$$t f(x) + \underbrace{\phi(x)}_{\text{Barrier function}}$$

$t \leftarrow \mu t$ "Barrier" parameter, $\mu > 1$.

○

unstructured Hessian → first-order

structured Hessian / do I need high-accuracy solution's → second-order

first-order:

can I distribute my computation / do I want to solve an extension of my problem → ADMM

→ parallel CD is also possible

→ PGD (APGD), CD, SGD

do I want simple implementation / do I know CD updates in closed form

→ CD

→ PGD, SGD

→ if I need to

---

cvx prob $\Longrightarrow$ cvx solution set.

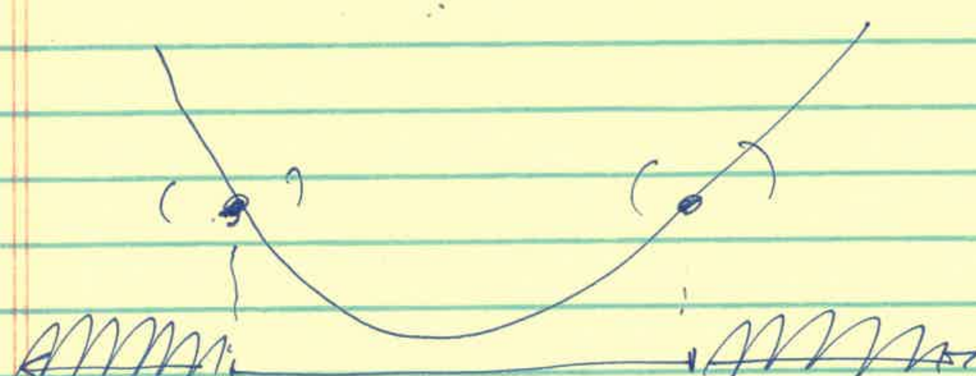
$$x^*(\varepsilon) = \arg\min_{\text{st } x \in C} f(x) + \frac{\varepsilon}{2}\|x\|_2^2$$

claim: as $\varepsilon \to 0$, $x^*(\varepsilon) \to$ min l_2-norm solution to $\min_{x \in C} f(x)$. | variational analysis |

$$X = \arg\min_{x \in C} f(x)$$

$$\min_{x \in X} \|x\|_2^2 \to \text{unique sol}$$

ie there a unique solution with minimal l_2 norm



$$f(y) \geq f(x) + \nabla f(x)^T (y-x) + \frac{L}{2}\|x-y\|_2^2$$

$\log\log\log(1/\varepsilon)$

<u>Thm</u> (Nesterov). For any first-order method, and any $x^{(0)}$. and iter k (not too large)

$\exists$ a function f. (cvx) with L-Lipschitz grad

s.t. $f(x^{(k)}) - f^* \geq \frac{c}{k^2}$.

First-order method: updates
$x^{(k+1)} \in \text{span}\{x^{(0)}, \nabla f(x^{(0)}), \ldots \nabla f(x^{(k)})\}$

$$x^T A x + (x_1 - 1)^2$$

$$A = \begin{bmatrix} -1 & 2 & -1 & & & \\ & -1 & 2 & -1 & & \\ & & & \ddots & & \\ & & & & \ddots & \end{bmatrix}$$

---

smoothed analysis for LPs (Spielman & Teng)

(Klee-Minty cube?)

$$\begin{aligned} \min \quad & c^T x \\ \text{s.t.} \quad & Ax \geq b \\ & x \geq 0. \end{aligned}$$

$$\min f(x) \quad \text{s.t.} \quad Ax = b.$$

$$L(x,u) = f(x) + u^T(Ax - b)$$
$$= f(x) + (A^Tu)^Tx - b^Tu$$

$$g(u) = \min_x L(x,u)$$
$$= -f^*(-A^Tu) - b^Tu$$

$$f^*(y) = \max_x \left(x^Ty - f(x)\right)$$
$$-f^*(y) = \min_x \left(f(x) - x^Ty\right)$$

$$u^+ = u + t\nabla g(u)$$
dual gradient ascent.

$$\nabla g(u) = A\nabla f^*(-A^Tu) - b.$$

Fact. $x \in \partial f(y) \iff y \in \partial f^*(x)$.

special case. f is strictly cvx $\Rightarrow$ f^* is diff.

$$x \in \partial f(y) \iff y = \nabla f^*(x)$$
$$\Updownarrow$$
$$y = \underset{z}{\operatorname{argmin}}\ f(z) - x^Tz$$

sg: $0 \in \partial f(y) - x$

i.e. $x \in \partial f(y)$

if f also diff.

$x = \nabla f(y) \iff y = \nabla f^*(x)$

i.e. $(\nabla f)^{-1} = \nabla f^*$

$$\nabla f^*(-A^Tu) = \underset{z}{\operatorname{argmin}}\ f(z) + u^TAz$$

dual grad ascent:

$$x^+ = \underset{z}{\operatorname{argmin}} \; f(z) + u^T A z$$

$$u^+ = u + t(Ax^+ - b)$$

ADMM: suppose
x replaced by (x, z)

& augment criterion by $\frac{\rho}{2} \| Ax + Bz - c \|_2^2$

only diff between dual ascent and ADMM:

- dual ascent: compute full gradient
- ADMM: splits up min over x & z into two steps

mirror descent:

given f

You define ϕ.

$$\nabla \phi^* (\nabla \phi(x) - \nabla f(x))$$