Lecture 12.

Huge-Scale Optimization Problems:

Tactics and Strategy.

– Random Optimization Approach

– Gaussian Smoothing

– Random Gradient-free oracles

– Random search for nonsmooth and stochastic optimization

– Simple subgradient methods

– Solving the huge-scale optimization problems

– Random sparse block-coordinate methods

– Examples:Google problem

Notation:

$$\|x\| =< Bx, x >^{1/2},\ x \in E, \qquad \|s\|_* =< s, B^{-1}s >^{1/2},\ s \in E$$

where $B = B^* \succ 0$ is a linear operator from $E \to E^*$.

Denote $uu^* : E^* \to E : uu^*(s) = u\cdot < s, u >,\ s \in E^*$.

We consider functions with different level of smoothness.

- $$f \in C^{0,0}(E) \text{ if } |f(x) - f(y)| \le L_0(f)\|x - y\|,\ x, y \in E$$

- $$f \in C^{1,1}(E) \text{ if } \|\nabla f(x) - \nabla f(y)|_* \le L_1(f)\|x - y\|,\ x, y \in E$$

  This conditions is equivalent to

  $$|f(y) - f(x) - < \nabla f(x), y - x > | \le \frac{1}{2} L_1(f)\|x - y\|^2,\ x, y \in$$

Gaussian smoothing

Assumption: $f : E \to R$ is differentiable along any direction $\forall x \in E$.

Let us form its *Gaussian approximation*

$$f_\mu(x) = \frac{1}{\kappa} \int\limits_E f(x + \mu u) e^{-\frac{1}{2}\|u\|^2} du,$$

$$\text{where } \kappa = \int\limits_E e^{-\frac{1}{2}\|u\|^2} du \quad (1)$$

Note: If f is convex and $g \in \partial f(x)$, then

$$f_\mu(x) \geq \frac{1}{\kappa} \int\limits_E [f(x) + \mu < g, u >] e^{-\frac{1}{2}\|u\|^2} du = f(x) \quad (2)$$

f_μ is better than f!

- If f is convex, then f_μ is also convex
- If $f \in C^{0,0}$, then $f_\mu \in C^{0,0}$ and $L_0(f_\mu) \leq L_0(f)$.
- If $f \in C^{1,1}$, then $f_\mu \in C^{1,1}$ and $L_1(f_\mu) \leq L_1(f)$.

Taking ln from (1) and differenting this identity in B, we get

$$\frac{1}{\kappa} \int\limits_E u u^* e^{-\frac{1}{2}\|u\|^2} du = B^{-1}. \quad (3)$$

Taking a scalar product with B, we get

$$\frac{1}{\kappa} \int\limits_E \|u\|^2 e^{-\frac{1}{2}\|u\|^2} du = n. \quad (4)$$

So,

$$M_0 \overset{(1)}{=\!=} 1, \; M_2 \overset{(4)}{=\!=} n.$$

Lemma 1 *For* $p \in [0,2]$, *we have* $M_p \in n^{p/2}$.

$$\textit{If } p \geq 2, \textit{ then } n^{p/2} \leq M_p \leq (p+n)^{p/2}. \tag{5}$$

Theorem 1 *Let* $f \in C^{0,0}(E)$, *then*

$$|f_\mu(x) - f(x)| \leq \mu L_0(f) n^{1/2}, \ x \in E. \tag{6}$$

If $f \in C^{1,1}(E)$, *then*

$$|f_\mu(x) - f(x)| \leq \frac{\mu^2}{2} L_1(f) n, \ x \in E. \tag{7}$$

For any $\mu > 0$, function f_μ is differentiable and

$$\nabla f_\mu(x) = \frac{1}{\kappa} \int_E \frac{f(x+\mu u) - f(x)}{\mu} e^{-\frac{1}{2}\|u\|^2} B u du. \tag{8}$$

Lemma 2 *Let* $f \in C^{0,0}(E)$ *and* $\mu > 0$. *Then* $f_\mu \in C^{1,1}(E)$ *with*

$$L_1(f_\mu) = \frac{2n^{1/2}}{\mu} L_0(f). \tag{9}$$

Denote

$$f'(x,u) = \lim_{\alpha \to 0} \frac{1}{\alpha}[f(x+\alpha u) - f(x)].$$

$$\nabla f_0(x) = \frac{1}{\kappa} \int_E f'(x,u) e^{-\frac{1}{2}\|u\|^2} B u du. \tag{10}$$

If f is differentiable at x,then

$$\nabla f_0(x) = \frac{1}{\kappa} \int_E < \nabla f(x), u > e^{-\frac{1}{2}\|u\|^2} B u du \overset{(3)}{=} \nabla f(x). \tag{11}$$

Random gradient-free oracle

Let random vector $u \in E$ have Gaussian distribution with correlation operator B^{-1}.

- 1. Generate random $u \in E$ and return $g_\mu(x) = \frac{f(x+\mu u)-f(x)}{\mu} \cdot Bu$
- 2. Generate random $u \in E$ and return

$$\bar{g}_\mu(x) = \frac{f(x+\mu u) - f(x-\mu u)}{2\mu} \cdot Bu. \qquad (12)$$

- 3. Generate random $u \in E$ and return $g_0(x) = f'(x,u) \cdot Bu$

Note:

$$E_u(g_0(x)) = \nabla f_0(x) \in \partial f(x).$$

Theorem 2 *1. If f is differentiable at x, then*

$$E_u(\|g_0(x)\|_*^2) \leq (n+4)\|\nabla f_0(x)\|_*^2. \qquad (13)$$

2. Let f be convex. Denote $D(x) = diam \partial f(x)$. Then, for any $x \in E$ we have

$$E_u(\|g_0(x)\|_*^2) \leq (n+4)(\|\nabla f_0(x)\|_*^2 + nD^2(x)). \qquad (14)$$

3. If $f \in C^{0,0}(E)$, then

$$E_u(\|g_\mu(x)\|_*^2) \leq L_0^2(f)(n+4)^2. \qquad (15)$$

Random search for nonsmooth optimization

Problem:

$$\min \{ f(x) : x \in Q \}$$

where Q ш $\subseteq E$ is a closed convex set, and f is a nonsmooth convex function on E.

Method $\mathcal{RS}_\mu$: Choose $x_0 \in Q$. If $\mu = 0$, we need $D(x_0) =$
Iteration $(k \geq 0)$**:** **a).** Generate u_k and corresponding $g_\mu(x_k)$. **b).** Compute $x_{k+1} = \pi_Q(x_k - h_k B^{-1} g_\mu(x_k))$.

(16)

Denote: $\mathcal{U}_k = (u_0, .., u_k), \phi_0 = f(x_0), \phi_k = E_{\mathcal{U}_{k-1}}(f(x_k))$, $\bar{x}_N = argmin[f(x) : x \in x_0, .., x_N]$.

Theorem 3 *Let sequence* $\{x_k\}_{k \geq 0}$ *be generated by* $\mathcal{RS}_0$. *Then, for any* $N \geq 0$ *we have*

$$\sum_{k=0}^{N} h_k(\phi_k - f^*) \leq \frac{1}{2}\|x_0 - x^*\|^2 + \frac{n+4}{2} L_0^2(f) \sum_{k=0}^{N} h_k^2. \quad (17)$$

Note:

If $\|x_0 - x^*\| \leq R$, $h_k = R/((n+1)^{1/2}(N+1)^{1/2} L_0(f)$,

$$N = \frac{n+4}{\epsilon^2} L_0^2(f) R^2 \Longrightarrow E_{\mathcal{U}_{N-1}}(f(\bar{x}_N)) - f^* \leq \epsilon.$$

Denote $S_N = \Sigma_{k=0}^N h_k$.

Theorem 4 *Let sequence* $\{x_k\}_{k\geq 0}$ *be generated by* $\mathcal{RS}_\mu$ *with* $\mu > 0$*. Then, for any* $N \geq 0$ *we have*

$$\frac{1}{S_N} \sum_{k=0}^{N} h_k(\phi_k - f^*) \leq \mu L_0(f) n^{1/2} + \frac{1}{S_N}\left[\frac{1}{2}\|x_0 - x^*\|^2 + \right.$$
$$\left. + \frac{n+4^2}{2} L_0^2(f) \sum_{k=0}^{N} h_k^2 \right]. \tag{18}$$

Proof. Denote $r_k = \|x_k - x^*\|$. Then

$$r_{k+1}^2 \leq \|x_k - h_k g_\mu(x_k) - x^*\|^2 =$$
$$= r_k^2 - 2h_k < g_\mu(x_k), x_k - x^* > + h_k^2 \|g_\mu(x_k)\|_*^2.$$

Using representation (8)and (15), we get

$$E_{U_k}(r_{k+1}^2) \leq r_k^2 - 2h_k < \bigtriangledown f_\mu(x_k), x_k - x^* > +$$

$$+ h_k^2 (n+4)^2 L_0^2(f) \overset{(2)}{\leq} r_k^2 - 2h_k(f(x_k) - f_\mu(x^*)) + h_k^2(n+4)^2 L_0^2(f).$$

Taking now expectation in U_{k-1}, we obtain

$$E_{U_k}(r_{k+1}^2) \leq E_{U_{k-1}}(r_k^2) - 2h_k(\phi_k - f_\mu(x^*)) + h_k^2 (n+4)^2 L_0^2(f).$$

It remains to note that $f_\mu(x^*) \overset{(6)}{\leq} f^* + \mu L_0(f) n^{1/2}$. $\square$

Note: We can choose

$$\mu = \frac{\epsilon}{2L_0(f)n^{1/2}}, \;\; h_k = \frac{R}{(n+4)(N+4)^{1/2} L_0(f)},$$
$$N = \frac{4(n+4)^2}{\epsilon^2} L_0^2(f) R^2. \tag{19}$$

Random search for stochastic optimization

Problem:

$$\min \{ \ f(x) : \ x \in Q \ \}$$

$$f(x) = E_{\xi}[F(x,\xi)] = \int_{\Xi} F(x,\xi) dP(\xi), \qquad (20)$$

where ξ is a random vector with probability distribution $P(\xi)$, $\xi \in \Xi$.

Assumption: $f \in C^{0,0}(E)$ - convex.

Random stochastic gradient-free oracles:

- 1. Generate random $u \in E$, $\xi \in \Xi$. Return $s_{\mu}(x) = \frac{F(x+\mu u,\xi) - F(x,\xi)}{\mu} \cdot Bu$
- 2. Generate random $u \in E$, $\xi \in \Xi$. Return

$$\bar{s}_{\mu}(x) = \frac{F(x+\mu u,\xi) - F(x-\mu u,\xi)}{2\mu} \cdot Bu. \qquad (21)$$

- 3. Generate random $u \in E$, $\xi \in \Xi$ and return $s_0(x) = F'(x,\xi) \cdot Bu$

Method $\mathcal{SS}_\mu$: Choose $x_0 \in Q$.

Iteration $(k \geq 0)$**:**

a). For $x_k \in Q$, generate random vectors $\xi_k \in \Xi$ and u_k.

b). Compute $s_\mu(x_k)$, and $x_{k+1} = \pi_Q(x_k - h_k B^{-1} s_\mu(x_k))$.

$$\tag{22}$$

Theorem 5 *Let* $L_0(F(\cdot,\xi)) \leq L \ \forall \xi \in Xi$*. Assume the sequence* $\{x_k\}_{k\geq 0}$ *be generated by* $\mathcal{SS}_\mu$ *with* $\mu > 0$*. Then, for any* $N \geq 0$ *we have*

$$\frac{1}{S_N}\sum_{k=0}^{N} h_k(\phi_k - f^*) \leq \mu L n^{1/2} + \frac{1}{S_N}\left[\frac{1}{2}\|x_0 - x^*\|^2 + + \frac{n+4^2}{2} L^2 \sum_{k=0}^{N} h_k^2\right], \tag{23}$$

where $\phi_k = E_{\mathcal{U}_{k-1},\mathcal{P}_{k-1}}(f(x_k))$*, and* $\mathcal{P}_k = \{\xi_0, .., \xi_k\}$*.*

Proof.

$$r_{k+1}^2 \leq r_k^2 - 2h_k < s_\mu(x_k), x_k - x^* > + h_k^2\|s_\mu(x_k)\|_*^2.$$

By assumption $\|s_\mu(x_k)\| \leq L\|u_k\|^2$. Since $E_\xi(s_\mu(x)) = g_\mu(x)$, we have

$$E_{u_k,\xi_k}(r_{k+1}^2) \leq r_k^2 + E_{u_k}(-2h_k < g_\mu(x_k), x_k - x^* >) +$$

$$+ h_k^2 L^2 \|u_k\|^4 \overset{(8),(5)}{\leq} r_k^2 - 2h_k(\nabla f_\mu(x_k), x_k - x^* > +$$

$$+ h_k^2(n+4)L^2 \leq r_k^2 - 2h_k(f_\mu(x_k) - f_\mu(x^*)) + h_k^2(n+4)L^2.$$

Taking now expectation in U_{k-1} and $\mathcal{P}_{k-1}$, we get

$$E_{\mathcal{U}_k,\mathcal{P}_k}(r_{k+1}^2) \overset{(2)}{\leq} E_{\mathcal{U}_{k-1},\mathcal{P}_{k-1}}(r_k^2) - 2h_k(\phi_k - f_\mu(x^*)) + h_k^2(n+4)L^2.$$

It remains to note that $f_\mu(x^*) \overset{(6)}{\leq} f^* + \mu L n^{1/2}$. □

Note: *Choosing the parameters in accordance to (19), we can solve (20) in* $O(\frac{n^2}{\epsilon^2})$ *iterations.*

Simple subgradient methods

Problem

$$\min \{ \ f(x) : \ x \in Q \ \}$$

where $Q \subseteq R^n$ is a closed convex set, and f is a convex function. *Assumptions:* 1) We are able to compute a subgradient $f'(x)$ $\forall x \in Q$
2) $\|f'(x)\| \leq L(f), \ x \in Q$
3) Q is simple
4) The optimal value f^* is known.
Consider the following optimization scheme:

$$x_0 \in Q, \ x_{k+1} = \pi_Q \left(x_k - \frac{f(x_k) - f^*}{\|f'(x_k)\|^2} f'(x_k) \right), \quad k \geq 0. \tag{24}$$

Denote $f_k^* = \min_{0 \leq i \leq k} f(x_i), \ L_k(f) = \max_{0 \leq i \leq k} \|f'(x_i)\|$.

Theorem 6

$$f_k^* - f^* \leq \frac{L_k(f)\|x_0 - \pi_{X^*}(x_0)}{(k+1)^{1/2}}. \tag{25}$$

Moreover, we have

$$\|x_k - x^*\| \leq \|x_0 - x^*\| \quad \forall x^* \in X^*. \tag{26}$$

Hence, $\{x_k\}$ *converges to a single point in* X^**.*

Corollary *Assume that* X^* *has a recession direction* d_**. Then*

$$\|x_k - \pi_{X^*}(x_0)\| \leq \|x_0 - \pi_{X^*}(x_0)\|, \tag{27}$$

$$< d_*, x_k > \geq < d_*, x_0 > .$$

Problem

$$\min_{x \in Q} \{ f(x) : g(x) \leq 0 \},$$

where $Q \subseteq R^n$ is a closed convex set, and f and g are closed convex function.

Method $SG_N(h)$.
For $k = 0, .., N-1$ iterate: If $g(x_k) > h \| g'(x_k) \|$, then (A): $x_{k+1} = \pi_Q(x_k - \frac{g(x_k)}{\|g'(x_k)\|^2} g'(x$ else (B): $x_{k+1} = \pi_Q(x_k - \frac{h}{\|f'(x_k)\|} f'(x_k))$.

(28)

Denote $\mathcal{F}_k \subseteq \{0, .., k\}-$ the set of numbers of iterations (B)
Denote

$$f_k^* = \min_{i \in \mathcal{F}_k} f(x_i), \quad L_k(f) = \max_{i \in \mathcal{F}_k} \| f'(x_i) \|,$$
$$g_k^* = \max_{i \in \mathcal{F}_k} g(x_i), \quad L_k(g) = \max_{i \notin \mathcal{F}_k} \| f'(x_i) \|.$$

Theorem 7 *IF* $N > r_0^2(x)/h^2$, *then* $\mathcal{F}_N \neq 0$ *and*

$$f_N^* - f(x) \leq h L_N(f), \; g_N^* \leq h L_N(g). \tag{29}$$

Our goal: Find an $\epsilon-$ solution $\bar{x} \in Q$:

$$f(\bar{x}) - f(x^*) \leq \epsilon, \; g(\bar{x}) \leq \epsilon. \tag{30}$$

If $L(f), \; L(g)$ are known, we can take

$$h = \frac{\epsilon}{\max\{L(f), L(g)\}}. \tag{31}$$

Solving the huge-scale optimization problems

Assumption: The main operator $A(x) = Ax + b$ has a *block structure.*
$A \in R^{M\times N}$ is divided on mn blocks $A_{i,j} \in R^{r_i \times g_j}$

$$x = (x^1, .., x^n) \in R^N, \; u = (u^1, .., u^m) \in R^M,$$

$$A_{i,j} \neq 0, \; j \in \sigma_b(A_i) \subseteq \{1, .., n\}.$$

$$\textbf{Problem: } \min\{f(x) \equiv f_0(u^0(x))\} \tag{32}$$

$$\psi(u) \equiv \max_{1\leq i\leq m} f_i(u^i), \; g(x) \equiv \psi(u(x)) \leq 0,$$

$$u^i(x) = \sum_{j\in\sigma_b(A_i)} A_{i,j}x^j - b^i, \quad i = 1, .., m$$

$$x^j \in Q_j, \quad j = 1, .., n.$$

Goal: Solve this problem by (20).

$$x_{k+1} = \pi_Q(x_k + d_k)$$

Assume that d_k *is already computed and is block-sparse. We must update* x_{k+1} *and* u_{k+1}.

$$x^j_{k+1} = \pi_{Q_j}(x^j_k + d^j_k), \quad j \in \sigma_b(d_k), \tag{33}$$

$$x^j_{k+1} = x^j_k, \quad \text{otherwise}$$

Define $\delta^j_k = x^j_{k+1} - x^j_k, \quad j \in \sigma_b(d_k)$.
If $u_k = Ax_k - b$ is already computed, then $u_{k+1} = Ax_{k+1} - b$ can be obtained by a *sequence of recursive updates.*

Start: $u_+ = u_k$.
For $j \in \sigma_b(d_k),\ i \in \sigma_b(A^j)$ iterate: **1.** Update $u_+^i = u_+^i A_{i,j}\delta_k^j$ **2.** Compute $f_i(u_+^i)$ and $f_i'(u_+^i)$. **3.** Update $\psi(u_+)$ and $i_+ = argmax_{1\leq l\leq m} f_l(u_+^l)$.

(34)

Final:$u_{k+1} = u_+$.
Thus,

$$g(x_{k+1}) = \psi(u_+), \quad g'(x_{k+1}) = A_{i_+}^T f_{i_+}'(u_{k+1}^{i_+}), \tag{35}$$

$$\|g'(x_{k+1})\|^2 = \sum_{j\in\sigma_b(A_{l_+})} \|g'(x_{k+1})^j\|^2.$$

If $u_+^0 \neq u_k^0$ and $g(x_{k+1}) \leq h\|g'(x_{k+1})\|$, then $i_+ = 0$ and we need to compute

$$f'(x) = A_0^T f_0'(u_+^0).$$

Note: d_{k+1} have the same block sparsity pattern as $A_{i_+}^T$.

Theorem 8 *Let* $r_i \equiv 1$ *and* $q_j \equiv 1$.
Assume that filling of matrix A *is uniform:*

$$p(A_i) \leq c_r,\ i = 1, .., m, \quad p(A^j) \leq c_q,\ j = 1, .., n$$

Then the computational costs (20) does not exceed

$$(1 + log_2 M)c_q c_r.$$

The computational cost of one iteration grows logarithmicaly with dimension of the image space.

Coordinate descent schemes

Problem:

$$\min_{x \in R^n} f(x),$$

where f is convex and differentiable on R^n.

Decomposition of R^n on n subspaces:

$$R^N = \bigoplus_{i=1}^{n} R^{n_i}, \quad N = \sum_{i=1}^{n} n_i.$$

$$I_N = (U_1, .., U_n) \in R^{N \times N}, \; U_i \in R^{N \times n_i}, \quad i = 1, .., n.$$

$$x = \sum_{i=1}^{n} U_i x^{(i)}, \; x^{(i)} \in R^{n_i}, \; i = 1, .., n.$$

Define *the Partial gradient of $f(x)$ in $x^{(i)}$* :

$$f_i'(x) = U_i^T \nabla f(x) \in R^{n_i}, \; x \in R^N.$$

Assumption

$$\|f_i'(x+U_i h_i) - f_i'(x)\|_{(i)}^* \leq L_i \|h_i\|_{(i)}, \quad h_i \in R^{n_i}, \; i = 1, .., n, x \in \tag{36}$$

Hence, we can prove

$$f(x+U_i h_i) \leq f(x) + < f_i'(x), h_i > + \frac{L_i}{2} \|h_i\|_{(i)}^2, \; x \in R^N, \; h_i \in \tag{37}$$

Define the optimal coordinate steps:

$$T_i(x) = x - \frac{1}{L_i} U_i f_i'(x), \; i = 1, .., n.$$

Then, in view (37), we get

$$f(x) - f(T_i(x)) \geq \frac{1}{2L_i} \left(\|f_i'(x)\|_{(i)}^* \right)^2, \; i = 1, .., n. \tag{38}$$

Define a random counter $\mathcal{R}_\alpha$, $\alpha \in R$, *which generates an integer number* $i \in 1,..,n$ *with probability*

$$p_\alpha^{(i)} = L_i^\alpha \cdot \left[L_j^\alpha \right]^{-1}, \; i = 1,..,n. \tag{39}$$

Method $RCDM(\alpha, x_0)$.
For $k \geq 0$ iterate: 1) Choose $i_k = \mathcal{R}_\alpha$. 2) Update $x_{k+1} = T_{i_k}(x_k)$.

(40)

$$\|x\|_\alpha = \left[\sum_{i=1}^{n} L_i^\alpha \|x^{(i)}\|_{(i)}^2 \right]^{1/2}, \; x \in R^N, \tag{41}$$

$$\|g\|_\alpha^* = \left[\sum_{i=1}^{n} L_i^{-\alpha} \left(\|g^{(i)}\|_{(i)}^* \right)^2 \right]^{1/2}, \; g \in R^N$$

Define

$$\xi_k = \{i_0, .., i_k\}, \quad \phi_k = E_{\xi_{k-1}} f(x_k).$$

Theorem 9

$$\phi_k - f^* \leq \frac{2}{k+4} \left[\sum_{j=1}^{n} L_j^\alpha \right] \cdot R_{1-\alpha}^2(x_0), \tag{42}$$

where $R_\beta(x_0) = \max_x \{ \max_{x_* \in X^*} \|x - x^*\|_\beta : \; f(x) \leq f(x_0) \}$.

Random sparse block-coordinate methods

$$\textbf{Problem:}\ \min\{g(x) \equiv \max_{1\le i\le m} f_i(u^i))\} \qquad (43)$$

$$u^i(x) = \sum_{j\in\sigma_b(A_i)} A_{i,j}x^j - b^i, \quad i = 1,..,m$$

$$x^j \in Q_j, \quad j = 1,..,n,$$

where convex functions $f_i(u_i)$, $i = 1,..,m$, have bounded subgradients, and the sets $Q_j \subseteq R^{q_j}$, $j = 1,..,n$ are closed and convex. Define an active index $i(x)$: $g(x) = f_{i(x)}(u^{(i(x))}(x))$. Then

$g'(x) = A_i^T f'_{i(x)}(u^{(i(x))}(x))$, $\sigma_b(g'(x)) \subseteq \sigma_b(A_{i(x)})$, $p_b(g'(x)) \le p$

Define a random variable $\xi(x)$, which generates indexes from $\sigma_b(A_{i(x)})$ with probabilities $1/p_b(A_{i(x)})$.

Assume the optimal value g^* of the problem is known. *Define* a random vector variable $Next(x)$:

Next(x).

1. Compute $h(x) = \frac{g(x)-g^*}{\|g'(x)\|^2}$. Generate $j(x) = \xi(x)$.

2. Define $[Next(x)]^{j(x)} = \pi_{Q_{j(x)}}(x - h(x)A_i^T f'_{i(x)}(u^{(i(x))}(x)))$

3. For other indexes $j \ne j(x)$, define $[Next(x)]^j = x^j$.

(44)

Method RSBC: Choose $x_0 \in Q \equiv \Pi_{j=1}^n Q_j$. Compute $u_0 =$
kth iteration ($k \geq 0$). **a.** Generate $j_k = \xi(x_k)$ and update $x_{k+1} = Next(x_k)$. **b.** Update $u_{k+1} = u_k + A^{j_k}\left(x_{k+1}^{j_k} - x_k^{j_k}\right)$, computing in para values $f_i(u_{k+1}^i),\ i \in \sigma_b(A^{j_k})$ with immediate evaluation of $g($

(45)

Denote $g_k^* = \min_{0 \leq i \leq k} g(x_i)$.

Theorem 10 *Let* $p_b(A_i) \leq r,\ i = 1, .., m$. *Then, for any* $k \geq 0$ *we have*

$$E\left([g_k^* - g^*]^2\right) \leq \frac{rL^2(g)\|x_0 - \pi_{X^*}(x_0)\|^2}{k+1}, \tag{46}$$

$$E\left([x_k - x_*]^2\right) \leq \|x_0 - x_*\|^2,\ \forall x_* \in X_*. \tag{47}$$

The computational costs of one iteration:

Let $r_i \equiv q_j \equiv 1$, and $p(A^j) \leq q \Longrightarrow (1 + \log_2 m)q$. Hence, we get

$$E\left([g_k^* - g^*]^2\right) \leq \epsilon^2 \qquad \text{in} \quad O\left(\frac{qrL(g)^2\|x_0 - x_*\|^2}{\epsilon^2} \log_2 m\right) \text{ oper}$$

Google problem

Let $E \in R^{N \times N}$ be an incidence matrix of a graph. Denote $\bar{E} = E \cdot diag(E^T e)^{-1}$, where $e = (1, .., 1)^T$.

Since,$\bar{E}^T e = e \Longrightarrow \bar{E}$ - stochastic.
Our Goal: Find $x^* \geq 0 : \;\; \bar{E}x^* = x^*, \;\; x^* \neq 0.$
Note: If the degree of each node is small, then the computational of partial derivatives is cheap.
Let $\rho(A)$ - the maximal eigenvalue of a nonegative A. Then,

$$\rho(A) = \min_{x \geq 0} \max_{1 \leq i \leq n} \frac{1}{x^{(i)}} < A^T e_i, x >$$

Moreover, the minimum is attained at the corresponding eigenvector. Since $\rho(\bar{E}) = 1$, then x^* is a solution to the following

$$\textbf{Problem}: \;\; g(x) = \max_{1 \leq i \leq N} [< A^T e_i, x > - x^{(i)}] \rightarrow \min_{x \geq 0}. \tag{48}$$

Note $g^* = 0$ and X^* - convex cone. Let us solve the problem by method (24).
Let $x_0 = e$. Note:$\{x_k\}$ is well separated from zero: from (27) we have

$$< x^*, e > \leq < x^*, x_k > \leq \|x^*\|_1 \cdot \|x_k\|_\infty = < x^*, e > \cdot \|x_k\|_\infty.$$

Thus,
Our goal: Find $\bar{x} \geq 0 : \;\; \|\bar{x}\|_\infty \geq 1$ and $g(\bar{x}) \leq \epsilon$