Consensus Optimization

6:36 PM

Consensus Optimizat...

Consensus using Monotone Operator Splitting

Consensus Optimization:

Compare this with
[Consensus using Proximal ADMM]

$$\left(\begin{array}{l} \gamma \sum_{i=1}^{n} f_i(x_i) = f(x) \\ \\ x_1 = x_2 = \dots = x_n \end{array}\right) = \gamma\left(\sum_{i=1}^{n} f_i(x) + I_C(x)\right) \quad \text{if } C = \{(x_1,\dots,x_n) \mid x_1 = \dots = x_n\}$$

Consensus Optimization Problem

the minimizer x^* ⊡ $\underbrace{\partial\left(\sum_{i=1}^{n} f_i(x_i)\right)}_{B(x)} + \underbrace{\underbrace{\partial I_C(x)}_{N_C(x)}}_{A(x)} \ni 0$

thm: proximal operator is the resolvent of subdifferential operator

- the resolvent of a subdifferential operator w.r.t function ψ is the proximal map, i.e.

$$R_{\partial\psi}(⊟) = (I+\lambda\psi)^{-1}⊟ = \operatorname{prox}_{\lambda\psi}(⊟) = \underset{⊟}{\operatorname{argmin}}\left(\lambda\psi(⊟) + \tfrac{1}{2}\|⊟ - ⊟\|_2^2\right) = \underset{⊟}{\operatorname{argmin}}\left(\psi(⊟) + \tfrac{1}{2\lambda}\|⊟ - ⊟\|_2^2\right)$$

by definition $\operatorname{prox}_{\varpi}(⊟) = \underset{⊟}{\operatorname{argmin}}\left(\varpi(⊟) + \tfrac{1}{2}\|⊟ - ⊟\|_2^2\right)$

- $R_{N_C}(x) = \Pi_C(x)$ •resolvent of the normal cone operator of the normal

so, to solve $B(x) + A(x) \ni 0$ we apply D-R splitting as follows: [# eq: Douglas-Rachford splitting (Ernest's notation]

1) $x^{k+\frac{1}{2}} = R_B(z^k) = R_{\partial f}(z^k) = \operatorname{prox}_{\lambda f}(z^k) = \underset{⊟}{\operatorname{argmin}}\left(f(⊟) + \frac{1}{2\lambda}\|⊟ - z^k\|_2^2\right)$ # now note that,

$$\gamma\; \underbrace{f(y)}_{\sum_{i=1}^{n} f_i(y_i)} + \frac{1}{2\lambda}\underbrace{\|y - z^k\|_2^2}_{\sum_{i=1}^{n}\|y_i - z_i^k\|_2^2} = \gamma\left(\sum_{i=1}^{n} \underbrace{\left(f_i(y_i) + \frac{1}{2\lambda}\|y_i - z_i^k\|_2^2\right)}_{\text{note that this is seperable so each part can be solved independently}}\right) = \sum_{i=1}^{n} \gamma\left(f_i(y_i) + \frac{1}{2\lambda}\|y_i - z_i^k\|_2^2\right)$$

$\left\|\begin{bmatrix} a \\ b \end{bmatrix}\right\|_2^2 = a_1^2 + \dots + a_p^2 + b_1^2 + \dots + b_q^2 = \|a\|_2^2 + \|b\|_2^2$

so we can split the iterate $x^{k+\frac{1}{2}}$ into n seperate local iterates,

$\forall_{i\in\{1,\dots,n\}}\quad x_i^{k+\frac{1}{2}} = \underset{y_i}{\operatorname{argmin}}\left(f_i(y_i) + \frac{1}{2\lambda}\|y_i - z_i^k\|_2^2\right)$

each of them are local iteration

2) $\tilde{z}^{k+\frac{1}{2}} := 2x^{k+\frac{1}{2}} - z^k = \begin{bmatrix} 2x_1^{k+\frac{1}{2}} - z_1^k \\ \vdots \\ 2x_n^{k+\frac{1}{2}} - z_n^k \end{bmatrix} \leftrightarrow \forall_{i\in\{1,\dots,n\}}\quad \tilde{z}_i^{k+\frac{1}{2}} = 2x_i^{k+\frac{1}{2}} - z_i^k$

3) $x^{k+1} = R_A(\tilde{z}^{k+\frac{1}{2}}) = R_{N_C}(\tilde{z}^{k+\frac{1}{2}}) = \Pi_C(\tilde{z}^{k+\frac{1}{2}})$

Now we want to find the projection on the consensus set: $C = \{x \mid x_1 = x_2 = \dots = x_n\}$

$\Pi_C(x) = \begin{bmatrix} \bar{x} \\ \vdots \\ \bar{x} \end{bmatrix}$ where $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ # Though this is intuitive, I need to figure out a proof later

$\Pi_C(\tilde{z}^{k+\frac{1}{2}}) = \left(\frac{1}{n}\sum_{i=1}^{n}\tilde{z}_i^{k+\frac{1}{2}}\right)\mathbf{1}_{i=1}^{n}$ # so we can split this one to as well, and localize in

n blocks, note that this is a vector

$\forall_{i\in\{1,\dots,n\}}\quad x_i^{k+1} = \frac{1}{n}\sum_{i=1}^{n}\tilde{z}_i^{k+1/2}$

4) $z^{k+1} := z^k + x^{k+1} - x^{k+\frac{1}{2}}$

As we have decentralized 1), 2), 3), we can apply the same splitting to the 4th iteration as well:

$$\begin{aligned}
\forall_{i\in\{1,\dots,n\}}\quad z_i^{k+1} &= z_i^k + x_i^{k+1} - x_i^{k+1/2} \\
&= z_i^k + \frac{1}{n}\sum_{i=1}^{n}\tilde{z}_i^{k+1/2} - x_i^{k+1/2} \\
&= z_i^k + \frac{1}{n}\sum_{i=1}^{n}\left(2x_i^{k+\frac{1}{2}} - z_i^k\right) - x_i^{k+1/2} \\
&= z_i^k + 2\underbrace{\left(\frac{1}{n}\sum_{i=1}^{n} x_i^{k+1/2}\right)}_{\bar{x}^{k+1/2}} - \underbrace{\left(\frac{1}{n}\sum_{i=1}^{n} z_i^k\right)}_{\bar{z}^k} - x_i^{k+1/2}
\end{aligned}$$

the average of x_is and z_is at kth iterate

$= z_i^k + 2\bar{x}^{k+1/2} - \bar{z}^k - x_i^{k+1/2}$

$\therefore \forall_{i\in\{1,\dots,n\}}\quad z_i^{k+1} = z_i^k + 2\bar{x}^{k+1/2} - \bar{z}^k - x_i^{k+1/2}$

note that $\tilde{x}, \tilde{z}$ related iterateগুলোর
তেমন প্রয়োজন নাই, and as we are
interested in $x = R_B(z)$, $z = C_A C_B(z)$ in
the end, in our iteration for consensus optimization

$\therefore \forall_{i \in \{1,\ldots,n\}} \quad z_i^{k+1} = z_i^k + 2\bar{x}^{k+1/2} - \bar{z}^k - x_i^{k+1/2}$

note that x, z related iterateগুলোর
যেগুলো দরকার নাই, and as we are
interested in $x = R_B(z)$, $z = C_A C_B(z)$ in
the end, in our iteration for consensus optimization
we need only 1) and 4) th iteration.

In Summary, D-R consensus iteration can be written as:

$\forall_{i \in \{1,\ldots,n\}}$

$$x_i^{k+\frac{1}{2}} = \underset{y_i}{\operatorname{argmin}} \left(f_i(y_i) + \frac{1}{2\lambda} \| y_i - z_i^k \|_2^2 \right)$$

$$z_i^{k+1} = z_i^k + 2\bar{x}^{k+1/2} - \bar{z}^k - x_i^{k+1/2}$$

note that the entire process is
decentralized, there is no need for a central unit!

D-R consensus

here $\bar{x}^{k+1/2} = \frac{1}{n} \sum_{i=1}^{n} x_i^{k+1/2}$, $\bar{z}^k = \frac{1}{n} \sum_{i=1}^{n} z_i^k$

* Distributed QP:

$$\begin{pmatrix} \gamma \sum_{i=1}^{n} \frac{1}{2} \| A_i x - b_i \|_2^2 \\ \urcorner \\ \forall_{i \in \{1,\ldots,n\}} \quad F_i x \leq g_i \end{pmatrix} = \left(\underset{x}{\gamma} \sum_{i=1}^{n} \frac{1}{2} \| A_i x - b_i \|_2^2 + \sum_{i=1}^{n} I_{F_i \square - g_i \leq 0}(x) \right)$$

y_i গুলো আসলে x_i [illegible]

$$= \underset{x}{\gamma} \sum_{i=1}^{n} \left(\frac{1}{2} \| A_i x - b_i \|_2^2 + I_{F_i \square - g_i \leq 0}(x) \right)$$

Remember the key to distributed optimization is having uncoupled objective (See Primal Dual decomposition). Now in our case we do not have uncoupled objective, we uncouple them by providing local copy of the variable, and maintain the equality in terms of a consensus constraint such as local_copy_1=local_copy_2=... =local_copy_n

$$= \begin{pmatrix} \gamma \sum_{i=1}^{n} \frac{1}{2} \| A_i y_i - b_i \|_2^2 \\ \urcorner \forall_{i \in \{1,\ldots,n\}} \quad F_i y_i \leq g_i \\ y_1 = y_2 = \cdots = y_n \quad \text{\# all of them equal to } x \end{pmatrix}$$

note $y_1, \ldots, y_n$ are vectors themselves with same dimension as x

$$\begin{pmatrix} \gamma \sum_{i=1}^{n} \frac{1}{2} \| A_i x_i - b_i \|_2^2 + \sum_{i=1}^{n} I_{\{\square | F_i \square \leq g_i\}}(x_i) \\ \urcorner \quad x_1 = x_2 = \cdots = x_n \end{pmatrix}$$

$$= \begin{pmatrix} \gamma \sum_{i=1}^{n} \overbrace{\left(\frac{1}{2} \| A_i x_i - b_i \|_2^2 + I_{\{\square | F_i \square \leq g_i\}}(x_i) \right)}^{f_i(x_i)} \\ \urcorner \\ x_1 = x_2 = \cdots = x_n \end{pmatrix}$$

this form of the optimization
problem is in the consensus optimization form # Consensus Optimization Problem

so, we can apply D-R consensus form # D-R consensus

Then we will have the following optimization problem:

$\forall_{i \in \{1,\ldots,n\}}$

$$x_i^{k+\frac{1}{2}} = \underset{y_i}{\operatorname{argmin}} \left(f_i(y_i) + \frac{1}{2\lambda} \| y_i - z_i^k \|_2^2 \right) = \underset{y_i}{\operatorname{argmin}} \left(\frac{1}{2} \| A_i y_i - b_i \|_2^2 + I_{\{\square | F_i \square \leq g_i\}}(y_i) + \frac{1}{2\lambda} \| y_i - z_i^k \|_2^2 \right)$$

$$= \underset{y_i \in \{\square | F_i \square \leq g_i\}}{\operatorname{argmin}} \left(\frac{1}{2} \| A_i y_i - b_i \|_2^2 + \frac{1}{2\lambda} \| y_i - z_i^k \|_2^2 \right)$$

$$z_i^{k+1} = z_i^k + 2\bar{x}^{k+1/2} - \bar{z}^k - x_i^{k+1/2}$$