LECTURE 19

LECTURE OUTLINE

- Proximal minimization algorithm
- Extensions

****

Consider minimization of closed proper convex $f : \Re^n \mapsto (-\infty, +\infty]$ using a different type of approximation:

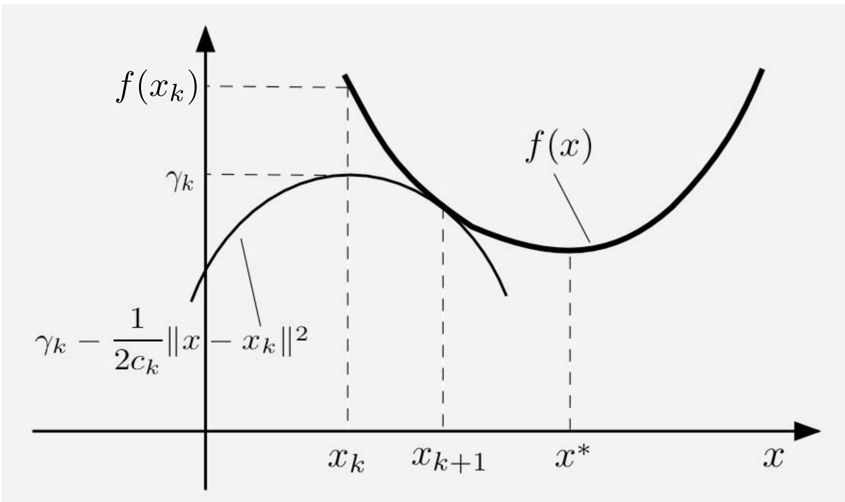
- Regularization in place of linearization
- Add a quadratic term to f to make it strictly convex and "well-behaved"
- Refine the approximation at each iteration by changing the quadratic term

PROXIMAL MINIMIZATION ALGORITHM

- A general algorithm for convex fn minimization

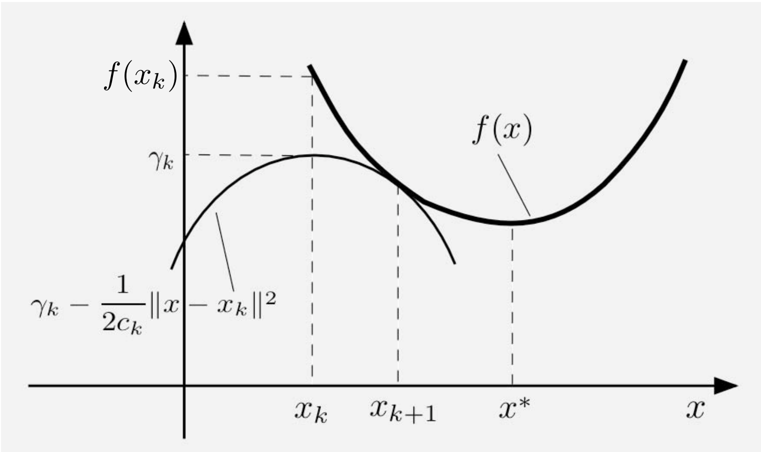
$$x_{k+1} \in \arg\min_{x \in \Re^n} \left\{ f(x) + \frac{1}{2c_k} \|x - x_k\|^2 \right\}$$

  - $f : \Re^n \mapsto (-\infty, \infty]$ is closed proper convex
  - c_k is a positive scalar parameter
  - x_0 is arbitrary starting point



- x_{k+1} exists because of the quadratic.
- Note it does not have the instability problem of cutting plane method
- If x_k is optimal, $x_{k+1} = x_k$.
- If $\sum_k c_k = \infty$, $f(x_k) \to f^*$ and $\{x_k\}$ converges to some optimal solution if one exists.

CONVERGENCE



- Some basic properties: For all k

$$(x_k - x_{k+1})/c_k \in \partial f(x_{k+1})$$

so x_k to x_{k+1} move is "nearly" a subgradient step.

- For all k and $y \in \Re^n$

$$\|x_{k+1}-y\|^2 \leq \|x_k-y\|^2-2c_k\big(f(x_{k+1})-f(y)\big)-\|x_k-x_{k+1}\|^2$$

Distance to the optimum is improved.

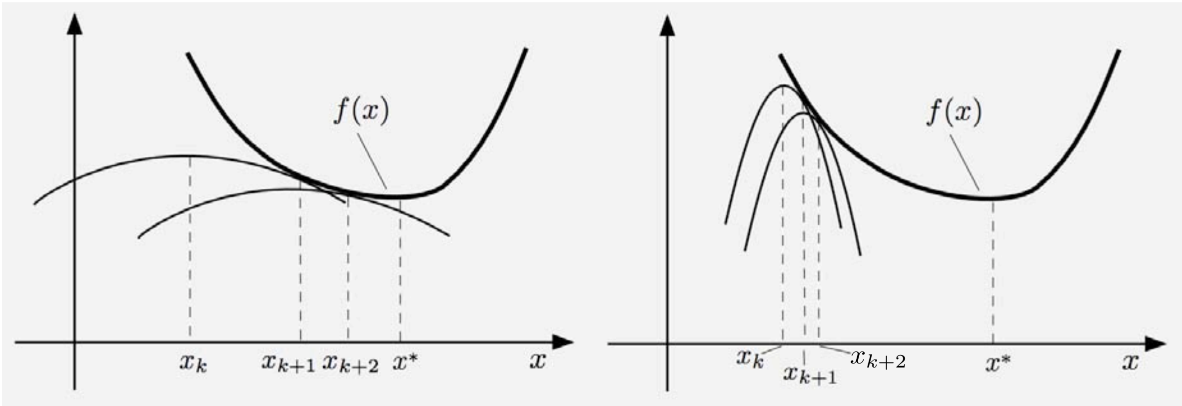
- Convergence mechanism:

$$f(x_{k+1}) + \frac{1}{2c_k}\|x_{k+1} - x_k\|^2 < f(x_k).$$

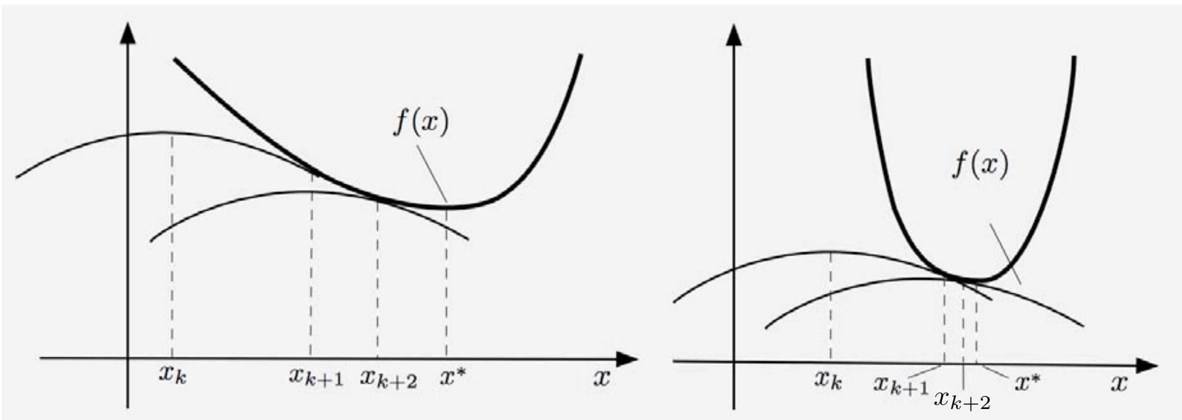
Cost improves by at least $\frac{1}{2c_k}\|x_{k+1} - x_k\|^2$, and this is sufficient to guarantee convergence.

RATE OF CONVERGENCE I

- Role of penalty parameter c_k:



- Role of growth properties of f near optimal solution set:



RATE OF CONVERGENCE II

- Assume that for some scalars $\beta > 0$, $\delta > 0$, and $\alpha \geq 1$,

$$f^* + \beta\big(d(x)\big)^\alpha \leq f(x), \quad \forall\, x \in \Re^n \;\; \text{with } d(x) \leq \delta$$

where

$$d(x) = \min_{x^* \in X^*} \|x - x^*\|$$

i.e., **growth of order α from optimal solution set X^*.**

- If $\alpha = 2$ and $\lim_{k\to\infty} c_k = \bar{c}$, then

$$\limsup_{k\to\infty} \frac{d(x_{k+1})}{d(x_k)} \leq \frac{1}{1 + \beta\bar{c}}$$

linear convergence.

- If $1 < \alpha < 2$, then

$$\limsup_{k\to\infty} \frac{d(x_{k+1})}{\big(d(x_k)\big)^{1/(\alpha-1)}} < \infty$$

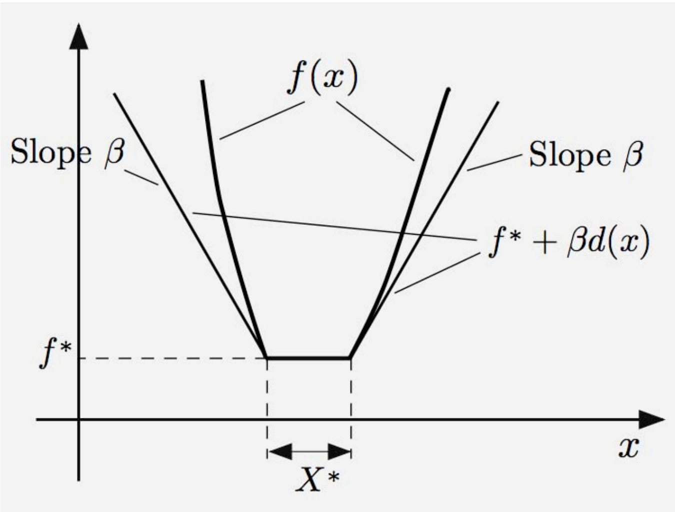
superlinear convergence.

FINITE CONVERGENCE

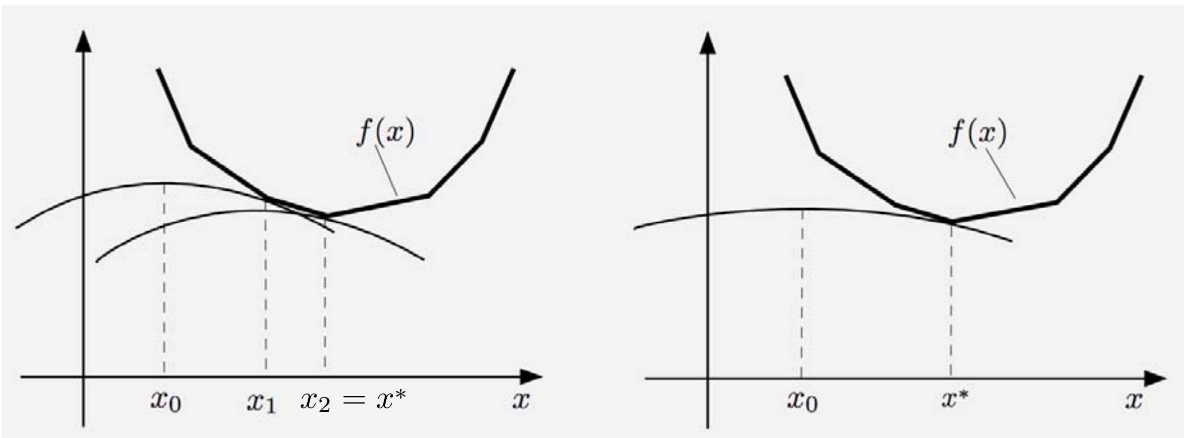
- Assume growth order $\alpha = 1$:

$$f^* + \beta d(x) \leq f(x), \qquad \forall\ x \in \Re^n,$$

e.g., f is polyhedral.

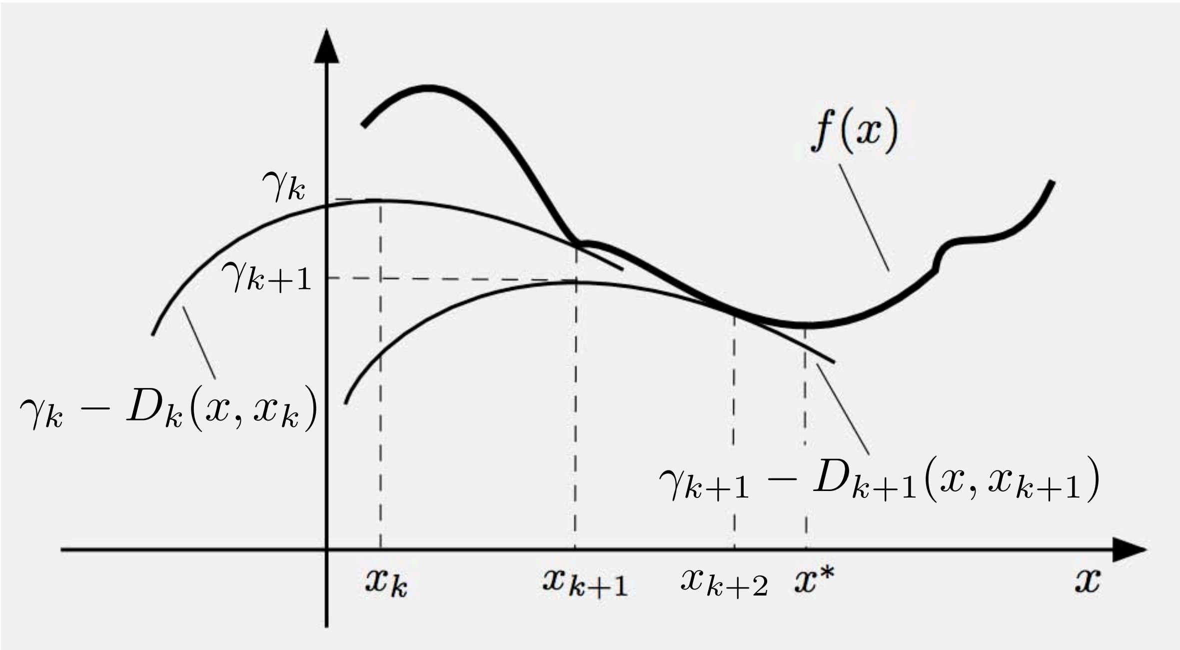


- Method converges finitely (in a single step for c_0 sufficiently large).



IMPORTANT EXTENSIONS

- Replace quadratic regularization by more general proximal term.
- Allow nonconvex f.



- Combine with linearization of f (we will focus on this first).

6.253 Convex Analysis and Optimization
Spring 2012