1 定义

def 1.

$$\forall i, x_i(k+1) = x_i(k) + \mathrm{I}_{\{i \in T_k\}} U_i\left(x^i(k)\right) \tag{1}$$

T_k 表示第 k 步 *master* 接受更新的 *workers* 的集合。

i 个 *worker* 上面的 *local* 数值

$$x^i(k) \tag{2}$$

第 k 步的全局 x 值：

$$x(k) = (x_1(k), \cdots, x_n(k)) \tag{3}$$

def 2.

$$\left[x_i^i(k)\right]' = prox_{g_i}^{\eta}\left(x_i^i(k) - \eta \nabla_i f\left(x^i(k)\right)\right) \tag{4}$$

$$\left[x_i(k)\right]' = prox_{g_i}^{\eta}\left(x_i(k) - \eta \nabla_i f(x(k))\right) \tag{5}$$

$$\begin{aligned} U_i\left(x^i(k)\right) &= prox_{g_i}^{\eta}\left(x_i^i(k) - \eta \nabla_i f\left(x^i(k)\right)\right) - x_i^i(k) \\ &= \left[x_i^i(k)\right]' - x_i^i(k) \end{aligned} \tag{6}$$

def 3.

$$M_F(x(k+1)) = x(k+1) - prox_g^{\eta}\{x(k+1) - \eta \nabla f(x(k+1))\} \tag{7}$$

$$M_F(x(k+1))_i = x_i(k+1) - prox_{g_i}^{\eta}\{x_i(k+1) - \eta \nabla_i f(x(k+1))\} \tag{8}$$

2 算法描述

首先按照 workers 的数目 n，把变量 x 分成 n 部分，即 $x = (x_1, x_2, \cdots, x_n)$。master 上面储存有真正的 x 值，也叫做全局 x 值，第 k 步的全局 x 值记作 $x(k)$；每个 workers 上面储存的是局部 x 值，对于第 i 个 worker，第 k 步的局部 x 值记作 $x^i(k)$。每一步迭代中，每一个 worker 处理其中的一部分，也就是说 $i-th$worker 处理 x_i，第 k 步时，计算更新程度 $U_i(x^i(k))$。对于 master，每次接收 workers 上的更新程度，并作更新 $x_i(k+1) = x_i(t) + \mathrm{I}_{\{i \in T_k\}} U_i(x^i(t))$（$T_k$ 表示第 k 步时，master 会接受的 workers 集合）。因为第 i 个 worker 只负责 x_i 的更新，因而有 $x_i^i(k) = x_i(k)$，也就是说，$x_i^i(k)$ 的值保持最新。

Algorithm 1 Decouped Asyn-SCD, For worker i

1: **repeat**
2: Obtain the parameter x from the master(shared memory or parameter server).
3: Evaluate the gradient of the i^{th} component over parameter x, denoted by $\nabla_i f(x)$.
4: Evaluate the proximal operate $[x_i]^{'} = prox_{g_i}^{\eta}(x_i - \eta \nabla_i f(x))$.
5: Send the update information $U_i(x) = [x_i]^{'} - x_i$.
6: **until** procedure of master ends

Algorithm 2 Decouped Asyn-SCD, For master

1: **for** $k = 1, 2, \cdots$ **do**
2: Get $U_i(x^i) = [x_i^i]^{'} - x_i^i$.
3: Update parameter with $x_i(k+1) = x_i(t) + U_i(x^i(t))$.
4: $k = k + 1$
5: **end for**

3 定理

Assumption 1. *For all $x \in R^n$ and all $t \in R$, we have*

$$|\nabla_i g(x + te_i) - \nabla_i g(x)| \leq L_i |t| \tag{9}$$

L_i is called component Lipshitz constants. $L_{max} = \max L_i, \quad i = 1, 2, \cdots, n$

Assumption 2. *For all $x, y \in R^n$, we have*

$$\|\nabla f(x) - \nabla f(y)\| \leq L_f \|x - y\| \tag{10}$$

Assumption 3. *For all $x, y \in R^n$, we have*

$$\|\nabla g(x) - \nabla g(y)\| \leq L \|x - y\| \tag{11}$$

Assumption 4. *F is bounded from below F^*.*

Assumption 5. $\forall i, j$ $x_i(k)$ 对应的迭代次数和 $x_i^j(k)$ 对应的迭代次数的差小于 s，s 称为最大 *delay.*

Th 1. 在 *Assumption1*、*2*、*3*、*4*、*5*满足的情形下，有：

$$\frac{1}{T}\sum_{k=0}^{T}[M_F(x(k))]^2 \leq \frac{2ns}{T}(\eta L_{max} + 1)^2 \frac{2}{\frac{1}{\eta} - L_f - 2Ls}[F(x(0)) - F^*] \tag{12}$$

4 证明

$$\begin{aligned}
M_F\left(x\left(k+1\right)\right)_i =& x_i\left(k+1\right) - prox_{g_i}^{\eta}\left\{x_i\left(k+1\right) - \eta\nabla_i f\left(x\left(k+1\right)\right)\right\} \\
\leq& x_i\left(k\right) + u_i\left(x^i\left(k\right)\right) - prox_{g_i}^{\eta}\left\{x_i\left(k+1\right) - \eta\nabla_i f\left(x\left(k+1\right)\right)\right\} \\
=& x_i\left(k\right) + \left[x_i^i\left(k\right)\right]' - x_i^i\left(k\right) - \left[x_i\left(k+1\right)\right]' \\
=& x_i\left(k\right) - x_i^i\left(k\right) + \left[x_i^i\left(k\right)\right]' - \left[x_i\left(k+1\right)\right]'
\end{aligned} \tag{13}$$

Now, Focus on $\left[x_i^i\left(k\right)\right]' - \left[x_i\left(k+1\right)\right]'$

$g_i\left(x\right)$ 是凸函数，根据凸函数的单调性，有

$$\forall\, u \in \partial g_i\left(\left[x_i\left(k+1\right)\right]'\right) \quad \forall\, v \in \partial g_i\left(\left[x_i^i\left(k\right)\right]'\right) \tag{14}$$

使得

$$\left(u-v\right)\left[\left[x_i\left(k+1\right)\right]' - \left[x_i^i\left(k\right)\right]'\right] \geq 0 \tag{15}$$

根据 $\left[x_i\left(k+1\right)\right]'$ 以及 $\left[x_i^i\left(k\right)\right]'$ 的定义。

有

$$0 \in \partial g_i\left(\left[x_i\left(k+1\right)\right]'\right) + \nabla_i f^T\left(x\left(k\right)\right) + \frac{1}{\eta}\left[x_i\left(k+1\right)\right]' \tag{16}$$

以及

$$0 \in \partial g_i\left(\left[x_i^i\left(k\right)\right]'\right) + \nabla_i f^T\left(x\left(k\right)\right) + \frac{1}{\eta}\left[x_i^i\left(k\right)\right]' \tag{17}$$

取

$$\hat{u} \in \partial g_i\left(\left[x_i\left(k+1\right)\right]'\right) \quad \text{sit. } \hat{u} + \nabla_i f^T\left(x\left(k\right)\right) + \frac{1}{\eta}\left[x_i\left(k+1\right)\right]' \tag{18}$$

$$\hat{v} \in \partial g_i\left(\left[x_i^i\left(k\right)\right]'\right) \quad \text{sit. } \hat{v} + \nabla_i f^T\left(x\left(k\right)\right) + \frac{1}{\eta}\left[x_i^i\left(k\right)\right]' \tag{19}$$

带入15即得：

$$\left\{\partial_i f\left(x^i\left(k\right)\right) - \partial_i f\left(x\left(k+1\right)\right) + \frac{1}{\eta}\left(\left[x_i^i\left(k\right)\right]' - \left[x_i\left(k+1\right)\right]'\right)\right\}^T\left(\left[x_i\left(k+1\right)\right]' - \left[x_i^i\left(k\right)\right]'\right) \geq 0 \tag{20}$$

移项化简之即得：

$$\left\|\left[x_i^i\left(k\right)\right]' - \left[x_i\left(k+1\right)\right]'\right\|_2^2 \quad \leq \quad \eta\left[\nabla_i f\left(x^i\left(k\right)\right) - \nabla_i f\left(x\left(k\right)\right)\right]^T\left(\left[x_i\left(k+1\right)\right]' - \left[x_i^i\left(k\right)\right]'\right) \tag{21}$$

即

$$\left\|\left[x_i^i\left(k\right)\right]' - \left[x_i\left(k+1\right)\right]'\right\|_2^2 \leq \left\{\eta\left[\nabla_i f\left(x^i\left(k\right)\right) - \nabla_i f\left(x\left(k\right)\right)\right] + \left[x_i\left(k\right) - x_i^i\left(k\right)\right]\right\}^T\left(\left[x_i\left(k+1\right)\right]' - \left[x_i^i\left(k\right)\right]'\right) \tag{22}$$

$$\text{(Cauchy inequality)} \leq \left\|\eta\left[\nabla_i f\left(x^i\left(k\right)\right) - \nabla_i f\left(x\left(k\right)\right)\right] + \left(x_i\left(k\right) - x_i^i\left(k\right)\right)\right\|_2\left\|\left[x_i\left(k+1\right)\right]' - \left[x_i^i\left(k\right)\right]'\right\|_2 \tag{23}$$

即

$$\begin{aligned}
\left\|\left[x_i^i(k)\right]' - \left[x_i(k+1)\right]'\right\|_2 &\leq \left\|\eta\left[\nabla_i f\left(x^i(k)\right) - \nabla_i f(x(k))\right] + \left(x_i(k) - x_i^i(k)\right)\right\|_2 \\
\text{(triangle inequality)} &\leq \eta\left\|\nabla_i f\left(x^i(k)\right) - \nabla_i f(x(k))\right\|_2 + \left\|\left[x_i(k+1)\right]' - \left[x_i^i(k)\right]'\right\|_2 \\
(\nabla_i\ f\text{Lipschitz Continuous Gradient}) &\leq \eta L_i\left\|x(k) - x^i(k)\right\|_2 + \left\|x_i(k) - x_i^i(k)\right\|_2 \\
&\leq (\eta L_i + 1)\left\|x(k) - x_i^i(k)\right\|_2
\end{aligned} \tag{24}$$

将24带入 $\left[M_F(x(k+1))\right]_i$ 之中

$$\begin{aligned}
\left[M_F(x(k+1))\right]_2^2 &\leq \left\|x_i(k) - x_i^i(k)\right\|_2^2 + \left\|\left[x_i^i(k)\right]' - \left[x_i(k+1)\right]'\right\|_2^2 \\
&\leq 2(\eta L_i + 1)^2\left\|x(k) - x^i(k)\right\|_2^2
\end{aligned} \tag{25}$$

Difference between local $x^{i_k}(k)$ and global $x(k)$

$$\begin{aligned}
\left\|x(k) - x^i(k)\right\|_2 &= \sqrt{\sum_{j=1}^{n}\left\|x_j(k) - x_j^i(k)\right\|_2^2} \\
&\leq \sqrt{\sum_{j=1}^{n}\left\|\sum_{t=(k-s)_+}^{k-1}\left(x_i(t+1) - x_i(t)\right)\right\|_2^2} \\
&= \left\|\sum_{t=(k-s)_+}^{k-1}\left(x(t+1) - x(t)\right)\right\|_2 \\
\text{(triangle inequality)} &\leq \sum_{t=(k-s)_+}^{k-1}\left\|x(t+1) - x(t)\right\|_2
\end{aligned} \tag{26}$$

As $F(x) = f(x) + \sum_{i=1}^{n} g_i(x_i)$, Then

$$F(x(k+1)) - F(x(k)) = f(x(k+1)) - f(x(k)) + \sum_{i=1}^{n}\left[g_i(x_i(k+1)) - g_i(x_i(k))\right] \tag{27}$$

For $f(x_i(k+1)) - f(x_i(k))$, 根据 ∇f Lipschitz Continuous Gradient

$$f(x(k+1)) - f(x(k)) \leq \nabla f(x(k))^T(x(k+1) - x(k)) + \frac{L_f}{2}\left\|x(k+1) - x(k)\right\|_2^2 \tag{28}$$

设 $i \in T_k$，则

$$x_i(k+1) = \operatorname{prox}_{g_i}^{\eta}\left(x_i(k) - \eta\nabla_i f\left(x^i(k)\right)\right) \tag{29}$$

则对于 $\forall z \in \operatorname{dom} g$

$$\begin{aligned}
&g_i(x_i(k+1)) + \nabla_i f\left(x^i(k)\right)^T(x_i(k+1) - x_i(k)) + \frac{1}{2\eta}\left\|x_i(k+1) - x_i(k)\right\|_2^2 \\
&\leq g_i(z) + \nabla_i f\left(x^i(k)\right)^T(z - x_i(k)) + \frac{1}{2\eta}\left\|z - x_i(k)\right\|_2^2
\end{aligned} \tag{30}$$

取 $z=x_i(k)$，则

$$g_i(x_i(k+1))-g_i(x_i(k))\leq -\left\{\nabla_i f\left(x^j(k)\right)^T(x_i(k+1)-x_i(k))+\frac{1}{2\eta}\|x_i(k+1)-x_i(k)\|_2^2\right\} \tag{31}$$

把2831带入27之中，得

$$\begin{aligned}
F(x(k+1))\leq\ & \nabla f(x(k))^T(x(k+1)x(k))+\frac{L_f}{2}\|x(k+1)-x(k)\|_2^2\\
&-\sum_{i=1}^{n}\left\{\frac{1}{2\eta}\|(x_i(k+1)-x_i(k)\|_2^2+\nabla_i f\left(x^i(k)\right)^T(x_i(k+1)-x_i(k))\right\}\\
&=\frac{1}{2}\left(L_f-\frac{1}{\eta}\right)\|(x(k+1)-x(k))\|_2^2+\sum_{i=1}^{n}\left(\nabla_i f(x(k))-\nabla_i f\left(x^i(k)\right)\right)^T(x_i(k+1)-x_i(k))\\
&\leq\frac{1}{2}\left(L_f-\frac{1}{\eta}\right)\|(x(k+1)-x(k))\|_2^2+\sum_{i=1}^{n}\left\|\nabla_i f(x(k))-\nabla_i f\left(x^i(k)\right)\right\|_2\|x_i(k+1)-x_i(k)\|_2\\
&\leq\frac{1}{2}\left(L_f-\frac{1}{\eta}\right)\|(x(k+1)-x(k))\|_2^2+\sum_{i=1}^{n}L_i\|x(k+1)-x(k)\|_2\|x_i(k+1)-x_i(k)\|_2\\
&\leq\frac{1}{2}\left(L_f-\frac{1}{\eta}\right)\|(x(k+1)-x(k))\|_2^2+\sum_{i=1}^{n}L_i\|x_i(k+1)-x_i(k)\|_2\sum_{t=(k-1)_+}^{k-1}\|x(t+1)-x(t)\|_2\\
&\leq\frac{1}{2}\left(L_f-\frac{1}{\eta}\right)\|(x(k+1)-x(k))\|_2^2+L\|x(k+1)-x(k)\|_2\sum_{t=(k-1)_+}^{k-1}\|x(t+1)-x(t)\|_2\\
&\leq\frac{1}{2}\left(L_f-\frac{1}{\eta}\right)\|(x(k+1)-x(k))\|_2^2+\frac{L}{2}\sum_{t=(k-1)}^{k-1}\left(\|x(k+1)-x(k)\|_2+\|x(t+1)-x(t)\|_2\right)\\
&\leq\frac{1}{2}\left(L_f-\frac{1}{\eta}+Ls\right)\|(x(k+1)-x(k))\|_2^2+\frac{L}{2}\sum_{t=(k-1)}^{k-1}\|x(t+1)-x(t)\|_2
\end{aligned} \tag{32}$$

改变求和顺序，有不等式:

$$\sum_{k=m}^{n}\sum_{t(k-1)_+}^{k-1}\|x(t+1)-x(t)\|_2^2\leq s\sum_{k=m}^{n}\|x(k)-x(k-1)\|_2^2 \tag{33}$$

对于32，从 $k=m$ 加到 $k=n$，得

$$\begin{aligned}
F((n)-(m))&\leq\frac{1}{2}\left(L_f-\frac{1}{\eta}+Ls\right)\sum_{k=m}^{n-1}\|x(k+1)-x(k)\|_2^2+\frac{L}{2}\sum_{k=m}^{n-1}\sum_{t=(k-s)_+}^{k-1}\|x(t+1)-x(t)\|_2^2\\
&\leq\frac{1}{2}\left(L_f-\frac{1}{\eta}+Ls\right)\sum_{k=m}^{n-1}\|x(k+1)-x(k)\|_2^2+\frac{Ls}{2}\sum_{k=m}^{n}\|x(k)-x(k-1)\|_2^2\\
&=\frac{1}{2}\left(L_f-\frac{1}{\eta}+2Ls\right)\sum_{k=m}^{n}\|x(k+1)-x(k)\|_2^2
\end{aligned} \tag{34}$$

取 $\eta < \dfrac{1}{L_f + 2Ls}$，则 $L_f - \dfrac{1}{\eta} + 2Ls < 0$，则

$$\sum_{k=m}^{n} \|x(k+1) - x(k)\|_2^2 \le \frac{2}{\frac{1}{\eta} - L_f - 2Ls} \left(F(x(m)) - F(x(n))\right) \tag{35}$$

$$\begin{aligned}
\frac{1}{T}\sum_{k=0}^{T} \left[M_F(x(k))\right]^2 &\le \frac{2}{T}(\eta L_{\max}+1)^2 \sum_{k=0}^{T}\sum_{i=1}^{n} \left\|x(k) - x^i(k)\right\|_2^2 \\
&\le \frac{2}{T}(\eta L_{\max}+1)^2 \sum_{k=0}^{T}\sum_{i=1}^{n} \sum_{t=(k-s)_+}^{k-1} \|x(t+1) - x(t)\|_2^2 \qquad (36) \\
&= \frac{2n}{T}(\eta L_{\max}+1)^2 \sum_{k=0}^{T} \sum_{t=(k-s)_+}^{k-1} \|x(k+1) - x(k)\|_2^2 \\
&\le \frac{2ns}{T}(\eta L_{\max}+1)^2 \sum_{k=0}^{T} \|x(k+1) - x(k)\|_2^2 \\
&\le \frac{2ns}{T}(\eta L_{\max}+1)^2 \frac{2}{\frac{1}{\eta} - 2L_f - 2Ls}\left(F(x(0)) - F(x(T))\right) \quad \le \frac{2ns}{T}(\eta L_{\max}+1)^2 \frac{2}{\frac{1}{\eta} - 2L_f - 2Ls}\,(F
\end{aligned}$$

(37)