11.5 COMPLEXITY ANALYSIS VIA SELF CONCORDANCE

CONVERGENCE ANALYSIS:

OF OUTER ITERATIONS:

GAP: $\frac{m}{\mu^k t} < \varepsilon \rightsquigarrow \left\lceil \frac{\log\left(\frac{m}{t\varepsilon}\right)}{\log \mu} \right\rceil = K$

THE FIRST ITER.

HOW TO BOUND INNER ITERATIONS?:

NESTEROV & NEMIROVSKI

DEF: CONVEX

$f: \mathbb{R} \to \mathbb{R}$ IS SELF CONCORDANT IF

$|f'''(x)| \leq 2\, f''(x)^{3/2}$

EX:

- f LINEAR ... $0 \leq 0$
- f QUADRATIC ... $0 \leq k \cdot 2 \quad k > 0$
- $f = -\log x$

$f''(x) = \frac{1}{x^2}, \quad f'''(x) = -2/x^3$

$|2/x^3| = 2 \cdot \left(\frac{1}{x^2}\right)^{3/2}$

IN DEFINITION

: 2 IS JUST FOR EASIER COMPUTATION.

COULD BE ANY K

- LET $|f'''(x)| \le K \cdot f''(x)^{3/2}$

$g = \frac{K^2}{4} f(x)$

$|g'''(x)| = \frac{K^2}{4}|f'''(x)| \le \frac{K^3}{4}\left(f''(x)\right)^{3/2}$

$= \frac{K^3}{4}\left(\frac{4}{K^2} \cdot g''(x)\right)^{3/2} = 2\, g''(x)^{3/2}$

→ JUST BY SCALING

THE $3/2$ IS THE IMPORTANT

Q. $g(y) = f(ay+b)$ IS SELF CONCORDANT IFF f IS.

$g''(y) = a^2 f''(x)$

$g'''(y) = a^3 f'''(x) \quad x = ay+b$

$|g'''(y)| \le 2|g''(x)|^{3/2}$

$|a^3 f'''(x)| \le 2(a^2 f''(x))^{3/2}$

BOUND ON THIRD DERIVATIVE INDEPENDENT OF AFFINE COORD. CHANGE

DEF: $f: \mathbb{R}^n \to \mathbb{R}$ IS SELF-CONCORDANT IF $g(t) = f(x + tv)$ IS SELFCONCORDANT FOR EVERY $v \in \mathbb{R}^n$.

f, g SELF-CONCORDANT $\Rightarrow$ af, $f+g$, $f(Ax+b)$ ARE SELF-CONCORDANT $(a \geq 0)$

$f(x) = -\sum \log(b_i - a_i^T x)$ IS SELF CONCORDANT

USAGE: f STRICTLY CONVEX & SC $\frac{1}{2}$

$$\lambda(x) = \left(\nabla f(x)^T Hf(x)^{-1} \nabla f(x)\right)$$

POS DEF FOR f SELF-CONCORD

$g: \mathbb{R} \to \mathbb{R}$: SELF CONCORD: $\forall\ 0 \leq t \leq g''(0)^{-\frac{1}{2}}$

$$\frac{g''(0)}{\left(1 + t\, g''(0)^{1/2}\right)^2} \leq g''(t) \leq \frac{g''(0)}{\left(1 - t\, g''(0)^{1/2}\right)^2}$$

LET $g(t) = f(x + tv)$ $\quad g : \mathbb{R} \to \mathbb{R}$

$\Rightarrow$ BY ALOT COMPUTATION: $\quad v$ DESCENT DIRECTION

$$\inf_{t \geq 0} g(t) \geq g(0) + \lambda(x) + \log(1 - \lambda(x))$$

$$\Rightarrow f(x^*) \geq f(x) + \lambda(x) + \log(1 - \lambda(x))$$

FOR $0 < \lambda \leq 0.68$

$$-(\lambda + \log(1-\lambda)) \leq \lambda^2$$

$\Rightarrow$

$$f(x^*) \geq f(x) - \lambda(x)^2$$

USEFUL FOR BOUNDING GAP.

IF $\lambda(x)^2 < \varepsilon \Rightarrow$ CLOSE TO OPTIMUM.

NEWTON' METHOD (GUARDED / DAMPED)

X_0 STARTING POINT, ε TOLERANCE

$\Delta x_k = -Hf(x_k)^{-1} \nabla f(x_k)$

$\lambda^2 = \nabla f(x_k) Hf(x_k)^{-1} \nabla f(x_k)$

IF $\lambda^2/2 \le \varepsilon$ STOP X_k IS GOOD

FIND t S.T. ← INSTEAD OF MINIMIZING t

$f(x_k + t\,\Delta x_k) \le f(x_k) + \alpha t \nabla f(x) \Delta x_k$

(START $t=1$ IF NOT $t = t \cdot \beta$

$\alpha \in (0, 0.5)$ $\beta \in (0,1)$)

CRITERIUM 4)

$X_{k+1} = X_k + t \Delta x_k$

THEOREM:

ETA :-)

$\exists \gamma > 0$, $\exists\, 0 < \eta \le 1/4$ DEPENDING ONLY ON α, β S.T.

IF $\lambda(x_k) > \eta$: $f(x_{k+1}) - f(x_k) \le -\gamma$

IF $\lambda(x_k) \le \eta$:

$$2\lambda(X_{k+1}) \le \left(2\lambda(X_k)\right)^2$$

↳ METHOD HAS "TWO PHASES"

LET $\lambda(x_k) \le \eta \quad \forall l > k \quad \lambda(x_l) \le \eta$

$$2\lambda(x_l) \le (2\lambda(x_k))^{2^{l-k}} \le (2\eta)^{2^{l-k}} \le \left(\frac{1}{2}\right)^{2^{l-k}}$$

$$\Rightarrow f(x_l) - f(x^*) \le \left(\frac{1}{2}\right)^{2^{l-k+1}} \le \varepsilon \quad ?$$

$$l-k \ge \log_2 \log_2 \frac{1}{\varepsilon}$$

$\Rightarrow$ # ITERATIONS:

$$\frac{f(x_0) - f(x^*)}{\gamma} + \log_2 \log_2 {}^{1}/_{\varepsilon}$$

1 CENTERING STEP:

- SOLVED x FOR $tf(x)+\phi(x)$
- SOLVING FOR $\mu t f(x)+\phi(x)$ .. x^+

NEED TO BOUND: (USE KKT)

$$\mu t f(x)+\phi(x)-\mu t f(x^+)-\phi(x^+) =$$

$$= \mu t f(x) - \mu t f(x^+) - \sum_{i=1}^{m} \log(-g_i(x)) + \sum \log(-g_i(x^+)) =$$

$$= \mu t f(x) - \mu t f(x^+) + \sum \log\left(\frac{-\mu t}{-g_i(x)\mu t} g_i(x^+)\right) =$$

$$= \mu t f(x) - \mu t f(x^+) + \sum \log(-\mu t \lambda_i g_i(x^+)) - m \log \mu \le$$

WHERE $\lambda_i = -\frac{1}{t g_i(x)}$

$$\le \mu t f(x) - \mu t f(x^+) - \mu t \sum \lambda_i g_i(x^+) - m - m \log \mu$$

SINCE $\log a \le a-1$ FOR $a>0$

$$= \mu t f(x_0) - \mu t \left(\boxed{f(x^*) + \sum_{i=1}^{m} \lambda_i g_i(x^*)} \right) - m - m \log \mu$$

$$h(\lambda) = \inf_z \left(f(z) + \sum \lambda_i g_i(z) \right) \le$$

$$\le \mu t f(x_0) - \mu t h(\lambda) - m - m \log \mu$$

$$\left[h(\lambda) = f(x) - m/t \right]$$

$$= \mu m - m - m \log \mu = m(\mu - 1 - \log \mu)$$

$\Rightarrow$ # STEPS FOR 1 INNER ITERATION IS

$$\frac{m(\mu - 1 - \log \mu)}{\gamma} + c$$

DOES NOT DEPEND ON t OR f ! (OR $\mathbb{R}^n \to \mathbb{R}$)

LARGE μ = MORE STEPS (AS INTUITION)

SUMMING UP:

$$N = \left\lceil \frac{\log(m/t \cdot \varepsilon)}{\log \mu} \right\rceil \cdot \left(\frac{m(\mu - 1 - \log \mu)}{\gamma} + c \right)$$

FOR $\mu = 1 + 1/\sqrt{m}$:

INITIAL t

$$N = O\left(\sqrt{m} \log\left(\frac{m/t}{\varepsilon}\right)\right)$$

[IN PRACTICE CHOOSE $\mu = 10, \ldots 20$ JUST FOR COMPLEXITY]

STEPS FOR FIXED GAP REDUCTION $O(\sqrt{m})$