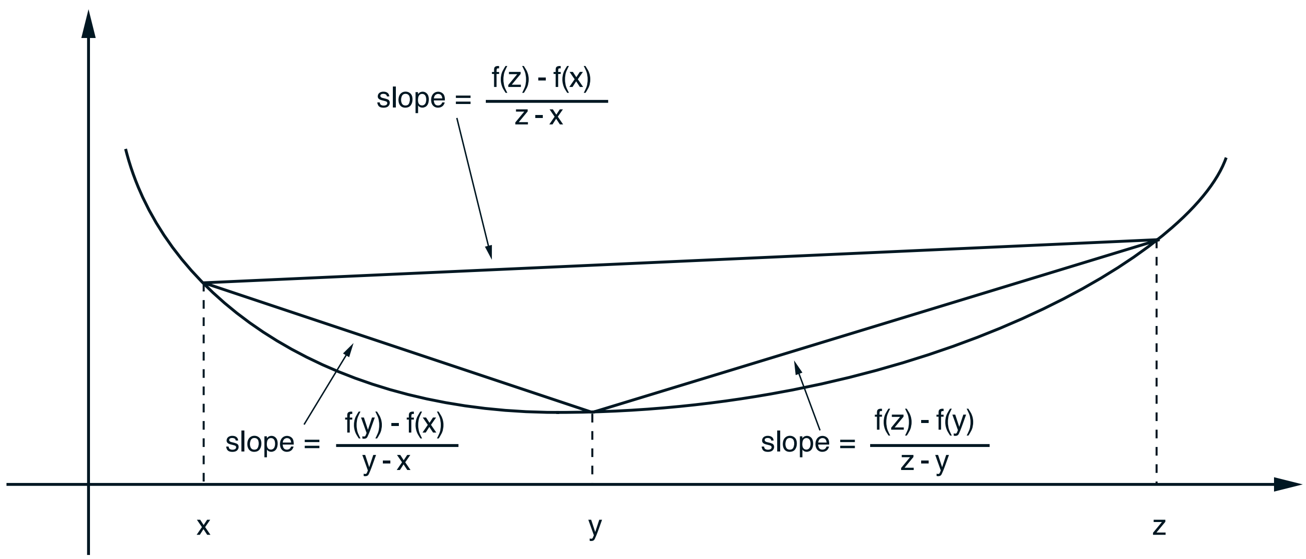
LECTURE 13

LECTURE OUTLINE

- Directional derivatives of one-dimensional convex functions
- Directional derivatives of multi-dimensional convex functions
- Subgradients and subdifferentials
- Properties of subgradients

ONE-DIMENSIONAL DIRECTIONAL DERIVATIVES

- Three slopes relation for a convex $f : \Re \mapsto \Re$:



$$\frac{f(y)-f(x)}{y-x} \leq \frac{f(z)-f(x)}{z-x} \leq \frac{f(z)-f(y)}{z-y}$$

- Right and left directional derivatives exist

$$f^{+}(x) = \lim_{\alpha \downarrow 0} \frac{f(x+\alpha)-f(x)}{\alpha}$$

$$f^{-}(x) = \lim_{\alpha \downarrow 0} \frac{f(x)-f(x-\alpha)}{\alpha}$$

MULTI-DIMENSIONAL DIRECTIONAL DERIVATIVES

- For a convex $f : \Re^n \mapsto \Re$

$$f'(x; y) = \lim_{\alpha \downarrow 0} \frac{f(x + \alpha y) - f(x)}{\alpha},$$

is the directional derivative at x in the direction y.

- Exists for all x and all directions.

- f is differentiable at x if $f'(x; y)$ is a linear function of y denoted by

$$f'(x; y) = \nabla f(x)' y,$$

where $\nabla f(x)$ is the gradient of f at x.

- Directional derivatives can be defined for extended real-valued convex functions, but we will not pursue this topic (see the book).

SUBGRADIENTS

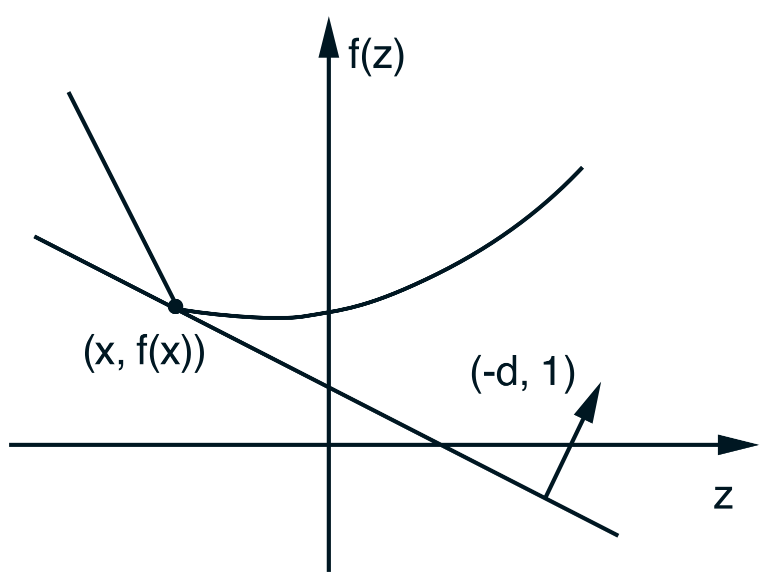
- Let $f : \Re^n \mapsto \Re$ be a convex function. A vector $d \in \Re^n$ is a *subgradient* of f at a point $x \in \Re^n$ if

$$f(z) \geq f(x) + (z - x)'d, \qquad \forall\, z \in \Re^n.$$

- d is a subgradient if and only if

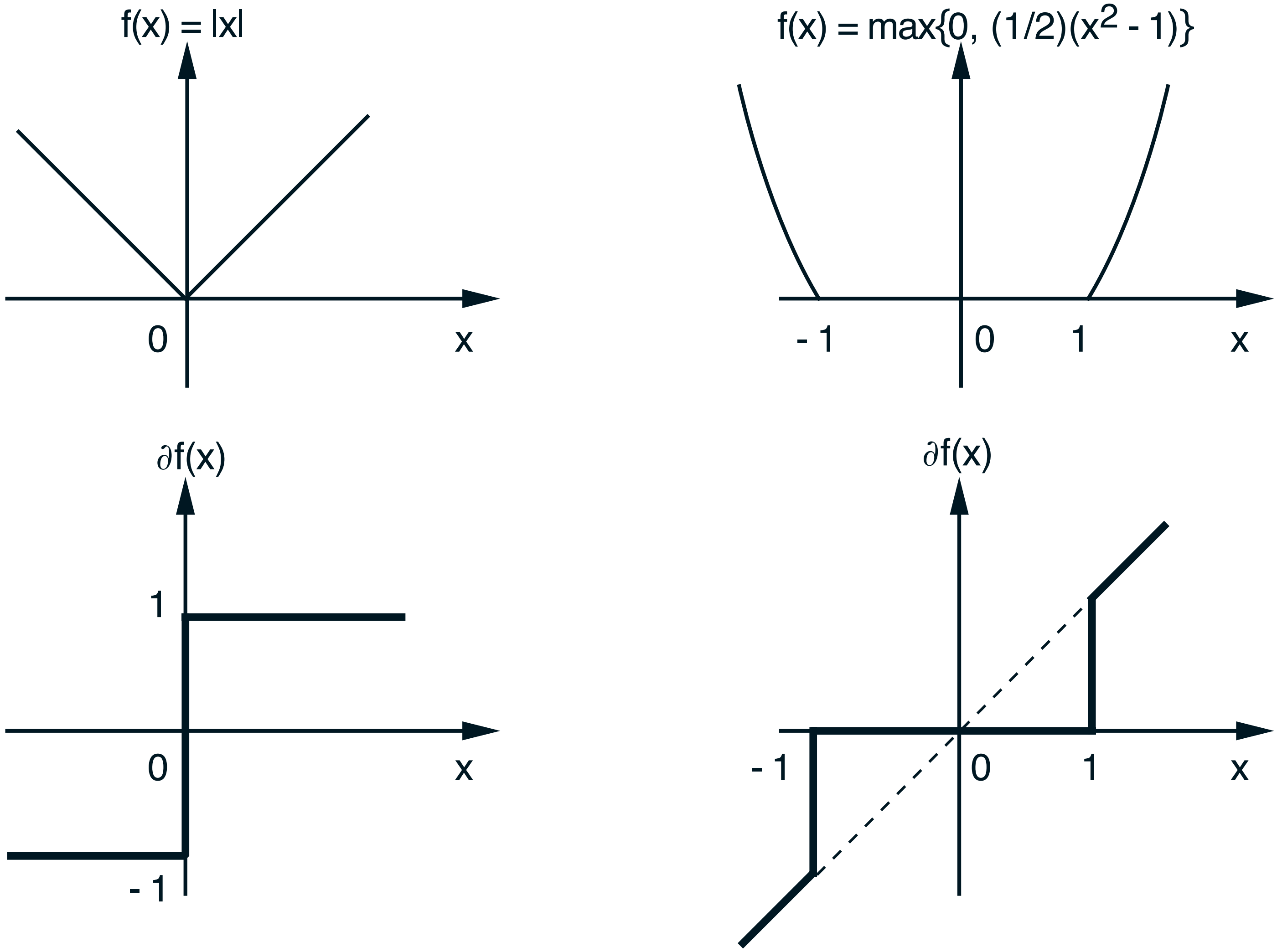
$$f(z) - z'd \geq f(x) - x'd, \qquad \forall\, z \in \Re^n$$

so d is a subgradient at x if and only if the hyperplane in $\Re^{n+1}$ that has normal $(-d, 1)$ and passes through $\big(x, f(x)\big)$ supports the epigraph of f.



SUBDIFFERENTIAL

- The set of all subgradients of a convex function f at x is called the *subdifferential* of f at x, and is denoted by $\partial f(x)$.
- Examples of subdifferentials:



PROPERTIES OF SUBGRADIENTS I

- $\partial f(x)$ is nonempty, convex, and compact.

Proof: Consider the min common/max crossing framework with

$$M = \{(u, w) \mid u \in \Re^n,\ f(x+u) \le w\}.$$

Min common value: $w^* = f(x)$. Crossing value function is $q(\mu) = \inf_{(u,w)\in M}\{w + \mu' u\}$. We have $w^* = q^* = q(\mu)$ iff $f(x) = \inf_{(u,w)\in M}\{w + \mu' u\}$, or

$$f(x) \le f(x+u) + \mu' u, \qquad \forall\, u \in \Re^n.$$

Thus, the set of optimal solutions of the max crossing problem is precisely $-\partial f(x)$. Use the Min Common/Max Crossing Theorem II: since the set

$$D = \{u \mid \text{there exists } w \in \Re \text{ with } (u, w) \in M\} = \Re^n$$

contains the origin in its interior, the set of optimal solutions of the max crossing problem is nonempty, convex, and compact. **Q.E.D.**

PROPERTIES OF SUBGRADIENTS II

- For every $x \in \Re^n$, we have

$$f'(x; y) = \max_{d \in \partial f(x)} y'd, \qquad \forall\, y \in \Re^n.$$

- f is differentiable at x with gradient $\nabla f(x)$, if and only if it has $\nabla f(x)$ as its unique subgradient at x.

- If $f = \alpha_1 f_1 + \cdots + \alpha_m f_m$, where the $f_j : \Re^n \mapsto \Re$ are convex and $\alpha_j > 0$,

$$\partial f(x) = \alpha_1 \partial f_1(x) + \cdots + \alpha_m \partial f_m(x).$$

- Chain Rule: If $F(x) = f(Ax)$, where A is a matrix,

$$\partial F(x) = A' \partial f(Ax) = \big\{ A'g \mid g \in \partial f(Ax) \big\}.$$

- Generalizes to functions $F(x) = g\big(f(x)\big)$, where g is smooth.

ADDITIONAL RESULTS ON SUBGRADIENTS

- **Danskin's Theorem**: Let Z be compact, and $\phi : \Re^n \times Z \mapsto \Re$ be continuous. Assume that $\phi(\cdot, z)$ is convex and differentiable for all $z \in Z$. Then the function $f : \Re^n \mapsto \Re$ given by

$$f(x) = \max_{z \in Z} \phi(x, z)$$

is convex and for all x

$$\partial f(x) = \mathrm{conv}\big\{\nabla_x \phi(x, z) \mid z \in Z(x)\big\}.$$

- The subdifferential of an extended real valued convex function $f : \Re^n \mapsto (-\infty, \infty]$ is defined by

$$\partial f(x) = \big\{d \mid f(z) \geq f(x) + (z - x)'d,\ \forall\, z \in \Re^n\big\}.$$

- $\partial f(x)$, is closed but may be empty at relative boundary points of $\mathrm{dom}(f)$, and may be unbounded.

- $\partial f(x)$ is nonempty at all $x \in \mathrm{ri}\big(\mathrm{dom}(f)\big)$, and it is compact if and only if $x \in \mathrm{int}\big(\mathrm{dom}(f)\big)$. The proof again is by Min Common/Max Crossing II.