Lecture 23

Last time

- Clarke subdifferential
- Basic measure theory.

Today

- Ekeland's variational principle
- Inverse Problems.

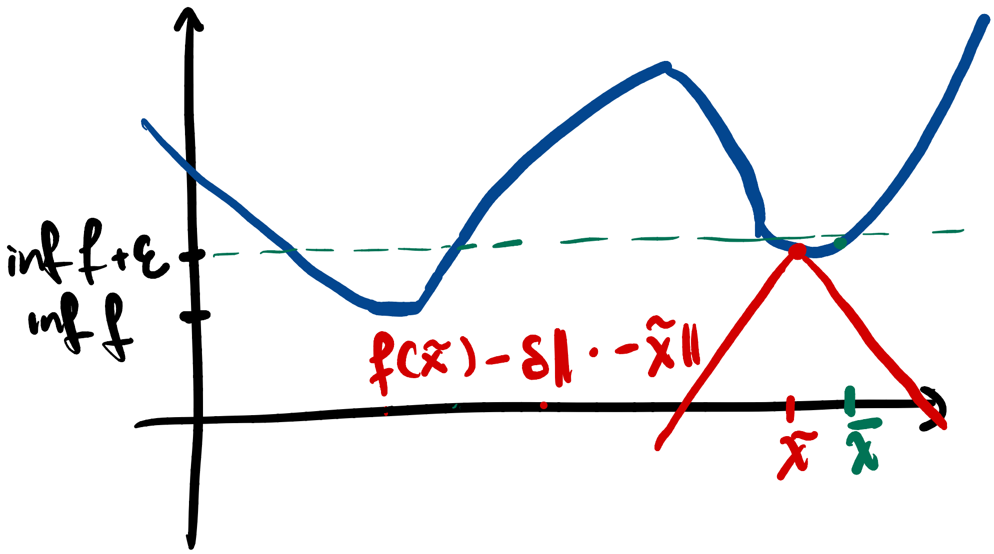
Ekeland Variational principle

To prove many of the theorems we have covered we used the existence of a minimizer for closed, coercive functions. The following priciple gets ride of coercivity.

Theorem: Let $f : E \to \mathbb{R} \cup \{+\infty\}$ be a proper closed function. Fix a point $\bar{x}$ s.t. $f(\bar{x}) \leq \inf f + \varepsilon$ for some $\varepsilon > 0$. Then for any $\delta > 0$, there is a point $\hat{x}$ satisfying

1. $\|\bar{x} - \hat{x}\| \leq \varepsilon/\delta$
2. $f(\hat{x}) \leq f(\bar{x})$
3. $\{\hat{x}\} = \operatorname{argmin}\{f(x) + \delta\|x - \hat{x}\|\}$

Intuition



Proof: We want to find a conic minorant as in our intuition. Define the function

$$c(x) = f(\bar{x}) - \delta\|x - \bar{x}\|.$$

If $f(x) > c(x)$ $\forall x \neq x$, then, we can take $\tilde{x} = x$. Suppose that this is not the case. Define

$$\begin{aligned} W &:= \{x \mid f(x) \le c(x)\} \\ &= \{x \mid f(x) + \delta\|x - \bar{x}\| \le f(\bar{x})\}. \\ &\subseteq \{x \mid \delta\|x - \bar{x}\| \le f(\bar{x}) - \inf f\} \\ &\subseteq \{x \mid \|x - \bar{x}\| \le \varepsilon/\delta\} \end{aligned}$$

Notice that W is nonempty, bounded

since $f+\delta\|\cdot-\bar{x}\|$ is closed

and closed. Therefore, the function $f(x)+\tau_W(x)$ achieves a minimizer $\tilde{x}$. Since $\bar{x}\in W$, we have that

$$(\because)\qquad f(\tilde{x})+\delta\|\tilde{x}-\bar{x}\|\le f(\bar{x}).$$

It is then immediate that 1. and 2. hold. Moreover, for $x\in W\setminus\tilde{x}$ we have $f(\tilde{x})\le f(x)$ and so

$$f(\tilde{x})+0<f(x)+\delta\|x-\tilde{x}\|.$$

For $x\in W^c$, then

$$\begin{aligned}
f(\tilde{x}) &\overset{(\because)}{\le} f(\bar{x})-\delta\|\tilde{x}-\bar{x}\| \\
&\overset{x\in W^c}{<} f(x)+\delta\|x-\bar{x}\|-\delta\|\tilde{x}-\bar{x}\| \\
&\le f(x)-\delta\|x-\tilde{x}\|
\end{aligned}$$

where the last line follows by the reverse triangle inequality. Thus 3. follows.

$\square$

It is often enough to take $\delta=\sqrt{\varepsilon}$. The following is a useful corollary.

Corollary. For a closed f, if $f(\bar{x}) \leq \inf f + \varepsilon$. Then, there is $\hat{x} \in \bar{x} + \sqrt{\varepsilon} B$ and

$$y \in \partial_L f(\tilde{x}) \quad \text{with} \quad \|y\| \leq \varepsilon.$$

Proof: Apply Ekeland's with the chain rule to conclude that

$$0 \in \partial_L f(\tilde{x}) + \sqrt{\varepsilon} B.$$

$\square$

Inverse Problems

One of the core goals of variational analysis is to understand the behavior of solutions of equations

$$\text{(P)} \qquad F(x) = y$$

for a given y. As a first step let's assume that $F: E \to Y$ is C^1 and suppose there exist $(\bar{x}, \bar{y}) \in E \times Y$ s.t. $F(\bar{x}) = \bar{y}$. The inverse function theorem gives a way to reason about "nearby problems".

Theorem (Inverse Function Theorem): Suppose F is a C^1 map and $\nabla F(\bar{x})$ is an invertible operator onto $\mathbf{Y}$. Then, there is a map $G: U \to \underline{E}$ defined in a neighborhood U of $\bar{y} = F(\bar{x})$ s.t. it is differentiable around $\bar{y}$ and

$$\nabla G(\bar{y}) = \nabla F(\bar{x})^{-1} \text{ and } F \circ G(y) = y \quad \forall y \in U.$$

$\dashv$

Intuitively, if we perturb $\bar{y}$ to y, there will be a nearby solution x to $F(x) = y$.

Corollary: There exists a constant $K > 0$ $\left(\sigma_{min}(\nabla F(x))/2\right)$ such that if (x, y) are close to $(\bar{x}, \bar{y})$ then

$$(*) \inf_{z \in F^{-1}(y)} \|z - x\| =: \operatorname{dist}(x, F^{-1}(y)) \leq K \|F(x) - y\|.$$

Hint: Start with F linear and then use the fact that you can linearly approximate G.

Question: What happens when we have more general equations $F(x)=y$? What properties imply $(\star)$?

We will take a more general template where we consider a set-valued map $\Phi: E \rightrightarrows Y$.

Def: We say that Φ is metrically regular at $\bar{x}$ for $\bar{y} \in \Phi(\bar{y})$ if there exists $\kappa>0$ so that

$$(\star^2) \qquad \operatorname{dist}(x, \Phi^{-1}(y)) \le \kappa \operatorname{dist}(\Phi(x), y),$$

for all (x,y) near $(\bar{x},\bar{y})$.

In optimization this is useful whenever we are interested in finding a critical point $0 \in \partial f(\bar{x})$, (or a saddle). It tells us that the problem is rather stable.
In turn it also have algorithmic consequences (that we will cover next class).

To try to answer our question, let's take a step back and consider F C^1.

Note that x solves (P) if and only if it minimizes $\|F(\cdot)-y\|$.

Moreover, x solves (P) if and only if it minimizes $\|F(\cdot)-y\|+\delta\|\cdot-x\|$ for any $\delta>0$ small enough. To see this, suppose x minimizes

$$\|F(\cdot)-y\|+\delta\|\cdot-x\|$$

but $F(x)\neq y$. Then, subdifferential calculus yield

$$\nabla F(x)^T \underset{\text{unit norm}}{u} \in \delta B. \qquad (\circ\!\circ\!\circ)$$

Since $\nabla F(\cdot)$ is continuous we have $\|\nabla F(x)^T u\| \geq \dfrac{\sigma_{min}(\nabla F(\bar{x}))}{2}$ for x close to $\bar{x}$ and so $(\circ\!\circ\!\circ)$ fails for δ small.

This characterization motivated Ioffe to generalize the invertibility of $\nabla F(\bar{x})$.

Lemma (Ioffe '79) Suppose Φ is given by

$$\Phi(x) = \begin{cases} F(x) & x \in S \\ \emptyset & \text{otherwise} \end{cases}$$

where F is continuous and $S \subseteq E$ is a closed set. Suppose that Φ is not metrically regular at $\bar{x} \in S$. Then, there is an arbitrarily small $\delta > 0$, a y close to $\bar{y} = F(\bar{x})$ and x close to $\bar{x}$ minimizing

$$\|F(\cdot) - y\| + \delta \|\cdot - x\| \text{ over } S,$$

but $F(x) \neq y$. $\dashv$