11 - Introduzione al Calcolo Differenziale negli Spazi di Banach

Premesse

> Definizione: Segmento, Poligonale

Sia E uno spazio vettoriale.
Siano $\mathbf{x}, \mathbf{y} \in E$.

Si dice **segmento** di estremi $\mathbf{x}$ e $\mathbf{y}$ l'insieme
$[\mathbf{x}, \mathbf{y}] := \{\mathbf{x} + \lambda(\mathbf{y} - \mathbf{x}) \mid \lambda \in [0; 1]\}$.

Siano $\mathbf{x}_1, \dots, \mathbf{x}_n \in E$.

Si dice **poligonale** di vertici $\mathbf{x}_1, \dots, \mathbf{x}_n$ l'insieme
$[\mathbf{x}_1, \dots, \mathbf{x}_n] := [\mathbf{x}_1, \mathbf{x}_2] \cup \dots \cup [\mathbf{x}_{n-1}, \mathbf{x}_n]$.

I punti $\mathbf{x}_1$ e $\mathbf{x}_n$ sono detti **estremi** della poligonale.

Derivabilità secondo Gateaux e secondo Fréchet

⌘ **Definizione: Derivabilità e derivata, secondo Gateaux**

Siano $(X, \|\cdot\|_X)$ e $(Y, \|\cdot\|_Y)$ due spazi normati.
Sia $A \subseteq X$.

Sia $\mathbf{x}_0 \in \overset{\circ}{A}$.
Sia $f : A \to Y$ una funzione.

f si dice **derivabile secondo Gateaux** (o G-derivabile) in $\mathbf{x}_0$ quando esiste $\varphi \in \mathcal{L}(X, Y)$ tale che $\lim\limits_{\lambda \to 0^+} \frac{f(\mathbf{x}_0 + \lambda \mathbf{v}) - f(\mathbf{x}_0)}{\lambda} = \varphi(\mathbf{v})$ per ogni $\mathbf{v} \in X$.

Tale φ è unico (per unicità del limite); esso prende il nome di **derivata secondo Gateaux** di f in $\mathbf{x}_0$, e si denota con $f'(\mathbf{x}_0)$.

Osservazione

Siano $(X, \|\cdot\|_X)$ e $(Y, \|\cdot\|_Y)$ due spazi normati.
Sia $A \subseteq X$.
Sia $\mathbf{x}_0 \in \overset{\circ}{A}$.
Sia $f : A \to Y$ una funzione G-derivabile in $\mathbf{x}_0$.

Allora, per ogni $\mathbf{v} \in X$ si ha $\lim\limits_{\lambda \to 0} \frac{f(\mathbf{x}_0 + \lambda \mathbf{v}) - f(\mathbf{x}_0)}{\lambda} = f'(\mathbf{x}_0)(\mathbf{v})$.

Infatti, si ha $\lim\limits_{\lambda \to 0^+} \frac{f(\mathbf{x}_0 + \lambda \mathbf{v}) - f(\mathbf{x}_0)}{\lambda} = f'(\mathbf{x}_0)(\mathbf{v})$ per definizione di derivata secondo Gateaux in $\mathbf{x}_0$, che esiste per ipotesi.

D'altra parte, si ha
$\lim\limits_{\lambda \to 0^-} \frac{f(\mathbf{x}_0 + \lambda \mathbf{v}) - f(\mathbf{x}_0)}{\lambda} = \lim\limits_{\lambda \to 0^+} \frac{f(\mathbf{x}_0 - \lambda \mathbf{v}) - f(\mathbf{x}_0)}{-\lambda} = -\lim\limits_{\lambda \to 0^+} \frac{f(\mathbf{x}_0 + \lambda(-\mathbf{v})) - f(\mathbf{x}_0)}{\lambda} = -f'(\mathbf{x}_0)(-\mathbf{v})$, che per linearità di $f'(\mathbf{x}_0)$ è pari a $f'(\mathbf{x}_0)(\mathbf{v})$.

⌘ Definizione: Derivabilità e derivata, secondo Fréchet

Siano $(X, \|\cdot\|_X)$ e $(Y, \|\cdot\|_Y)$ due spazi normati.
Sia $A \subseteq X$.

Sia $\mathbf{x}_0 \in \overset{\circ}{A}$.
Sia $f : A \to Y$ una funzione.

f si dice **derivabile secondo Fréchet** (o F-derivabile) in $\mathbf{x}_0$ quando esiste $\psi \in \mathcal{L}(X, Y)$ tale che $\lim_{\mathbf{v} \to \mathbf{0}_X} \frac{f(\mathbf{x}_0+\mathbf{v})-f(\mathbf{x}_0)-\psi(\mathbf{v})}{\|\mathbf{v}\|_X} = \mathbf{0}_Y$.

Un tale ψ prende il nome di **derivata secondo Fréchet** di f in $\mathbf{x}_0$.

Proposizione 11.1: F-derivabilità implica G-derivabilità

Siano $(X, \|\cdot\|_X)$ e $(Y, \|\cdot\|_Y)$ due spazi normati.
Sia $A \subseteq X$.
Sia $\mathbf{x}_0 \in \overset{\circ}{A}$.
Sia $f : A \to Y$ una funzione F-derivabile in $\mathbf{x}_0$.

Si hanno i seguenti fatti:

- f è G-derivabile in $\mathbf{x}_0$;
- La derivata secondo Frechet di f in $\mathbf{x}_0$ è unica, e coincide con $f'(\mathbf{x}_0)$.

Dimostrazione

Sia ψ una derivata secondo Fréchet di f in $\mathbf{x}_0$, che esiste per ipotesi di F-derivabilità.
Sia $\mathbf{v} \in X$.

Se $\mathbf{v} = \mathbf{0}_X$, si ha $\psi(\mathbf{0_X}) = \mathbf{0}_Y$ per linearità; d'altra parte, si ha $\lim_{\lambda \to 0^+} \frac{f(\mathbf{x}_0+\lambda\mathbf{0}_X)-f(\mathbf{x}_0)}{\lambda} = \lim_{\lambda \to 0^+} \frac{\mathbf{0}_Y}{\lambda} = \mathbf{0}_Y$, per cui $f'(\mathbf{x}_0)(\mathbf{0}_X) = \mathbf{0}_Y$.

Ne segue che $\psi(\mathbf{0}_X) = f'(\mathbf{x}_0)(\mathbf{0}_X)$.

Si supponga ora $\mathbf{v} \neq \mathbf{0}_X$, cosicché $\|\mathbf{v}\|_X \neq 0$.

Si ha $\lim\limits_{\mathbf{u}\to\mathbf{0}_X} \frac{f(\mathbf{x}_0+\mathbf{u})-f(\mathbf{x}_0)-\psi(\mathbf{u})}{\|\mathbf{u}\|_X} = \mathbf{0}_Y$ per definizione di ψ;

restringendo tale espressione ai vettori del tipo $\lambda\mathbf{v}$, con $\lambda > 0$, si ottiene che $\mathbf{0}_X$ è ancora punto di accumulazione per questo insieme, e il limite di sopra effettuato sulla restrizione, che dunque ha senso, è lo stesso.
Risulta cioè che

$$\lim_{\lambda\to 0^+} \frac{f(\mathbf{x}_0+\lambda\mathbf{v})-f(\mathbf{x}_0)-\psi(\lambda\mathbf{v})}{\|\lambda\mathbf{v}\|_X} = \mathbf{0}_Y.$$

Si osservi ora che $\frac{f(\mathbf{x}_0+\lambda\mathbf{v})-f(\mathbf{x}_0)-\psi(\lambda\mathbf{v})}{\|\lambda\mathbf{v}\|_X} = \frac{1}{\|\mathbf{v}\|_X}\left(\frac{f(\mathbf{x}_0+\lambda\mathbf{v})-f(\mathbf{x}_0)}{\lambda} - \psi(\mathbf{v})\right)$ per ogni $\lambda > 0$, per linearità di ψ e per assoluta omogeneità di $\|\cdot\|_X$.

Si ha allora che $\lim\limits_{\lambda\to 0^+} \frac{f(\mathbf{x}_0+\lambda\mathbf{v})-f(\mathbf{x}_0)}{\lambda} = \lim\limits_{\lambda\to 0^+} \|\mathbf{v}\|_X \cdot \underbrace{\frac{f(\mathbf{x}_0+\lambda\mathbf{v})-f(\mathbf{x}_0)-\psi(\lambda\mathbf{v})}{\|\lambda\mathbf{v}\|_X}}_{\to\mathbf{0}_Y} + \psi(\mathbf{v}) = \psi(\mathbf{v}).$

Dall'arbitrarietà di $\mathbf{v} \in X$ segue allora che f è G-derivabile in $\mathbf{x}_0$, e $f'(\mathbf{x}_0)(\mathbf{v}) = \psi(\mathbf{v})$ per ogni $\mathbf{v} \in X$, da cui viene l'unicità di ψ per unicità della derivata secondo Gateaux.
■

In virtù della proposizione precedente, la derivata secondo Gateaux e secondo Fréchet verranno denominate semplicemente "derivata".

Osservazione: Operatori lineari continui sono F-derivabili

Siano $(X, \|\cdot\|_X)$ e $(Y, \|\cdot\|_Y)$ due spazi normati.
Sia $T \in \mathcal{L}(X, Y)$.

Allora, T è F-derivabile in X, con derivata pari a T.

Infatti, per ogni $\mathbf{v} \in X \setminus \{\mathbf{0}\}$ si ha $\frac{T(\mathbf{x}+\mathbf{v})-T(\mathbf{x})-T(\mathbf{v})}{\|\mathbf{v}\|_X} = \mathbf{0}_Y$ per linearità di T.

Ne segue quindi che $\lim\limits_{\mathbf{v}\to\mathbf{0}_X} \frac{T(\mathbf{x}+\mathbf{v})-T(\mathbf{x})-T(\mathbf{v})}{\|\mathbf{v}\|_X} = \mathbf{0}_Y$.

■

Proposizione 11.2: F-Derivabilità implica continuità

Siano $(X, \|\cdot\|_X)$ e $(Y, \|\cdot\|_Y)$ due spazi normati.
Sia $A \subseteq X$.
Sia $\mathbf{x}_0 \in \overset{\circ}{A}$.
Sia $f : A \to Y$ una funzione F-derivabile in $\mathbf{x}_0$.

Allora, f è continua in $\mathbf{x}_0$.

Dimostrazione

Si ha $\lim\limits_{\mathbf{v}\to\mathbf{0}_X} \frac{f(\mathbf{x}_0+\mathbf{v})-f(\mathbf{x}_0)-f'(\mathbf{x}_0)(\mathbf{v})}{\|\mathbf{v}\|_X} = \mathbf{0}_Y$ per ipotesi.

Per acquisire la continuità di f in $\mathbf{x}_0$, si mostri che $\lim\limits_{\mathbf{v}\to\mathbf{0}_X} f(\mathbf{x}_0+\mathbf{v}) - f(\mathbf{x}_0) = \mathbf{0}_Y$.

Si osservi intanto che $\lim\limits_{\mathbf{v}\to\mathbf{0}_X} f'(\mathbf{x}_0)(\mathbf{v}) = f'(\mathbf{x}_0)(\mathbf{0}_X) = \mathbf{0}_Y$, essendo $f'(\mathbf{x}_0) \in \mathcal{L}(X, Y)$, cioè lineare e continuo.

Si ha $\lim\limits_{\mathbf{v}\to\mathbf{0}_X} f(\mathbf{x}_0+\mathbf{v}) - f(\mathbf{x}_0) = \lim\limits_{\mathbf{v}\to\mathbf{0}_X} \|\mathbf{v}\|_X \underbrace{\frac{f(\mathbf{x}_0+\mathbf{v})-f(\mathbf{x}_0)-f'(\mathbf{x}_0)(\mathbf{v})}{\|\mathbf{v}\|_X}}_{\to\mathbf{0}_Y} + \underbrace{f'(\mathbf{x}_0)(\mathbf{v})}_{\to\mathbf{0}_Y} = \mathbf{0}_Y$.

■

Osservazione

La G-derivabilità non implica la continuità (dunque neanche la F-derivabilità, per il teorema precedente).

Ad esempio, si definisca $f : \mathbb{R}^2 \to \mathbb{R}$ ponendo $f(x,y) = \begin{cases} \frac{y^3}{x}, & x \neq 0 \\ 0, & x = 0 \end{cases}$.

f è G-derivabile in $(0,0)$.

Fissati infatti $(u,v) \in \mathbb{R}^2$ e $\lambda \neq 0$, si ha $\frac{f\big((0,0)+\lambda(u,v)\big)-f(0,0)}{\lambda} = \begin{cases} \lambda\frac{v^3}{u}, & u \neq 0 \\ 0, & u = 0 \end{cases}$; ne segue che $\lim\limits_{\lambda \to 0} \frac{f\big((0,0)+\lambda(u,v)\big)-f(0,0)}{\lambda} = \mathbf{0}_Y$ per ogni $(u,v) \in \mathbb{R}^2$.

Pertanto, $f'(0,0) = \mathbf{0}_{(\mathbb{R}^2)^*}$.

f non è continua in $(0,0)$.

Infatti, restringendo f all'insieme $\gamma_1 = \{(t,t) \mid t \in \mathbb{R}\}$ si ottiene $\lim\limits_{(x,y)\to(0,0)} f_{|\gamma_1}(x,y) = \lim\limits_{t\to 0} f(t,t) = \lim\limits_{t \to 0} \frac{t^3}{t} = 0$.

D'altra parte, restringendo f all'insieme $\gamma_2 = \{(t^3,t) \mid t \in \mathbb{R}\}$ si ottiene $\lim\limits_{(x,y)\to(0,0)} f_{|\gamma_2}(x,y) = \lim\limits_{t\to 0} f(t^3,t) = \lim\limits_{t \to 0} \frac{t^3}{t^3} = 1$.

Osservazione

Sia $(X, \langle\cdot,\cdot\rangle)$ uno spazio di Hilbert.
Sia $A \subseteq X$.
Sia $\mathbf{x}_0 \in \mathring{A}$.
Sia $f : A \to \mathbb{R}$ una funzione G-derivabile in $\mathbf{x}_0$.

Per il [Teorema 9.15], esiste un unico $\tilde{\mathbf{x}} \in X$ che identifica $f'(\mathbf{x}_0)$, tramite l'isometria lineare biunivoca $X \to X^* : \mathbf{x} \mapsto \langle\cdot,\mathbf{x}\rangle$.

Derivata di funzioni di variabile reale

⌘ Definizione: Derivabilità e derivata, per funzioni di variabile reale

Sia $(Y, \|\cdot\|)$ uno spazio normato.
Sia $I \subseteq \mathbb{R}$ un intervallo.
$x_0 \in I$.
Sia $f : I \to Y$ una funzione.

f si dice **derivabile** in x_0 quando esiste $\boldsymbol{\omega} \in Y$ tale che $\lim\limits_{\lambda \to 0} \frac{f(x_0+\lambda)-f(x_0)}{\lambda} = \boldsymbol{\omega}$.

Tale $\boldsymbol{\omega}$ è unico (per unicità del limite); esso prende il nome di **derivata** di f in x_0, e si denota con $\dot{f}(x_0)$.

Proposizione 11.3: Legame tra derivabilità e F-derivabilità

Sia $(Y, \|\cdot\|)$ uno spazio normato.
Sia $I \subseteq \mathbb{R}$ un intervallo.
$x_0 \in \overset{\circ}{I}$.
Sia $f : I \to Y$ una funzione.

Si hanno i seguenti fatti:

- f è derivabile in x_0 se e solo se f è F-derivabile in x_0;
- In caso di derivabilità, si ha $f'(x_0)(x) = \dot{f}(x_0)\, x$ per ogni $x \in \mathbb{R}$

Osservazioni preliminari

$f \in \mathcal{L}(\mathbb{R}, Y)$ se e solo se esiste $\boldsymbol{\alpha} \in Y$ tale che $f(x) = \boldsymbol{\alpha} x$ per ogni $x \in \mathbb{R}$.

Infatti, le funzioni di questo tipo sono lineari e continue.

Viceversa, data una funzione $f \in \mathcal{L}(\mathbb{R}, Y)$, si ha $f(x) = f(x \cdot 1) = x f(1)$ per ogni $x \in \mathbb{R}$ per linearità di f.
Dunque, $f(x) = \boldsymbol{\alpha} x$ per ogni $x \in \mathbb{R}$, con $\boldsymbol{\alpha} = f(1)$.

Dimostrazione

Si supponga f F-derivabile in x_0.
Sia $\boldsymbol{\alpha} \in Y$ tale che $f'(x_0)(x) = \boldsymbol{\alpha} x$ per ogni $x \in \mathbb{R}$ (che esiste per l'osservazione preliminare).

Si ha allora $\lim\limits_{h \to 0} \frac{f(x_0+h)-f(x_0)-f'(x_0)(h)}{|h|} = \mathbf{0}$, ossia $\lim\limits_{h \to 0} \frac{f(x_0+h)-f(x_0)-\boldsymbol{\alpha} h}{|h|} = \mathbf{0}$

Allora, si ha

$$\lim_{h \to 0} \frac{f(x_0+h)-f(x_0)(h)}{h} = \lim_{h \to 0} \frac{|h|}{h} \underbrace{\frac{f(x_0+h)-f(x_0)(h)-\boldsymbol{\alpha} h}{|h|}}_{\to \mathbf{0}} + \boldsymbol{\alpha} = \boldsymbol{\alpha}.$$

Dunque, f è derivabile in x_0, e $\dot{f}(x_0) = \boldsymbol{\alpha}$; dunque, $f'(x_0)(x) = \dot{f}(x_0)\, x$ per ogni $x \in \mathbb{R}$.

Viceversa, si supponga f derivabile in x_0.

Sia $\psi : \mathbb{R} \to Y$ definita ponendo $\psi(x) = \dot{f}(x_0)\, x$ per ogni $x \in \mathbb{R}$; si ha $\psi \in \mathcal{L}(\mathbb{R}, Y)$ per l'osservazione preliminare.

Si ha $\lim\limits_{h \to 0} \frac{f(x_0+h)-f(x_0)-\psi(h)}{|h|} = \lim\limits_{h \to 0} \frac{f(x_0+h)-f(x_0)-\dot{f}(x_0)h}{|h|} = \lim\limits_{h \to 0} \frac{h}{|h|} \underbrace{\left(\frac{f(x_0+h)-f(x_0)}{h} - \dot{f}(x_0) \right)}_{\to \mathbf{0}} = \mathbf{0}.$

Dunque, $\psi = f'(\mathbf{x}_0)$.
■

Chiusure convesse e teorema di Lagrange

⌘ Definizione: Involucro convesso

Sia E uno spazio vettoriale.
Sia $A \subseteq E$.

Si dice **involucro convesso** (o inviluppo convesso) di A l'insieme
$\mathrm{conv}(A) := \bigcap\{C \subseteq E : C \text{ convesso}, C \supseteq A\}$.

Osservazione

$\mathrm{conv}(A)$ è il minimo sottoinsieme convesso di E contenente A, rispetto all'inclusione.

Infatti, $\mathrm{conv}(A)$ è contenuto in ogni sottoinsieme convesso di E contenente A per definizione;
inoltre, esso è convesso in quanto l'intersezione arbitraria di insiemi convessi è convessa (segue direttamente dalla definizione di convessità).

⌘ Definizione: Chiusura convessa

Sia $(X, \|\cdot\|)$ uno spazio normato.
Sia $A \subseteq X$.

Si dice **chiusura convessa** di A l'insieme
$\overline{\mathrm{conv}}(A) := \bigcap\{C \subseteq E : C \text{ chiuso e convesso}, C \supseteq A\}$.

Osservazione 1

$\overline{\mathrm{conv}}(A)$ è il minimo sottoinsieme chiuso e convesso di E contenente A, rispetto all'inclusione.

Infatti, $\overline{\mathrm{conv}}(A)$ è contenuto in ogni sottoinsieme chiuso e convesso di E contenente A per definizione;
inoltre, esso è chiuso e convesso in quanto l'intersezione arbitraria di insiemi chiusi e convessi è chiusa e convessa.

Osservazione 2

$\overline{\text{conv}}(A) = \overline{\text{conv}(A)}$.

Infatti, $\overline{\text{conv}(A)}$ è chiuso, e convesso essendo chiusura di un insieme convesso (Si veda l'Osservazione 2 sulla convessità, capitolo 7).
Dunque, $\overline{\text{conv}(A)} \supseteq \overline{\text{conv}}(A)$.

D'altra parte, $\overline{\text{conv}}(A)$ è convesso e contiene A;
allora, $\overline{\text{conv}}(A) \supseteq \text{conv}(A)$.
Essendo $\overline{\text{conv}}(A)$ anche chiuso, contenendo $\text{conv}(A)$ si ha $\overline{\text{conv}}(A) \supseteq \overline{\text{conv}(A)}$.

!thm] Teorema 11.4: Teorema di Lagrange
Siano $(X, \|\cdot\|_X)$ e $(Y, \|\cdot\|_Y)$ due spazi normati.
Sia $A \subseteq X$.
Sia $f : A \to Y$ una funzione, continua su $A \setminus \overset{\circ}{A}$, e G-derivabile su $\overset{\circ}{A}$.
Siano $\mathbf{x}, \mathbf{z} \in A$ tali che $\mathbf{x} + \lambda(\mathbf{z} - \mathbf{x}) \in \overset{\circ}{A}$ per ogni $\lambda \in \left]0; 1\right[$.
Sia $C = \left\{f'(\mathbf{x} + \lambda(\mathbf{z} - \mathbf{x}))(\mathbf{z} - \mathbf{x}) \mid \lambda \in \left]0; 1\right[\right\}$.

Allora, $f(\mathbf{z}) - f(\mathbf{x}) \in \overline{\text{conv}}(C)$.

Dimostrazione

Si proceda per assurdo, supponendo $f(\mathbf{z}) - f(\mathbf{x}) \notin \overline{\text{conv}}(C)$.

Dalle osservazioni sulla chiusura convessa, si ha che $\overline{\text{conv}}(C)$ è chiuso e convesso.
$\{f(\mathbf{z}) - f(\mathbf{x})\}$ è compatto e convesso, e disgiunto da $\overline{\text{conv}}(C)$ per ipotesi di assurdo.

Per il Teorema di Separazione ([Teorema 7.9]), esiste allora $\varphi \in Y^*$ tale che
$\sup_{\mathbf{y} \in \overline{\text{conv}}(C)} \varphi(\mathbf{y}) < \varphi(f(\mathbf{z}) - f(\mathbf{x}))$.

In particolare, essendo $C \subseteq \overline{\text{conv}}(C)$, ne segue che $\varphi\big(f'(\mathbf{x} + \lambda(\mathbf{z} - \mathbf{x}))(\mathbf{z} - \mathbf{x})\big) < \varphi(f(\mathbf{z}) - f(\mathbf{x}))$ per ogni $\lambda \in \,]0;1[$.

Si definisca ora la funzione $\gamma : [0;1] \to \mathbb{R}$ ponendo $\gamma(\lambda) = \varphi\big(f\big(\mathbf{x} + \lambda(\mathbf{z} - \mathbf{x})\big)\big)$ per ogni $\lambda \in [0;1]$.

Si provi che γ è derivabile in $]0;1[$.

Si fissi dunque $\lambda_0 \in \,]0;1[$; fissato $h \in \mathbb{R}$ dimodoché $\lambda_0 + h \in \,]0;1[$, si ha

$$\begin{aligned}
\frac{\gamma(\lambda_0 + h) - \gamma(\lambda_0)}{h} &= \frac{\varphi\big(f\big(\mathbf{x} + (\lambda_0 + h)(\mathbf{z} - \mathbf{x})\big)\big) - \varphi\big(f\big(\mathbf{x} + \lambda_0(\mathbf{z} - \mathbf{x})\big)\big)}{h} && \text{Definizione di } \gamma \\
&= \frac{\varphi\big(f\big(\mathbf{x} + \lambda_0(\mathbf{z} - \mathbf{x}) + h(\mathbf{z} - \mathbf{x})\big)\big) - \varphi\big(f\big(\mathbf{x} + \lambda_0(\mathbf{z} - \mathbf{x})\big)\big)}{h} \\
&= \varphi\left(\frac{f\big(\mathbf{x} + \lambda_0(\mathbf{z} - \mathbf{x}) + h(\mathbf{z} - \mathbf{x})\big) - f\big(\mathbf{x} + \lambda_0(\mathbf{z} - \mathbf{x})\big)}{h}\right) && \text{Per linearità di } \varphi
\end{aligned}$$

Per ipotesi, $\mathbf{x} + \lambda_0(\mathbf{z} - \mathbf{x}) \in \mathring{A}$, e f è G-derivabile su $\mathring{A}$;

pertanto, $\lim_{h \to 0} \frac{f(\mathbf{x}+\lambda_0(\mathbf{z}-\mathbf{x})+h(\mathbf{z}-\mathbf{x}))-f(\mathbf{x}+\lambda_0(\mathbf{z}-\mathbf{x}))}{h} = f'(\mathbf{x} + \lambda_0(\mathbf{z} - \mathbf{x}))(\mathbf{z} - \mathbf{x})$.

Dalla continuità di φ (essendo $\varphi \in Y^*$) segue allora che

$$\lim_{h \to 0} \frac{\gamma(\lambda_0+h)-\gamma(\lambda_0)}{h} = \lim_{h \to 0} \varphi\left(\frac{f(\mathbf{x}+\lambda_0(\mathbf{z}-\mathbf{x})+h(\mathbf{z}-\mathbf{x}))-f(\mathbf{x}+\lambda_0(\mathbf{z}-\mathbf{x}))}{h}\right) = \varphi\big(f'(\mathbf{x} + \lambda_0(\mathbf{z} - \mathbf{x}))(\mathbf{z} - \mathbf{x})\big).$$

Dunque, è stato ricavato che γ è derivabile in $]0;1[$, e che
$\dot{\gamma}(\lambda) = \varphi\big(f'(\mathbf{x} + \lambda_0(\mathbf{z} - \mathbf{x}))(\mathbf{z} - \mathbf{x})\big)$ per ogni $\lambda \in \,]0;1[$.

Si ha inoltre che γ è continua in 0 e 1.

Infatti, la funzione $\begin{matrix}[0;1] \to \mathbb{R} \\ \lambda \mapsto f(\mathbf{x}+\lambda(\mathbf{z}-\mathbf{x}))\end{matrix}$ è continua in 0 e 1;
questo fatto si acquisisce osservando che, su $\mathbf{x}$ e su $\mathbf{z}$, la restrizione di f al segmento $[\mathbf{x},\mathbf{z}]$ è continua per G-derivabilità di f, se $\mathbf{x},\mathbf{z} \in \mathring{A}$, oppure per ipotesi di continuità di f su $A \setminus \mathring{A}$, se $\mathbf{x},\mathbf{z} \notin \mathring{A}$.

Essendo γ composizione di φ, continua, con tale funzione, segue la continuità di γ in 0 e 1.

Allora, γ soddisfa le ipotesi del teorema di Lagrange.
Pertanto, esiste $\tilde{\lambda} \in \,]0;1[$ tale che $\gamma(1) - \gamma(0) = \dot{\gamma}(\tilde{\lambda})$, ossia

$\varphi(f(\mathbf{z})) - \varphi(f(\mathbf{x})) = \varphi\big(f'(\mathbf{x}+\tilde{\lambda}(\mathbf{z}-\mathbf{x}))(\mathbf{z}-\mathbf{x})\big)$;

sfruttando la linearità di φ al primo membro, si ottiene allora

$\varphi(f(\mathbf{z}) - f(\mathbf{x})) = \varphi\big(f'(\mathbf{x}+\tilde{\lambda}(\mathbf{z}-\mathbf{x}))(\mathbf{z}-\mathbf{x})\big)$, in contraddizione con la disuguaglianza ottenuta inizialmente per costruzione di φ.
■

Nel caso in cui $Y = \mathbb{R}$, il teorema di Lagrange si può dimostrare senza procedere per assurdo:

Teorema 11.4*: Teorema di Lagrange (per funzioni a valori reali)

Sia $(X, \|\cdot\|_X)$ uno spazio normato.
Sia $A \subseteq X$.
Sia $f : A \to \mathbb{R}$ una funzione, continua su $A \setminus \mathring{A}$, e G-derivabile su $\mathring{A}$.
Siano $\mathbf{x},\mathbf{z} \in A$ tali che $\mathbf{x} + \lambda(\mathbf{z}-\mathbf{x}) \in \mathring{A}$ per ogni $\lambda \in \,]0;1[$.
Sia $C = \{f'(\mathbf{x}+\lambda(\mathbf{z}-\mathbf{x}))(\mathbf{z}-\mathbf{x}) \mid \lambda \in \,]0;1[\}$.

Allora, $f(\mathbf{z}) - f(\mathbf{x}) \in \overline{\mathrm{conv}}(C)$.

Dimostrazione

Sia $\gamma : [0;1] \to \mathbb{R}$ definita ponendo $\gamma(\lambda) = f(\mathbf{x} + \lambda(\mathbf{z} - \mathbf{x}))$ per ogni $\lambda \in [0;1]$.

Tale funzione è derivabile in $]0;1[$;

infatti, fissato $\lambda_0 \in \,]0;1[$, si ha $\mathbf{x} + \lambda_0(\mathbf{z} - \mathbf{x}) \in \mathring{A}$, e f G-derivabile su $\mathring{A}$ per ipotesi; pertanto,

$$\lim_{h \to 0} \frac{\gamma(\lambda_0+h)-\gamma(\lambda_0)}{h} = \lim_{h \to 0} \frac{f(\mathbf{x}+(\lambda_0+h)(\mathbf{z}-\mathbf{x}))-f(\mathbf{x}+\lambda_0(\mathbf{z}-\mathbf{x}))}{h} =$$

$$= \lim_{h \to 0} \frac{f(\mathbf{x}+\lambda_0(\mathbf{z}-\mathbf{x})+h(\mathbf{z}-\mathbf{x}))-f(\mathbf{x}+\lambda_0(\mathbf{z}-\mathbf{x}))}{h} = f'(\mathbf{x} + \lambda_0(\mathbf{z} - \mathbf{x}))(\mathbf{z} - \mathbf{x}).$$

Inoltre, γ è continua in 0 e 1;
questo fatto si acquisisce osservando che, su $\mathbf{x}$ e su $\mathbf{v}$, la restrizione di f al segmento $[\mathbf{x}, \mathbf{z}]$ è continua per G-derivabilità di f, se $\mathbf{x}, \mathbf{z} \in \mathring{A}$, oppure per ipotesi di continuità di f su $A \setminus \mathring{A}$, se $\mathbf{x}, \mathbf{z} \notin \mathring{A}$.

Allora, γ soddisfa le ipotesi del teorema di Lagrange.
Pertanto, esiste $\tilde{\lambda} \in \,]0;1[$ tale che $\gamma(1) - \gamma(0) = \dot{\gamma}(\tilde{\lambda})$, ossia

$$f(\mathbf{z}) - f(\mathbf{x}) = f'(\mathbf{x} + \tilde{\lambda}(\mathbf{z} - \mathbf{x}))(\mathbf{z} - \mathbf{x}) \in C \subseteq \overline{\operatorname{conv}}(C).$$

■

Corollario 11.5: Corollario al teorema di Lagrange

Siano $(X, \|\cdot\|_X)$ e $(Y, \|\cdot\|_Y)$ due spazi normati.
Sia $A \subseteq X$.
Sia $f : A \to Y$ una funzione.
Si supponga f continua su $A \setminus \mathring{A}$, e G-derivabile su $\mathring{A}$.
Siano $\mathbf{x}, \mathbf{z} \in A$ tali che, per ogni $\lambda \in \,]0;1[$, si abbia $\mathbf{x} + \lambda(\mathbf{z} - \mathbf{x}) \in \mathring{A}$ ed esista $M > 0$ tale che $\|f'(\mathbf{x} + \lambda(\mathbf{z} - \mathbf{x}))\|_{\mathcal{L}(X,Y)} \leq M$
.

Allora, $\|f(\mathbf{z})-f(\mathbf{x})\|_Y \leq M\|\mathbf{z}-\mathbf{x}\|_X$.

Dimostrazione

Valgono tutte le ipotesi del teorema di Lagrange ([Teorema 11.4]);

dunque, $f(\mathbf{z})-f(\mathbf{x}) \in \overline{\operatorname{conv}}(C)$, dove $C=\left\{f'(\mathbf{x}+\lambda(\mathbf{z}-\mathbf{x}))(\mathbf{z}-\mathbf{x}) \mid \lambda \in\,]0;1[\right\}$.

Per ogni $\lambda \in\,]0;1[$, si ha

	$\|f'(\mathbf{x}+\lambda(\mathbf{z}-\mathbf{x}))(\mathbf{z}-\mathbf{x})\|_Y \leq \|f'(\mathbf{x}+\lambda(\mathbf{z}-\mathbf{x}))\|_{\mathcal{L}(X,Y)} \cdot \|\mathbf{z}-\mathbf{x}\|_X$	Per come è definita la norma $\|\cdot\|_{\mathcal{L}(X,Y)}$
$\Longrightarrow$	$\|f'(\mathbf{x}+\lambda(\mathbf{z}-\mathbf{x}))(\mathbf{z}-\mathbf{x})\|_Y \leq M\|\mathbf{z}-\mathbf{x}\|_X$	Essendo $\|f'(\mathbf{x}+\lambda(\mathbf{z}-\mathbf{x}))\|_{\mathcal{L}(X,Y)} \leq M$ per ipotesi
$\Longrightarrow$	$f'(\mathbf{x}+\lambda(\mathbf{z}-\mathbf{x}))(\mathbf{z}-\mathbf{x}) \in \overline{B}(\mathbf{0}_Y, M\|\mathbf{z}-\mathbf{x}\|_X)$	

Pertanto, si ha $C \subseteq B(\mathbf{0}_Y, M\|\mathbf{z}-\mathbf{x}\|_X)$.

Essendo $\overline{B}(\mathbf{0}_Y, M\|\mathbf{z}-\mathbf{x}\|_X)$ un intorno sferico chiuso in Y, esso è chiuso e convesso; contenendo C, si ha che $\overline{\operatorname{conv}}(C) \subseteq \overline{B}(\mathbf{0}_Y, M\|\mathbf{z}-\mathbf{x}\|_X)$.

Per quanto osservato inizialmente si ha allora $f(\mathbf{z})-f(\mathbf{x}) \in \overline{\operatorname{conv}}(C) \subseteq \overline{B}(\mathbf{0}_Y, M\|\mathbf{z}-\mathbf{x}\|_X)$, per cui vale $\|f(\mathbf{z})-f(\mathbf{x})\|_Y \leq M\|\mathbf{z}-\mathbf{x}\|_X$.

$\blacksquare$