 \in D$ e $\varepsilon > 0$. Se $x \in D$ e $\|x - x_0\| < \varepsilon/L$, allora
$$|f(x) - f(x_0)| \leq L\|x - x_0\| < L \cdot \frac{\varepsilon}{L} = \varepsilon.$$
2. Se x_0 è di accumulazione per D, questa è la definizione di $\lim_{x \to x_0} f(x) = f(x_0)$; se è isolato, f è continua in x_0 per definizione.

□

3 Differenziazione di funzioni di più variabili

Derivate del primo ordine di una funzione di più variabili

Definizione 3.3 – Derivata direzionale

Siano $f : \mathbb{R}^n \to \mathbb{R}$, $x \in \mathbb{R}^n$ e $d \in \mathbb{R}^n$. f ha **derivata direzionale** in x lungo la direzione d se esiste finito il limite

$$Df(x, d) = \lim_{t \to 0^+} \frac{f(x + t\,d) - f(x)}{t}.$$

Definizione 3.4 – Derivata parziale e gradiente

Siano $f : \mathbb{R}^n \to \mathbb{R}$ e $x \in \mathbb{R}^n$. f ha **derivata parziale** rispetto alla variabile x_j nel punto x se esiste finito il limite

$$\frac{\partial f}{\partial x_j}(x) = \lim_{t \to 0} \frac{f(x_1, \ldots, x_j + t, \ldots, x_n) - f(x_1, \ldots, x_j, \ldots, x_n)}{t}.$$

Se in x esistono le derivate parziali rispetto a tutte le variabili, il **gradiente** di f in x è il vettore colonna $\nabla f(x) = \left(\frac{\partial f}{\partial x_1}(x), \ldots, \frac{\partial f}{\partial x_n}(x)\right)^{\mathsf{T}}$.

Proposizione 3.5 – Derivata parziale e derivate direzionali

Siano $f : \mathbb{R}^n \to \mathbb{R}$ e $x \in \mathbb{R}^n$, e sia e_j il j-esimo asse coordinato, cioè il vettore con 1 nella componente j e 0 nelle altre. Esiste in x la derivata parziale $\dfrac{\partial f}{\partial x_j}(x)$ se e solo se

(i) esiste la derivata direzionale $Df(x, e_j)$;
(ii) esiste la derivata direzionale $Df(x, -e_j)$;
(iii) risulta $Df(x, e_j) = -Df(x, -e_j)$.

In tal caso $\dfrac{\partial f}{\partial x_j}(x) = Df(x, e_j)$.

Dimostrazione da fare

Dimostrazione

Il punto $x + t\, e_j$ differisce da x solo nella coordinata j, che diventa $x_j + t$, quindi

$$\frac{\partial f}{\partial x_j}(x) = \lim_{t \to 0} \frac{f(x + t\, e_j) - f(x)}{t}.$$

La derivata parziale esiste se e solo se i limiti destro e sinistro del rapporto $\frac{f(x+t\, e_j)-f(x)}{t}$ esistono finiti e sono uguali.

1. **Limite destro.** Per definizione è $Df(x, e_j)$.
2. **Limite sinistro.** Con $t = -h$, $t \to 0^-$ equivale a $h \to 0^+$ e
$$\lim_{t \to 0^-} \frac{f(x + t\, e_j) - f(x)}{t} = \lim_{h \to 0^+} \frac{f(x + h(-e_j)) - f(x)}{-h} = -Df(x, -e_j),$$
e il primo limite esiste finito se e solo se esiste $Df(x, -e_j)$.
3. ($\Rightarrow$) Se la derivata parziale esiste, i due limiti esistono e sono uguali: valgono (i)-(iii) e $\frac{\partial f}{\partial x_j}(x) = Df(x, e_j)$.
4. ($\Leftarrow$) Se valgono (i)-(iii), per i passi 1 e 2 i limiti destro e sinistro esistono e valgono $Df(x, e_j) = -Df(x, -e_j)$: la derivata parziale esiste.

□

Definizione 3.6 – Funzione differenziabile

$f : \mathbb{R}^n \to \mathbb{R}$ è **differenziabile** in $x \in \mathbb{R}^n$ se esiste un vettore $a \in \mathbb{R}^n$ tale che

$$\lim_{d \to 0} \frac{f(x + d) - f(x) - a^{\mathsf{T}} d}{\|d\|} = 0.$$

Corollario 3.11 – Funzioni di classe C^1

Sia $f : D \to \mathbb{R}$, con D sottoinsieme aperto di $\mathbb{R}^n$. Se f è di classe C^1 su D, cioè $f \in C^1(D)$, con il gradiente che esiste ed è continuo su D, allora f è differenziabile in ogni punto di D.

Solo enunciato

Derivate del secondo ordine di una funzione reale

Proposizione 3.25 – Funzioni due volte differenziabili e matrice hessiana

Siano $f : \mathbb{R}^n \to \mathbb{R}$ e $x \in \mathbb{R}^n$. Si ricorda che f è due volte differenziabile in x se f è differenziabile in x ed esiste una matrice $H \in \mathbb{R}^{n\times n}$ tale che

$$f(x+d) = f(x) + \nabla f(x)^\mathsf{T} d + \frac{1}{2} d^\mathsf{T} H\, d + \beta(x,d) \qquad \text{con} \qquad \lim_{d\to 0} \frac{\beta(x,d)}{\|d\|^2} = 0.$$

(i) **Condizione necessaria.** Se f è due volte differenziabile in x, allora ∇f esiste ed è continuo in x, esiste la matrice hessiana $\nabla^2 f(x)$, è simmetrica e $H = \nabla^2 f(x)$.

(ii) **Condizione sufficiente.** Se $\nabla^2 f$ esiste nei punti vicini a x ed è continua in x, allora f è due volte differenziabile in x.

Solo enunciato

Teorema della media e formula di Taylor

Teorema 3.1 – Teorema della media

Siano $D \subseteq \mathbb{R}^n$ aperto e convesso e $f \in C^1(D)$. Per ogni $x, y \in D$ esiste $\theta \in (0,1)$ tale che, con $z = x + \theta\,(y - x)$,

$$f(y) = f(x) + \nabla f(z)^\mathsf{T} (y - x).$$

Solo enunciato

Teorema 3.3 – Teorema di Taylor

Siano $D \subseteq \mathbb{R}^n$ aperto e convesso e $f \in C^2(D)$. Per ogni $x, y \in D$ esiste $\theta \in (0,1)$ tale che, con $z = x + \theta\,(y - x)$,

$$f(y) = f(x) + \nabla f(x)^\mathsf{T} (y - x) + \frac{1}{2} (y - x)^\mathsf{T} \nabla^2 f(z)\, (y - x).$$

Solo enunciato

4 Convessità di insiemi e funzioni

Insiemi convessi

Definizione 4.7 – Insieme convesso

$C \subseteq \mathbb{R}^n$ è **convesso** se contiene il segmento fra due suoi punti qualsiasi:

$$(1-\lambda)\,x + \lambda\,y \in C \qquad \forall x, y \in C,\ \forall \lambda \in [0,1].$$

Proposizione 4.12 – Convessità di una palla

Siano $\|\cdot\|$ una norma qualsiasi di $\mathbb{R}^n$, $\bar{x} \in \mathbb{R}^n$ e $\rho > 0$. La palla aperta e la palla chiusa

$$\mathcal{B}_\rho(\bar{x}) = \{x \in \mathbb{R}^n : \|x - \bar{x}\| < \rho\}, \qquad \overline{\mathcal{B}}_\rho(\bar{x}) = \{x \in \mathbb{R}^n : \|x - \bar{x}\| \leq \rho\}$$

sono insiemi convessi.

Solo enunciato

Funzioni convesse e funzioni concave

Definizione 4.13 – Funzione convessa e strettamente convessa

Siano $S \subseteq \mathbb{R}^n$ convesso e $f : S \to \mathbb{R}$. f è **convessa** su S se per ogni $x, y \in S$ e ogni $\lambda \in [0,1]$

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

è **strettamente convessa** se la disuguaglianza è stretta per ogni $x \neq y$ e ogni $\lambda \in (0,1)$. f è (strettamente) **concava** se $-f$ è (strettamente) convessa.

Convessità e concavità di funzioni differenziabili

Proposizione 4.17 – Condizione del primo ordine di convessità

Siano $S \subseteq \mathbb{R}^n$ aperto e convesso e $f \in C^1(S)$. Allora f è convessa su S se e solo se

$$f(y) \geq f(x) + \nabla f(x)^{\mathsf{T}}(y - x) \qquad \forall x, y \in S.$$

Dimostrazione da fare

Dimostrazione

Condizione necessaria. Sia f convessa e siano $x, y \in S$.

1. Per $\lambda \in (0, 1]$, scrivendo $(1-\lambda)\, x + \lambda\, y = x + \lambda(y - x)$, la definizione dà $f(x + \lambda(y-x)) \leq f(x) + \lambda\big(f(y) - f(x)\big)$.
2. Si porta $f(x)$ a sinistra e si divide per $\lambda > 0$:
$$\frac{f(x + \lambda(y - x)) - f(x)}{\lambda} \leq f(y) - f(x).$$
3. Per $\lambda \to 0^+$ il primo membro tende alla derivata direzionale lungo $d = y - x$, che per f differenziabile (Corollario 3.11) vale $\nabla f(x)^{\mathsf{T}}(y - x)$. Il secondo membro non dipende da λ, quindi $\nabla f(x)^{\mathsf{T}}(y - x) \leq f(y) - f(x)$.

Condizione sufficiente. Siano $x, y \in S$, $\lambda \in [0, 1]$ e $z = (1 - \lambda)\, x + \lambda\, y$, che sta in S perché S è convesso.

1. L'ipotesi applicata alle coppie (z, x) e (z, y) dà
$$f(x) \geq f(z) + \nabla f(z)^{\mathsf{T}}(x - z), \qquad f(y) \geq f(z) + \nabla f(z)^{\mathsf{T}}(y - z).$$
2. Si moltiplica la prima per $1 - \lambda \geq 0$, la seconda per $\lambda \geq 0$ e si somma:
$$(1 - \lambda) f(x) + \lambda f(y) \geq f(z) + \nabla f(z)^{\mathsf{T}}\big((1 - \lambda)\, x + \lambda\, y - z\big) = f(z),$$
perché $(1 - \lambda)\, x + \lambda\, y - z = \mathbf{0}$. È la definizione di convessità.

□

Proposizione 4.18 – Condizione del primo ordine di stretta convessità

Siano $S \subseteq \mathbb{R}^n$ aperto e convesso e $f \in C^1(S)$. Allora f è strettamente convessa su S se e solo se

$$f(y) > f(x) + \nabla f(x)^{\mathsf{T}}(y - x) \qquad \forall x, y \in S,\ x \neq y.$$

Solo enunciato

Proposizione 4.19 – Condizione del secondo ordine di convessità

Siano $S \subseteq \mathbb{R}^n$ aperto e convesso e $f \in C^2(S)$. Allora f è convessa su S se e solo se $\nabla^2 f(x)$ è semidefinita positiva per ogni $x \in S$.

Solo enunciato

Proposizione 4.20 – Convessità di una funzione quadratica

Sia $f(x) = \frac{1}{2} x^\mathsf{T} Q\, x + c^\mathsf{T} x + b$, con Q matrice simmetrica $n \times n$, $c \in \mathbb{R}^n$ e $b \in \mathbb{R}$. Allora f è convessa su $\mathbb{R}^n$ se e solo se Q è semidefinita positiva.

Solo enunciato

Proposizione 4.21 – Condizione sufficiente di convessità stretta

Siano $S \subseteq \mathbb{R}^n$ aperto e convesso e $f \in C^2(S)$. Se $\nabla^2 f(x)$ è definita positiva per ogni $x \in S$, allora f è strettamente convessa su S.

Dimostrazione da fare

Dimostrazione

Siano $x, y \in S$ con $x \neq y$.

1. Per il teorema di Taylor (Teorema 3.3), con $z = x + \theta\,(y - x)$ e $\theta \in (0, 1)$,
$$f(y) = f(x) + \nabla f(x)^\mathsf{T}(y - x) + \frac{1}{2}\,(y - x)^\mathsf{T} \nabla^2 f(z)\,(y - x).$$
2. $z \in S$ perché S è convesso, quindi $\nabla^2 f(z)$ è definita positiva e, poiché $y - x \neq \mathbf{0}$, l'ultimo termine è positivo.
3. Allora $f(y) > f(x) + \nabla f(x)^\mathsf{T}(y - x)$, e f è strettamente convessa per la Proposizione 4.18.

□

Proposizione 4.22 – Convessità stretta di una funzione quadratica

Sia $f(x) = \frac{1}{2} x^\mathsf{T} Q\, x + c^\mathsf{T} x + b$, con Q matrice simmetrica $n \times n$, $c \in \mathbb{R}^n$ e $b \in \mathbb{R}$. Allora f è strettamente convessa su $\mathbb{R}^n$ se e solo se Q è definita positiva.

Solo enunciato

Insiemi convessi definiti mediante equazioni e disequazioni

Proposizione 4.29 – Insieme definito da disequazioni convesse ed equazioni affini

Siano $g_1, \dots, g_m : \mathbb{R}^n \to \mathbb{R}$ funzioni convesse, e $a_j \in \mathbb{R}^n$, $b_j \in \mathbb{R}$ per $j = 1, \dots, p$. Allora l'insieme

$$S = \{x \in \mathbb{R}^n : g_i(x) \leq 0,\ i = 1, \dots, m;\ \ a_j^\top x = b_j,\ j = 1, \dots, p\}$$

è convesso.

Solo enunciato

Appendice: elementi di algebra lineare

Proposizione 4.34 – Criterio di Sylvester

Sia Q una matrice reale e simmetrica $n \times n$, e siano Δ_i, per $i = 1, \dots, n$, i minori principali di ordine i, cioè i determinanti delle sottomatrici formate dalle prime i righe e dalle prime i colonne di Q. Allora:

1. Q è definita positiva se e solo se $\Delta_i > 0$, $i = 1, \dots, n$;
2. Q è definita negativa se e solo se $(-1)^i \Delta_i > 0$, $i = 1, \dots, n$.

Solo enunciato

5 Massimi e minimi di funzioni di più variabili

Definizione 5.1-5.2 – Punto di minimo globale e locale

Siano $S \subseteq \mathbb{R}^n$ e $f : S \to \mathbb{R}$. Un punto $x^* \in S$ è un **punto di minimo globale** di f su S se

$$f(x^*) \leq f(x) \qquad \forall x \in S.$$

Il valore $f(x^*)$ è il **minimo globale**. Il punto è di **minimo globale stretto** se $f(x^*) < f(x)$ per ogni $x \in S$ con $x \neq x^*$.

Un punto $x^* \in S$ è un **punto di minimo locale** di f su S se esiste $\rho > 0$ tale che

$$f(x^*) \leq f(x) \qquad \forall x \in \mathcal{B}_\rho(x^*) \cap S,$$

con $\mathcal{B}_\rho(x^*) = \{y \in \mathbb{R}^n : \|y - x^*\| < \rho\}$. È di **minimo locale stretto** se $f(x^*) < f(x)$ per ogni $x \in \mathcal{B}_\rho(x^*) \cap S$ con $x \neq x^*$.

Le definizioni di massimo, globale e locale, stretto e non, si ottengono invertendo le disuguaglianze.

Coincidenza di minimi locali e globali nel caso convesso

Proposizione 5.5 – Punti di minimo nel caso convesso

Siano $S \subseteq \mathbb{R}^n$ convesso e f convessa su S.

- Ogni punto di minimo locale di f su S è di minimo globale.
- L'insieme dei punti di minimo globale di f su S è convesso.

Solo enunciato

Proposizione 5.6 – Unicità del punto di minimo

Siano $S \subseteq \mathbb{R}^n$ convesso e f strettamente convessa su S. Se x^* è un punto di minimo locale di f su S, allora x^* è l'unico punto di minimo locale e l'unico punto di minimo globale di f su S.

Solo enunciato

Condizioni di esistenza della soluzione

Proposizione 5.7 – Teorema di Weierstrass

Siano $S \subseteq \mathbb{R}^n$ non vuoto e compatto e $f : \mathbb{R}^n \to \mathbb{R}$ continua su S. Allora esistono $x_{\min}, x_{\max} \in S$ tali che

$$f(x_{\min}) \leq f(x) \leq f(x_{\max}) \qquad \forall x \in S.$$

Dimostrazione da fare

Dimostrazione

Basta il minimo, perché il massimo di f è il minimo della funzione continua $-f$. Sia $L = \inf_{x \in S} f(x)$, con $L = -\infty$ se f non è limitata inferiormente su S.

1. Per la definizione di estremo inferiore esiste una successione $\{x_k\}$ con $x_k \in S$ e $f(x_k) \to L$.
2. S è compatto, quindi, per il Corollario 2.26, esistono una sottosuccessione $\{x_k\}_{k \in K}$ e un punto $x^* \in S$ con $\lim_{k \in K,\, k \to \infty} x_k = x^*$. Il limite sta in S perché S è chiuso, ma la convergenza vale solo lungo la sottosuccessione.
3. f è continua in x^*, quindi $f(x_k) \to f(x^*)$ per $k \in K$, $k \to \infty$.
4. Per $k \in K$ vale anche $f(x_k) \to L$, perché è una sottosuccessione di una successione che tende a L. Per l'unicità del limite $f(x^*) = L$.
5. Allora L è finito ed è un minorante dei valori di f su S: $f(x^*) = L \leq f(x)$ per ogni $x \in S$. Quindi x^* è un punto di minimo globale.

□

Condizioni di ottimalità nel caso di problemi non vincolati: direzioni di discesa

Definizione 5.8 – Direzione di discesa

Siano $f : \mathbb{R}^n \to \mathbb{R}$ e $x \in \mathbb{R}^n$. Un vettore $d \in \mathbb{R}^n$, $d \neq \mathbf{0}$, è una **direzione di discesa** per f in x se esiste $\bar{t} > 0$ tale che

$$f(x + t\,d) < f(x) \qquad \forall t \in (0, \bar{t}\,).$$

Proposizione 5.9 – Condizione di discesa del primo ordine

Siano $f : \mathbb{R}^n \to \mathbb{R}$, $x \in \mathbb{R}^n$ e $d \in \mathbb{R}^n$ un vettore non nullo.

(a) Se f è differenziabile in x e
$$\nabla f(x)^\mathsf{T} d < 0,$$
allora d è una direzione di discesa per f in x.

(b) Se $f \in C^1(\mathbb{R}^n)$ è convessa su una palla aperta $\mathcal{B}_\rho(x)$ e d è una direzione di discesa per f in x, allora $\nabla f(x)^\mathsf{T} d < 0$.

Dimostrazione da fare

Dimostrazione

(a)

1. Per la derivata direzionale di una funzione differenziabile $\lim_{t\to 0^+} \dfrac{f(x + t\,d) - f(x)}{t} = \nabla f(x)^\mathsf{T} d < 0$.
2. Per la permanenza del segno esiste $\bar{t} > 0$ tale che $\dfrac{f(x + t\,d) - f(x)}{t} < 0$ per ogni $t \in (0, \bar{t}\,)$. Poiché $t > 0$, il numeratore è negativo, cioè $f(x + t\,d) < f(x)$, e d è una direzione di discesa.

(b)

1. Per $t > 0$ abbastanza piccolo $f(x + t\,d) < f(x)$ e, riducendo t, anche $t\,\|d\| < \rho$, cioè $x + t\,d \in \mathcal{B}_\rho(x)$.
2. La palla $\mathcal{B}_\rho(x)$ è aperta e convessa (Proposizione 4.12) e f è su di essa C^1 e convessa, quindi il criterio del primo ordine (Proposizione 4.17) dà $f(x+t\,d) \geq f(x)+t\,\nabla f(x)^\mathsf{T} d$.
3. Insieme alla discesa, $f(x) + t\,\nabla f(x)^\mathsf{T} d \leq f(x + t\,d) < f(x)$, quindi $t\,\nabla f(x)^\mathsf{T} d < 0$. Poiché $t > 0$, $\nabla f(x)^\mathsf{T} d < 0$.

□

Proposizione – L'antigradiente è una direzione di discesa

Sia $f : \mathbb{R}^n \to \mathbb{R}$ differenziabile in $x \in \mathbb{R}^n$ con $\nabla f(x) \neq \mathbf{0}$. Allora l'**antigradiente** $d = -\nabla f(x)$ è una direzione di discesa per f in x.

Dimostrazione da fare

Dimostrazione

Si ha $d \neq \mathbf{0}$ e

$$\nabla f(x)^{\mathsf{T}} d = -\nabla f(x)^{\mathsf{T}} \nabla f(x) = -\|\nabla f(x)\|^2 < 0.$$

Quindi d è una direzione di discesa per la Proposizione 5.9(a). □

Proposizione 5.11 – Condizione di discesa del secondo ordine

Siano $f \in C^2(\mathbb{R}^n)$, $x \in \mathbb{R}^n$ e $d \in \mathbb{R}^n$ tali che

$$\nabla f(x)^{\mathsf{T}} d = 0 \qquad \text{e} \qquad d^{\mathsf{T}} \nabla^2 f(x)\, d < 0.$$

Allora d è una direzione di discesa per f in x.

Dimostrazione da fare

Dimostrazione

Si ha $d \neq \mathbf{0}$, perché $d^{\mathsf{T}} \nabla^2 f(x)\, d \neq 0$.

1. $f \in C^2$, quindi è due volte differenziabile (Proposizione 3.25(ii)). Per $t > 0$

$$f(x + t\,d) = f(x) + t\,\nabla f(x)^{\mathsf{T}} d + \frac{1}{2} t^2\, d^{\mathsf{T}} \nabla^2 f(x)\, d + \beta(x, t\,d),$$

   con $\beta(x, t\,d)/t^2 \to 0$ per $t \to 0^+$, perché $\|t\,d\|^2 = t^2 \|d\|^2$.
2. Il termine del primo ordine è nullo. Dividendo per $t^2 > 0$

$$\frac{f(x + t\,d) - f(x)}{t^2} = \frac{1}{2}\, d^{\mathsf{T}} \nabla^2 f(x)\, d + \frac{\beta(x, t\,d)}{t^2}.$$

3. Il primo addendo è negativo e non dipende da t, mentre il secondo tende a 0: per t abbastanza piccolo $|\beta(x, t\,d)|/t^2 < \frac{1}{2} |d^{\mathsf{T}} \nabla^2 f(x)\, d|$, quindi il secondo membro è negativo.
4. Allora $f(x + t\,d) < f(x)$ per ogni $t > 0$ abbastanza piccolo, quindi d è una direzione di discesa.

□

Condizioni di minimo locale

Proposizione 5.12 – Condizione necessaria di minimo locale

Sia $f : \mathbb{R}^n \to \mathbb{R}$. Se $x^* \in \mathbb{R}^n$ è un punto di minimo locale di f su $\mathbb{R}^n$, non esiste una direzione di discesa per f in x^*.

Dimostrazione da fare

Dimostrazione

Per assurdo, sia d di discesa in x^*, con il suo $\bar{t}$, e sia $\mathcal{B}_\rho(x^*)$ la palla della definizione di minimo locale.

1. Si sceglie $\hat{t}$ con $0 < \hat{t} < \bar{t}$ e $\hat{t}\,\|d\| < \rho$.
2. Allora $x^* + \hat{t}\,d \in \mathcal{B}_\rho(x^*)$ e, per la discesa, $f(x^* + \hat{t}\,d) < f(x^*)$.
3. Questo contraddice $f(x^*) \leq f(x)$ per ogni $x \in \mathcal{B}_\rho(x^*)$.

□

Proposizione 5.13 – Condizione necessaria del primo ordine

Siano $f \in C^1(\mathbb{R}^n)$ e x^* un punto di minimo locale di f su $\mathbb{R}^n$. Allora

$$\nabla f(x^*) = \mathbf{0},$$

cioè x^* è un punto stazionario di f.

Dimostrazione da fare

Dimostrazione

Per assurdo, sia $\nabla f(x^*) \neq \mathbf{0}$.

1. L'antigradiente $d = -\nabla f(x^*)$ dà $\nabla f(x^*)^\mathsf{T} d = -\|\nabla f(x^*)\|^2 < 0$.
2. Quindi d è una direzione di discesa in x^* per la Proposizione 5.9.
3. Questo contraddice la Proposizione 5.12.

□

Proposizione 5.14 – Condizione necessaria del secondo ordine

Siano $f \in C^2(\mathbb{R}^n)$ e x^* un punto di minimo locale di f su $\mathbb{R}^n$. Allora

(a) $\nabla f(x^*) = \mathbf{0}$;
(b) $\nabla^2 f(x^*)$ è semidefinita positiva, cioè $y^\mathsf{T} \nabla^2 f(x^*)\, y \geq 0$ per ogni $y \in \mathbb{R}^n$.

Dimostrazione da fare

Dimostrazione

La (a) è la condizione del primo ordine (Proposizione 5.13). Per la (b), per assurdo $\nabla^2 f(x^*)$ non sia semidefinita positiva.

1. Per la definizione esiste d con $d^\mathsf{T} \nabla^2 f(x^*)\, d < 0$.
2. Per la (a), $\nabla f(x^*)^\mathsf{T} d = 0$.
3. Per la discesa del secondo ordine (Proposizione 5.11) d è di discesa in x^*, contro la Proposizione 5.12.

□

Proposizione 5.15 – Condizione sufficiente: convessità locale

Siano $f \in C^2(\mathbb{R}^n)$ e $\bar{x} \in \mathbb{R}^n$ con $\nabla f(\bar{x}) = \mathbf{0}$. Se $\nabla^2 f(x)$ è semidefinita positiva per ogni x di una palla $\mathcal{B}_\rho(\bar{x})$, allora $\bar{x}$ è un punto di minimo locale di f su $\mathbb{R}^n$.

Solo enunciato

Proposizione 5.16 – Condizione sufficiente del secondo ordine

Siano $f \in C^2(\mathbb{R}^n)$ e $\bar{x} \in \mathbb{R}^n$ con $\nabla f(\bar{x}) = \mathbf{0}$. Se $\nabla^2 f(\bar{x})$ è definita positiva, allora $\bar{x}$ è un punto di minimo locale di f su $\mathbb{R}^n$.

Solo enunciato

Condizioni di minimo globale nel caso convesso

Proposizione 5.17 – Condizione di minimo globale: caso convesso

Siano $f \in C^1(\mathbb{R}^n)$ convessa su $\mathbb{R}^n$ e $\bar{x} \in \mathbb{R}^n$. Allora $\bar{x}$ è un punto di minimo globale di f se e solo se

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

Dimostrazione da fare

Dimostrazione

1. **Solo se.** Un punto di minimo globale è di minimo locale, quindi $\nabla f(\bar{x}) = \mathbf{0}$ per la Proposizione 5.13.
2. **Se.** $\mathbb{R}^n$ è aperto e convesso e f è C^1 e convessa, quindi il criterio del primo ordine (Proposizione 4.17) dà, per ogni $y \in \mathbb{R}^n$,

$$f(y) \geq f(\bar{x}) + \nabla f(\bar{x})^\mathsf{T}(y - \bar{x}) = f(\bar{x}).$$

Quindi $\bar{x}$ è di minimo globale.

□

Proposizione 5.18 – Minimizzazione di una funzione quadratica

Sia $f(x) = \frac{1}{2}\, x^\mathsf{T} Q\, x + c^\mathsf{T} x$, con Q simmetrica e $c \in \mathbb{R}^n$. Allora:

(a) f ha un punto di minimo globale se e solo se Q è semidefinita positiva ed esiste x^* tale che $Q\, x^* + c = \mathbf{0}$;

(b) f ha un unico punto di minimo globale se e solo se Q è definita positiva.

Solo enunciato

Riepilogo

Dimostrazioni da fare

Proposizione 1.36 – Caratterizzazione degli insiemi chiusi 2
Teorema 2.25 – Bolzano-Weierstrass per successioni 4
Teorema 2.27 – Bolzano-Weierstrass per insiemi (anche Teorema 1.5) 5
Lemma 2.33 – Lipschitz implica continua 6
Proposizione 3.5 – Derivata parziale e derivate direzionali 8
Proposizione 4.17 – Condizione del primo ordine di convessità 11
Proposizione 4.21 – Condizione sufficiente di convessità stretta 12
Proposizione 5.7 – Teorema di Weierstrass 15

Proposizione 5.9 – Condizione di discesa del primo ordine · · · · · 16
Proposizione – L'antigradiente è una direzione di discesa · · · · · 17
Proposizione 5.11 – Condizione di discesa del secondo ordine · · · · · 17
Proposizione 5.12 – Condizione necessaria di minimo locale · · · · · 18
Proposizione 5.13 – Condizione necessaria del primo ordine · · · · · 18
Proposizione 5.14 – Condizione necessaria del secondo ordine · · · · · 19
Proposizione 5.17 – Condizione di minimo globale: caso convesso · · · · · 20

Solo enunciato

Proposizione 1.16 – Disuguaglianza di Cauchy-Schwarz · · · · · 1
Proposizione 2.14 – Successioni limitate e componenti · · · · · 4
Teorema 2.24 – Bolzano-Weierstrass in $\mathbb{R}$ · · · · · 4
Corollario 2.26 – Successioni in un compatto · · · · · 5
Corollario 3.11 – Funzioni di classe C^1 · · · · · 9
Proposizione 3.25 – Funzioni due volte differenziabili e matrice hessiana · · · · · 9
Teorema 3.1 – Teorema della media · · · · · 9
Teorema 3.3 – Teorema di Taylor · · · · · 9
Proposizione 4.12 – Convessità di una palla · · · · · 10
Proposizione 4.18 – Condizione del primo ordine di stretta convessità · · · · · 11
Proposizione 4.19 – Condizione del secondo ordine di convessità · · · · · 12
Proposizione 4.20 – Convessità di una funzione quadratica · · · · · 12
Proposizione 4.22 – Convessità stretta di una funzione quadratica · · · · · 12
Proposizione 4.29 – Insieme definito da disequazioni convesse ed equazioni affini · · · · · 13
Proposizione 4.34 – Criterio di Sylvester · · · · · 13
Proposizione 5.5 – Punti di minimo nel caso convesso · · · · · 14
Proposizione 5.6 – Unicità del punto di minimo · · · · · 14
Proposizione 5.15 – Condizione sufficiente: convessità locale · · · · · 19
Proposizione 5.16 – Condizione sufficiente del secondo ordine · · · · · 19
Proposizione 5.18 – Minimizzazione di una funzione quadratica · · · · · 20