15 - Convessità e Derivabilità

Proposizione 15.1: Maggiorazione della derivata direzionale in un punto di G-derivabilità di una funzione convessa

Sia $(X, \|\cdot\|)$ uno spazio normato.
Sia $A \subseteq X$ convesso.
Sia $\mathbf{x}_0 \in \overset{\circ}{A}$.
Sia $f : A \to \mathbb{R}$ una funzione convessa, G-derivabile in $\mathbf{x}_0$.

Si ha $f'(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0) \leq f(\mathbf{x}) - f(\mathbf{x}_0)$ per ogni $\mathbf{x} \in A$.

Dimostrazione

Sia $\mathbf{x} \in A$; si provi che $f'(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0) \leq f(\mathbf{x}) - f(\mathbf{x}_0)$.

Dalla definizione di G-derivabilità di f in $\mathbf{x}_0$, si ha

$$\lim_{\lambda \to 0} \frac{f(\mathbf{x}_0 + \lambda(\mathbf{x} - \mathbf{x}_0)) - f(\mathbf{x}_0)}{\lambda} = f'(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0).$$

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

$$\begin{aligned} & \frac{f(\mathbf{x}_0 + \lambda(\mathbf{x} - \mathbf{x}_0)) - f(\mathbf{x}_0)}{\lambda} = \frac{f(\lambda \mathbf{x} + (1-\lambda)\mathbf{x}_0) - f(\mathbf{x}_0)}{\lambda} \\ & \leq \frac{\lambda f(\mathbf{x}) + (1-\lambda) f(\mathbf{x}_0) - f(\mathbf{x}_0)}{\lambda} && \text{Per convessità di } f \text{, essendo } \lambda \in]0; 1] \\ & = \frac{\lambda f(\mathbf{x}) - \lambda f(\mathbf{x}_0)}{\lambda} = f(\mathbf{x}) - f(\mathbf{x}_0) \end{aligned}$$

Segue allora per confronto che

$$\lim_{\lambda\to 0^+} \frac{f(\mathbf{x}_0+\lambda(\mathbf{x}-\mathbf{x}_0))-f(\mathbf{x}_0)}{\lambda} \leq \lim_{\lambda\to 0^+} f(\mathbf{x}) - f(\mathbf{x}_0) = f(\mathbf{x}) - f(\mathbf{x}_0);$$

essendo anche $\lim_{\lambda\to 0^+} \frac{f(\mathbf{x}_0+\lambda(\mathbf{x}-\mathbf{x}_0))-f(\mathbf{x}_0)}{\lambda} = f'(\mathbf{x}_0)(\mathbf{x}-\mathbf{x}_0)$ per quanto visto prima, si ha allora

$f'(\mathbf{x}_0)(\mathbf{x}-\mathbf{x}_0) \leq f(\mathbf{x}) - f(\mathbf{x}_0)$.

■

Corollario 15.2: Punti in cui una funzione convessa è G-derivabile e ha derivata nulla sono di minimo assoluto

Sia $(X, \|\cdot\|)$ uno spazio normato.
Sia $A \subseteq X$ convesso.
Sia $\mathbf{x}_0 \in \overset{\circ}{A}$.
Sia $f : A \to \mathbb{R}$ una funzione convessa, G-derivabile in $\mathbf{x}_0$.
Si supponga che $f'(\mathbf{x}_0) = \mathbf{0}_{X^*}$.

Allora, $\mathbf{x}_0$ è di minimo assoluto per f.

Dimostrazione

Sia $\mathbf{x} \in A$; si provi che $f(\mathbf{x}) \geq f(\mathbf{x}_0)$.

Essendo f G-derivabile in $\mathbf{x}_0$ con $f'(\mathbf{x}_0) = \mathbf{0}_{X^*}$ per ipotesi, si ha in particolare che
$f'(\mathbf{x}_0)(\mathbf{x}-\mathbf{x}_0) = 0$.

Allora, dalla [Proposizione 15.1] segue che
$f(\mathbf{x}) - f(\mathbf{x}_0) \geq f'(\mathbf{x}_0)(\mathbf{x}-\mathbf{x}_0) = 0$.

■

Proposizione 15.2: Prima caratterizzazione della convessità di funzioni G-derivabili

Sia $(X, \|\cdot\|)$ uno spazio normato.
Sia $A \subseteq X$ aperto e convesso.
Sia $f : A \to \mathbb{R}$ una funzione G-derivabile in A.

Sono equivalenti le seguenti affermazioni:

1. f è convessa;
2. $f(\mathbf{y}) \geq f(\mathbf{x}) + f'(\mathbf{x})(\mathbf{y} - \mathbf{x})$ per ogni $\mathbf{x}, \mathbf{y} \in A$.

Dimostrazione ($1. \Rightarrow 2.$)

Si supponga f convessa.

Siano $\mathbf{x}, \mathbf{y} \in A$; si provi che $f(\mathbf{y}) \geq f(\mathbf{x}) + f'(\mathbf{x})(\mathbf{y} - \mathbf{x})$.

f è convessa per ipotesi, e sempre per ipotesi è anche G-derivabile in $\mathbf{x}$;
dalla [Proposizione 15.1] segue allora che
$f'(\mathbf{x})(\mathbf{y} - \mathbf{x}) \leq f(\mathbf{y}) - f(\mathbf{x})$.

■

Dimostrazione ($2. \Rightarrow 1.$)

Si supponga $f(\mathbf{y}) \geq f(\mathbf{x}) + f'(\mathbf{x})(\mathbf{y} - \mathbf{x})$ per ogni $\mathbf{x}, \mathbf{y} \in A$.

Siano $\mathbf{x}_1, \mathbf{x}_2 \in A$, e sia $\lambda \in [0; 1]$;
si provi che $f(\mathbf{x}_1 + \lambda(\mathbf{x}_2 - \mathbf{x}_1)) \leq \lambda f(\mathbf{x}_2) + (1 - \lambda) f(\mathbf{x}_1)$.

Si hanno le seguenti disuguaglianze:

$f(\mathbf{x}_1) \geq f\big(\mathbf{x}_1 + \lambda(\mathbf{x}_2 - \mathbf{x}_1)\big) + f'\big(\mathbf{x}_1 + \lambda(\mathbf{x}_2 - \mathbf{x}_1)\big)\big(\mathbf{x}_1 - \mathbf{x}_1 - \lambda(\mathbf{x}_2 - \mathbf{x}_1)\big)$ Per ipotesi

$= f\big(\mathbf{x}_1 + \lambda(\mathbf{x}_2 - \mathbf{x}_1)\big) + f'\big(\mathbf{x}_1 + \lambda(\mathbf{x}_2 - \mathbf{x}_1)\big)\big(-\lambda(\mathbf{x}_2 - \mathbf{x}_1)\big)$

$= f\big(\mathbf{x}_1 + \lambda(\mathbf{x}_2 - \mathbf{x}_1)\big) - \lambda f'\big(\mathbf{x}_1 + \lambda(\mathbf{x}_2 - \mathbf{x}_1)\big)(\mathbf{x}_2 - \mathbf{x}_1)$ Per linearità di $f'\big(\mathbf{x}_1 + \lambda(\mathbf{x}_2 - \mathbf{x}_1)\big)$

da cui segue che $(1-\lambda)f(\mathbf{x}_1) \geq (1-\lambda)f\big(\mathbf{x}_1 + \lambda(\mathbf{x}_2 - \mathbf{x}_1)\big) - \lambda(1-\lambda)f'\big(\mathbf{x}_1 + \lambda(\mathbf{x}_2 - \mathbf{x}_1)\big)(\mathbf{x}_2 - \mathbf{x}_1)$, moltiplicando ambo i membri per $1 - \lambda \geq 0$;

$f(\mathbf{x}_2) \geq f\big(\mathbf{x}_1 + \lambda(\mathbf{x}_2 - \mathbf{x}_1)\big) + f'\big(\mathbf{x}_1 + \lambda(\mathbf{x}_2 - \mathbf{x}_1)\big)\big(\mathbf{x}_2 - \mathbf{x}_1 - \lambda(\mathbf{x}_2 - \mathbf{x}_1)\big)$

$= f\big(\mathbf{x}_1 + \lambda(\mathbf{x}_2 - \mathbf{x}_1)\big) + f'\big(\mathbf{x}_1 + \lambda(\mathbf{x}_2 - \mathbf{x}_1)\big)\big((1-\lambda)(\mathbf{x}_2 - \mathbf{x}_1)\big)$

$= f\big(\mathbf{x}_1 + \lambda(\mathbf{x}_2 - \mathbf{x}_1)\big) + (1-\lambda)f'\big(\mathbf{x}_1 + \lambda(\mathbf{x}_2 - \mathbf{x}_1)\big)(\mathbf{x}_2 - \mathbf{x}_1)$ Per linearità di $f'\big(\mathbf{x}_1 + \lambda(\mathbf{x}_2 - \mathbf{x}_1)\big)$

da cui segue che $\lambda f(\mathbf{x}_2) \geq \lambda f\big(\mathbf{x}_1 + \lambda(\mathbf{x}_2 - \mathbf{x}_1)\big) + \lambda(1-\lambda)f'\big(\mathbf{x}_1 + \lambda(\mathbf{x}_2 - \mathbf{x}_1)\big)(\mathbf{x}_2 - \mathbf{x}_1)$, moltiplicando ambo i membri per $\lambda \geq 0$.

Sommando membro a membro le due disuguaglianze ottenute e semplificando, si ricava che

$\lambda f(\mathbf{x}_2) + (1-\lambda)f(\mathbf{x}_1) \geq f\big(\mathbf{x}_1 + \lambda(\mathbf{x}_2 - \mathbf{x}_1)\big)$,

che è esattamente ciò che si voleva provare.
■

⌘ Definizione: Monotonia

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

Un operatore $T : A \to X^*$ si dice **monotono** quando
$\big(T(\mathbf{x}) - T(\mathbf{y})\big)(\mathbf{x} - \mathbf{y}) \geq 0$ per ogni $\mathbf{x}, \mathbf{y} \in A$.

Osservazione

Si consideri $\mathbb{R}$ come spazio euclideo, che è di Hilbert rispetto al prodotto di numeri reali;
sia $\phi : \mathbb{R} \to \mathbb{R}^*$ la funzione definita ponendo $\phi(x) = \langle x, \cdot \rangle = x \cdot (\,\cdot\,)$, che è un'isometria lineare per la [Proposizione 10.14].
Sia $A \subseteq \mathbb{R}$.

Un operatore $T : A \to \mathbb{R}^*$ è monotono se e solo se la funzione $\tilde{T} := \phi^{-1} \circ T$ è non decrescente in A.

Dimostrazione

Siano $x, y \in A$.
Sia $x_0 = \tilde{T}(x)$, e sia $y_0 = \tilde{T}(y)$.

Si osserva intanto che

$$\begin{aligned}
&\big(T(x) - T(y)\big)(x - y) && \text{Per monotonia di } T \\
&= \big(\phi(x_0) - \phi(y_0)\big)(x - y) && \text{Per definizione di } x_0 \text{ e } y_0 \text{ e di } \tilde{T} \\
&= \big(\phi(x_0 - y_0)\big)(x - y) && \text{Per linearità di } \phi \\
&= (x_0 - y_0) \cdot (x - y) && \text{Per definizione di } \phi \\
&= \big(\tilde{T}(x) - \tilde{T}(y)\big) \cdot (x - y) && \text{Per definizione di } x_0 \text{ e } y_0
\end{aligned}$$

Si supponga dapprima T monotono.
Fissati $x, y \in A$ con $x < y$, si ha $\big(T(x) - T(y)\big)(x - y) \geq 0$ per ipotesi di monotonia di T, ossia
$\big(\tilde{T}(x) - \tilde{T}(y)\big) \cdot (x - y) \geq 0$ per la catena di uguaglianze osservata prima;
essendo $x < y$, segue allora $\tilde{T}(x) \leq \tilde{T}(y)$, da cui la non decrescenza di $\tilde{T}$ per arbitrarietà di $x, y \in A$.

Viceversa, se $\varphi^{-1} \circ T$ è non decrescente, si ha
$\big(\tilde{T}(x) - \tilde{T}(y)\big) \cdot (x - y) \geq 0$ per ogni $x, y \in A$ (si verifica per ispezione sui casi $x \leq y$ e $x > y$),
$\big(T(x) - T(y)\big)(x - y) \geq 0$ per la catena di uguaglianze osservata prima;

segue allora la monotonia di T per arbitrarietà di $x, y \in A$.

■

Proposizione 15.3: Seconda caratterizzazione della convessità di funzioni G-derivabili

Sia $(X, \|\cdot\|)$ uno spazio normato.
Sia $A \subseteq X$ aperto e convesso.
Sia $f : A \to \mathbb{R}$ una funzione G-derivabile in A.

Sono equivalenti le seguenti affermazioni:

1. f è convessa;
2. f' è un operatore monotono.

Osservazioni preliminari

Fissati $\mathbf{x}, \mathbf{y} \in A$, si definisca $\varphi_{\mathbf{x},\mathbf{y}} : [0;1] \to \mathbb{R}$ ponendo
$\varphi_{\mathbf{x},\mathbf{y}}(\lambda) = f\big(\mathbf{x} + \lambda(\mathbf{y} - \mathbf{x})\big)$ per ogni $\lambda \in [0;1]$.

Allora:

1. Per ogni $\mathbf{x}, \mathbf{y} \in A$, $\varphi_{\mathbf{x},\mathbf{y}}$ è derivabile in $[0;1]$, e si ha $\dot{\varphi}_{\mathbf{x},\mathbf{y}}(\lambda) = f'\big(\mathbf{x} + \lambda(\mathbf{y} - \mathbf{x})\big)(\mathbf{y} - \mathbf{x})$ per ogni $\lambda \in [0;1]$;
2. f è convessa in A se e solo se $\varphi_{\mathbf{x},\mathbf{y}}$ è convessa in $[0;1]$ per ogni $\mathbf{x}, \mathbf{y} \in A$;
3. f è convessa in A se e solo se $\dot{\varphi}_{\mathbf{x},\mathbf{y}}$ è non decrescente in $[0;1]$ per ogni $\mathbf{x}, \mathbf{y} \in A$.

Il punto 1. segue dal fatto che, fissati $\mathbf{x}, \mathbf{y} \in A$ e fissato $\lambda_0 \in [0;1]$, si ha

$$\lim_{\mu \to 0} \frac{\varphi_{\mathbf{x},\mathbf{y}}(\lambda_0+\mu) - \varphi_{\mathbf{x},\mathbf{y}}(\lambda_0)}{\mu} = \lim_{\mu \to 0} \frac{f\big(\mathbf{x}+\lambda_0(\mathbf{y}-\mathbf{x})+\mu(\mathbf{y}-\mathbf{x})\big) - f\big(\mathbf{x}+\lambda_0(\mathbf{y}-\mathbf{x})\big)(\mathbf{y}-\mathbf{x})}{\mu} \quad \text{Per definizione di } \varphi_{\mathbf{x},\mathbf{y}}$$

$= f'\big(\mathbf{x} + \lambda_0(\mathbf{y} - \mathbf{x})\big)(\mathbf{y} - \mathbf{x})$

Per G-derivabilità di f in $\mathbf{x} + \lambda_0(\mathbf{y} - \mathbf{x}) \in A$ (A è convesso, e $\mathbf{x}, \mathbf{y} \in A$), e per definizione di f'

Si mostri il punto 2.
Si fissino $\mathbf{x}, \mathbf{y} \in A$.
Se $\varphi_{\mathbf{x},\mathbf{y}}$ è convessa in $[0;1]$, si ha in particolare che

$f\big(\mathbf{x} + \lambda(\mathbf{y} - \mathbf{x})\big) = \varphi_{\mathbf{x},\mathbf{y}}(\lambda)$	Per definzione di $\varphi_{\mathbf{x},\mathbf{y}}$
$= \varphi_{\mathbf{x},\mathbf{y}}(0 + \lambda(1-0))$	
$\leq \lambda\varphi_{\mathbf{x},\mathbf{y}}(1) + (1-\lambda)\varphi_{\mathbf{x},\mathbf{y}}(0)$	Per convessità di $\varphi_{\mathbf{x},\mathbf{y}}$
$= \lambda f(\mathbf{y}) + (1-\lambda) f(\mathbf{x})$	Per definizione di $\varphi_{\mathbf{x},\mathbf{y}}$

Ne segue che, se $\varphi_{\mathbf{x},\mathbf{y}}$ è convessa in $[0;1]$ per ogni $\mathbf{x}, \mathbf{y} \in A$, f è convessa in A.

Viceversa, si supponga f convessa in A.

Siano $\mathbf{x}, \mathbf{y} \in A$;
siano $\lambda_1, \lambda_2 \in [0;1]$, e sia $\mu \in [0;1]$.

Si ha
$\varphi_{\mathbf{x},\mathbf{y}}\big(\lambda_1 + \mu(\lambda_2 - \lambda_1)\big) = f\big(\mathbf{x} + \big(\lambda_1 + \mu(\lambda_2 - \lambda_1)\big)(\mathbf{y} - \mathbf{x})\big)$	Per definizione di $\varphi_{\mathbf{x},\mathbf{y}}$
$= f\big(\mathbf{x} + \lambda_1(\mathbf{y} - \mathbf{x}) + \mu(\lambda_2 - \lambda_1)(\mathbf{y} - \mathbf{x})\big)$	
$= f\Big(\,\mathbf{x} + \lambda_1(\mathbf{y} - \mathbf{x}) + \mu\Big(\,\mathbf{x} + \lambda_2(\mathbf{y} - \mathbf{x}) - (\mathbf{x} + \lambda_1(\mathbf{y} - \mathbf{x}))\,\Big)\,\Big)$	
$\leq \mu f\big(\mathbf{x} + \lambda_2(\mathbf{y} - \mathbf{x})\big) + (1-\mu) f\big(\mathbf{x} + \lambda_1(\mathbf{y} - \mathbf{x})\big)$	Per convessità di f
$= \mu\,\varphi_{\mathbf{x},\mathbf{y}}(\lambda_2) + (1-\mu)\varphi_{\mathbf{x},\mathbf{y}}(\lambda_1)$	Per definizione di $\varphi_{\mathbf{x},\mathbf{y}}$

da cui segue la convessità di $\varphi_{\mathbf{x},\mathbf{y}}$.

Infine, essendo $\varphi_{\mathbf{x},\mathbf{y}}$ derivabile in $[0;1]$ per ogni $\mathbf{x},\mathbf{y} \in A$ e a valori reali, per un risultato noto si ha che $\varphi_{\mathbf{x},\mathbf{y}}$ è convessa se e solo se $\dot{\varphi}_{\mathbf{x},\mathbf{y}}$ è non decrescente in $[0;1]$.

Essendo la convessità di $\varphi_{\mathbf{x},\mathbf{y}}$ equivalente alla convessità di f per il punto precedente, è acquisito anche il punto 3.

Dimostrazione ($1. \Rightarrow 2.$)

Si supponga f convessa;
si provi che f' è monotona, cioè $\big(f'(\mathbf{x}) - f'(\mathbf{y})\big)(\mathbf{x} - \mathbf{y}) \geq 0$ per ogni $\mathbf{x},\mathbf{y} \in A$.

Si fissino dunque $\mathbf{x},\mathbf{y} \in A$.

In virtù delle osservazioni preliminari, si ha che $\varphi_{\mathbf{x},\mathbf{y}}$ è derivabile in $[0;1]$ con $\dot{\varphi}_{\mathbf{x},\mathbf{y}}(\lambda) = f'(\mathbf{x} + \lambda(\mathbf{y} - \mathbf{x}))(\mathbf{y} - \mathbf{x})$ per ogni $\lambda \in [0;1]$;
inoltre, essendo f convessa, si ha $\dot{\varphi}_{\mathbf{x},\mathbf{y}}$ non decrescente in $[0;1]$.

Ne viene in particolare che $\dot{\varphi}_{\mathbf{x},\mathbf{y}}(0) \leq \dot{\varphi}_{\mathbf{x},\mathbf{y}}(1)$, ossia
$f'(\mathbf{x})(\mathbf{y} - \mathbf{x}) \leq f'(\mathbf{y})(\mathbf{y} - \mathbf{x})$, da cui la tesi.
■

Dimostrazione ($2. \Rightarrow 1.$)

Si supponga f' monotona;
in virtù delle osservazioni preliminari, si per provare la convessità di f basta mostrare che $\dot{\varphi}_{\mathbf{x},\mathbf{y}}$ è non decrescente in $[0;1]$ per ogni $\mathbf{x},\mathbf{y} \in A$.

Siano dunque $\lambda, \mu \in [0;1]$, con $\lambda < \mu$;
si provi che $\dot{\varphi}_{\mathbf{x},\mathbf{y}}(\lambda) \leq \dot{\varphi}_{\mathbf{x},\mathbf{y}}(\mu)$.

Si ha

$0 \leq \big(f'(\mathbf{x}+\lambda(\mathbf{y}-\mathbf{x}))-f'(\mathbf{x}+\mu(\mathbf{y}-\mathbf{x}))\big)\big(\mathbf{x}+\lambda(\mathbf{y}-\mathbf{x})-(\mathbf{x}+\mu(\mathbf{y}-\mathbf{x}))\big)$	Per ipotesi di monotonia di f'
$= \big(f'(\mathbf{x}+\lambda(\mathbf{y}-\mathbf{x}))-f'(\mathbf{x}+\mu(\mathbf{y}-\mathbf{x}))\big)\big((\lambda-\mu)(\mathbf{y}-\mathbf{x})\big)$	
$= f'\big(\mathbf{x}+\lambda(\mathbf{y}-\mathbf{x})\big)\big((\lambda-\mu)(\mathbf{y}-\mathbf{x})\big)-f'\big(\mathbf{x}+\mu(\mathbf{y}-\mathbf{x})\big)\big((\lambda-\mu)(\mathbf{y}-\mathbf{x})\big)$	Per definizione di $f'(\mathbf{x}+\lambda(\mathbf{y}-\mathbf{x}))-f'(\mathbf{x}+\mu(\mathbf{y}-\mathbf{x}))$
$= (\lambda-\mu)f'\big(\mathbf{x}+\lambda(\mathbf{y}-\mathbf{x})\big)(\mathbf{y}-\mathbf{x})-(\lambda-\mu)f'(\mathbf{x}+\mu(\mathbf{y}-\mathbf{x}))(\mathbf{y}-\mathbf{x})$	Per linearità di $f'(\mathbf{x}+\lambda(\mathbf{y}-\mathbf{x}))$ e $f'(\mathbf{x}+\mu(\mathbf{y}-\mathbf{x}))$
$= (\lambda-\mu)\dot{\varphi}_{\mathbf{x},\mathbf{y}}(\lambda)-(\lambda-\mu)\dot{\varphi}_{\mathbf{x},\mathbf{y}}(\mu)$	Per legge di $\dot{\varphi}_{\mathbf{x},\mathbf{y}}$
$= (\lambda-\mu)\big(\dot{\varphi}_{\mathbf{x},\mathbf{y}}(\lambda)-\dot{\varphi}_{\mathbf{x},\mathbf{y}}(\mu)\big)$	

Da questa catena di disuguaglianze e dal fatto che $\lambda < \mu$, segue che $\dot{\varphi}_{\mathbf{x},\mathbf{y}}(\lambda)-\dot{\varphi}_{\mathbf{x},\mathbf{y}}(\mu) \leq 0$, che è ciò che si voleva mostrare.

■

Dimostrazione alternativa ($1. \Rightarrow 2.$) (menomale)

Si supponga f convessa;
si provi che f' è monotona, cioè $\big(f'(\mathbf{x})-f'(\mathbf{y})\big)(\mathbf{x}-\mathbf{y}) \geq 0$ per ogni $\mathbf{x},\mathbf{y} \in A$.

Si fissino dunque $\mathbf{x},\mathbf{y} \in A$.

Essendo f convessa e G-derivabile su A per ipotesi, dalla [Proposizione 15.2] segue che
$f(\mathbf{y}) \geq f(\mathbf{x})+f'(\mathbf{x})(\mathbf{y}-\mathbf{x})$,
e anche
$f(\mathbf{x}) \geq f(\mathbf{y})+f'(\mathbf{y})(\mathbf{x}-\mathbf{y})$.

Sommando membro a membro le due disuguaglianze ottenute e semplificando, si ottiene

$$
\begin{aligned}
0 &\geq f'(\mathbf{x})(\mathbf{y}-\mathbf{x}) + f'(\mathbf{y})(\mathbf{x}-\mathbf{y}) \\
&= -f'(\mathbf{x})(\mathbf{x}-\mathbf{y}) + f'(\mathbf{y})(\mathbf{x}-\mathbf{y}) && \text{Per linearità di } f'(\mathbf{x})
\end{aligned}
$$

Si ha dunque $f'(\mathbf{x})(\mathbf{x}-\mathbf{y}) - f'(\mathbf{y})(\mathbf{x}-\mathbf{y}) \geq 0$, che corrisponde a quanto si voleva mostrare.
■

Dimostrazione alternativa ($2. \Rightarrow 1.$) (menomale parte 2)

Si supponga f' monotona.

Essendo f G-derivabile su A, in virtù della [Proposizione 15.2] si provi che
$f(\mathbf{y}) \geq f(\mathbf{x}) + f'(\mathbf{x})(\mathbf{y}-\mathbf{x})$ per ogni $\mathbf{x} \in X$.

Si fissino $\lambda, \mu \in [0;1]$ con $\lambda < \mu$.

Si ha che

$$
\begin{aligned}
0 &\leq \big(f'\big(\mathbf{x}+\lambda(\mathbf{y}-\mathbf{x})\big) - f'\big(\mathbf{x}+\mu(\mathbf{y}-\mathbf{x})\big)\big)\big(\mathbf{x}+\lambda(\mathbf{y}-\mathbf{x}) - (\mathbf{x}+\mu(\mathbf{y}-\mathbf{x}))\big) && \text{Per ipotesi di monotonia di } f' \\
0 &\leq \big(f'(\mathbf{x}+\lambda(\mathbf{y}-\mathbf{x})) - f'(\mathbf{x}+\mu(\mathbf{y}-\mathbf{x}))\big)\big(\mathbf{x}+\lambda(\mathbf{y}-\mathbf{x}) - (\mathbf{x}+\mu(\mathbf{y}-\mathbf{x}))\big) && \text{Per ipotesi di monotonia di } f' \\
&= \big(f'(\mathbf{x}+\lambda(\mathbf{y}-\mathbf{x})) - f'(\mathbf{x}+\mu(\mathbf{y}-\mathbf{x}))\big)\big((\lambda-\mu)(\mathbf{y}-\mathbf{x})\big) \\
&= (\lambda-\mu)\big(f'(\mathbf{x}+\lambda(\mathbf{y}-\mathbf{x})) - f'(\mathbf{x}+\mu(\mathbf{y}-\mathbf{x}))\big)(\mathbf{y}-\mathbf{x}) && \text{Per linearità di } f'(\mathbf{x}+\lambda(\mathbf{y}-\mathbf{x})) - f'(\mathbf{x}+\mu(\mathbf{y}-\mathbf{x}))
\end{aligned}
$$

Avendo supposto $\lambda < \mu$, ne segue che
$\big(f'(\mathbf{x}+\lambda(\mathbf{y}-\mathbf{x})) - f'(\mathbf{x}+\mu(\mathbf{y}-\mathbf{x}))\big)(\mathbf{y}-\mathbf{x}) \leq 0$, ossia
$f'\big(\mathbf{x}+\lambda(\mathbf{y}-\mathbf{x})\big)(\mathbf{y}-\mathbf{x}) \leq f'\big(\mathbf{x}+\mu(\mathbf{y}-\mathbf{x})\big)(\mathbf{y}-\mathbf{x})$.

Ne segue che la funzione $[0;1] \to \mathbb{R} : \lambda \mapsto f'\big(\mathbf{x}+\lambda(\mathbf{y}-\mathbf{x})\big)(\mathbf{y}-\mathbf{x})$ è monotona non decrescente;
in particolare, si ha allora

$$\underbrace{f'\big(\mathbf{x}+0(\mathbf{y}-\mathbf{x})\big)(\mathbf{y}-\mathbf{x})}_{=f'(\mathbf{x})(\mathbf{y}-\mathbf{x})} \leq f'\big(\mathbf{x}+\lambda(\mathbf{y}-\mathbf{x})\big)(\mathbf{y}-\mathbf{x}) \leq \underbrace{f'\big(\mathbf{x}+1(\mathbf{y}-\mathbf{x})\big)(\mathbf{y}-\mathbf{x})}_{=f'(\mathbf{y})(\mathbf{y}-\mathbf{x})},$$

per ogni $\lambda \in [0;1]$.

Sia ora $C = \left\{f'(\mathbf{x}+\lambda(\mathbf{y}-\mathbf{x}))(\mathbf{y}-\mathbf{x}) \mid \lambda \in]0;1[\right\}$;
per quanto appena osservato, si ha $C \subseteq \left[f'(\mathbf{x})(\mathbf{y}-\mathbf{x})\,;\, f'(\mathbf{y})(\mathbf{y}-\mathbf{x})\right]$.

Perdipiù, essendo $\left[f'(\mathbf{x})(\mathbf{y}-\mathbf{x})\,;\, f'(\mathbf{y})(\mathbf{y}-\mathbf{x})\right]$ un insieme chiuso e convesso, ne viene che
$\overline{\operatorname{conv}}(C) \subseteq \left[f'(\mathbf{x})(\mathbf{y}-\mathbf{x})\,;\, f'(\mathbf{y})(\mathbf{y}-\mathbf{x})\right]$.

Si osserva infine che f soddisfa le ipotesi del teorema di Lagrange ([Teorema 11.4]) rispetto a $\mathbf{x}, \mathbf{y}$, pertanto

$f(\mathbf{y}) - f(\mathbf{x}) \in \overline{\operatorname{conv}}(C) \subseteq \left[f'(\mathbf{x})(\mathbf{y}-\mathbf{x})\,;\, f'(\mathbf{y})(\mathbf{y}-\mathbf{x})\right]$, da cui segue in ultima battuta che

$f(\mathbf{y}) - f(\mathbf{x}) \geq f'(\mathbf{x})(\mathbf{y}-\mathbf{x})$, che è ciò che si voleva mostrare.
$\blacksquare$