II

Lemma 1: if f is closed and Convex, then all the following are equivalent:

$$y \in \partial f(x) \longleftrightarrow x \in \partial f^*(y) \longleftrightarrow x^T y = f(x) + f^*(y)$$

Proof: if $y \in \partial f(x)$, then $f^*(y) = \max_u (y^T u - f(u)) = y^T x - f(x)$

$$\therefore f^*(v) = \max_v (v^T u - f(u))$$

$$\geq x^T v - f(x) \qquad , \; + x^T y - x^T y$$

$$\geq x^T(v-y) \underbrace{- f(x) + y^T x} \qquad , \; f^*(y) = y^T x - f(x)$$

$$\geq f^*(y) + x^T(v-y)$$

This holds for all v, therefor $x \in \partial f^*(y)$, by Comparison to the definition of the Subgradient.

Since $f^{**} = f$, if $x \in \partial f^*(y) \longrightarrow y \in \partial f(x)$

$\rightarrow$ Main proof: $x^T y - f^*(y)$ has a unique maximizer y for every x

y maximizes $x^T y - f^*(y)$ iff $x \in \partial f^*(y)$ [from Lemma 1]

$$x \in \partial f^*(y) \longleftrightarrow y \in \partial f(x) = \{\nabla f(x)\}$$

$$\boxed{\therefore \nabla f(x) = \underset{y}{\operatorname{argmax}} \left(x^T y - f^*(y)\right)}$$