15-459N

11/8/13

Nesterov's Gradient method

Goal: $\min_{y \in \mathbb{R}^m} f(y)$, f convex & diff

Properties of f.

1) Lipshitz of ∇ constant C_2

$$\|\nabla f(y) - \nabla f(z)\|_2 \leq C_2 \|y-z\|_2 \quad \forall y, z$$

2) Strong Convexity C_{sc}

$$\langle \nabla f(y) - \nabla f(z), y-z \rangle \geq \frac{C_{sc}}{2} \|y-z\|_2^2 \quad \forall y, z$$

---

eg $f(x) = \frac{1}{2} x^T A x$ A spd.

$\nabla f(x) = Ax$

1)

$$\|\nabla f(y) - \nabla f(z)\|_2 = \|Ay - Az\|_2 \leq C_2 \|y-z\|_2$$

$$\|A(y-z)\|_2 \leq C_2 \|y-z\|_2$$

$x = y - z$

$\|A(x)\|_2 \le C_2 \|x\|_2$

$\left\| \frac{A(x)}{x} \right\|_2 \le C_2 \Rightarrow \lambda_{max}(A) \le C_2$

---

2) $(A(y) - A(z))^T (y - z) \geq \frac{C_{SC}}{2} \|y - z\|_2^2$

$x = y - z$

$A(x)^T x \geq \frac{C_{SC}}{2} x^T x$

$\frac{C_{SS}}{2} \le \frac{x^T A x}{x^T x} \Rightarrow \frac{C_{SC}}{2} \le \lambda_{min}(A)$

Thm (Nesterov)

A) $f: \mathbb{R}^m \to \mathbb{R}$, ∇f, constants $C_L, C_{SC} > 0$

given $y_0 \in \mathbb{R}^m$ then $\forall y^* \in \mathbb{R}^m$

$$\#\text{iter} \equiv T = \sqrt{\frac{C_L}{C_{SC}}} \ln\left(\frac{4\|y_0 - y^*\|_2}{\varepsilon}\right)$$

returns y_T st $f(y_T) - f(y^*) \leq \varepsilon$

B) Same as A) but $C_{SC} = 0$

$$\#\text{iter} \quad T \equiv 2\sqrt{\frac{C_L}{\varepsilon}} \cdot \|y_0 - y^*\|_2$$

then $f(y_T) - f(y^*) \leq \varepsilon$

Alg: Maxflow (G: graph; $F>0$ flow value; $\varepsilon>0$)

1) Find potential flow y_0 of size F.

2) Run Nesterov from y_0

   1) from y_0
   2) $f = \Phi(y) = \frac{1}{2}\|\text{over}(y)\|_2^2$ retricted to $\tilde{\mathcal{Z}}_F$
   3) $\nabla f = \Pi^{\perp}(y - \text{over}(y))$
   4) $C_2 = 1$ & $C_{SC} = 0$
   5) $T = \frac{2}{\varepsilon}\sqrt{\frac{m}{F}}$
   6) returning $y_T \in \tilde{\mathcal{Z}}_F$

3) rescate of $\left(\frac{1}{1+\varepsilon}\right) y_T = \tilde{y}_T$

4) Drain off over$(\tilde{y}_T)$ say $\bar{y}_T$

5) return $\bar{y}_T$

<u>Thm</u> If F is feasible then Max flow then

1) flow $(\overline{Y}_T) \geq (1-4\varepsilon) F$

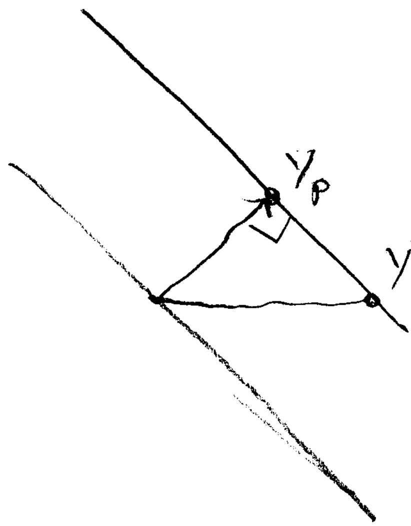
2) run time $O\left(\frac{L}{\varepsilon}\sqrt{\frac{m}{F}}\, m \log^2 n\right)$

<u>pf</u> 1) assume Y_0 is exact!

2) y^* is feasible F-flow.

$\|Y_0 - y^*\|_2 \leq \|y^*\|_2$

$\leq \sqrt{m}\, \|y^*\|_\infty \leq \sqrt{m} \quad y^* \in B_\infty^m$



Nesterov $T' = 2\sqrt{\frac{C_L}{\varepsilon'}}\, \|Y_0 - y^*\|_2$

$C_L = 1 \quad \|Y_0 - y^*\|_2 = \sqrt{m}$

$T' = 2\sqrt{1/\varepsilon'}\, \sqrt{m} \Rightarrow$ error ε'

We ran $T = \frac{2}{\varepsilon}\sqrt{\frac{m}{F}}$

$$\varepsilon' = \varepsilon^2 F$$

$$T' = 2\sqrt{\frac{1}{\varepsilon^2 F}}\sqrt{m} = \frac{2}{\varepsilon}\sqrt{\frac{m}{F}} = T$$

Thus

$$\Phi(Y_T) - \Phi(Y^*) \le \varepsilon^2 F$$

$$\Rightarrow \Phi(Y_T) = \|\mathrm{over}(Y_T)\|_2^2 \le 2\varepsilon^2 F$$

$$D = \{e : \mathrm{over}(Y_T(e)) > \varepsilon\}$$

Why the 2?

$$\varepsilon \sum_{e \in D} |\mathrm{over}(Y_T(e))| \le \sum_e (\mathrm{over}(Y_T)(e))^2 = 2\varepsilon^2 F$$

$$\Rightarrow \|\mathrm{over}(\tilde{Y}_T)\|_1 \le 2\varepsilon F$$

Drain off over $(\tilde{Y}_T)$ get

$$\frac{F}{1+\varepsilon} - 2\varepsilon F = \left(\frac{1-2\varepsilon(1+\varepsilon)}{1+\varepsilon}\right)F$$

$$= \frac{1-2\varepsilon-2\varepsilon^2}{1+\varepsilon} F \doteq (1-3\varepsilon)F$$

Due to solver error we get

$(1-4\varepsilon)F$ flow

Drain

flow γ with $\mathrm{ova}(f)$