Distributed Optimization by Circuits

12:31 PM

Decomposition_using...

Page 1 ✱ Distributed convex optimization:

- Convex problem

  $\downarrow$ Partition into coupled subsysystems

  $\boxed{\text{sub system}}$ $\boxed{\text{sub system}}$ . . . $\boxed{\text{sub system}}$

- divide variables, constraints, objective terms in two groups: $\boxed{\text{local}}$ variables, constraints, objective terms appearing in only one subsystem.

  $\boxed{\text{complicating}}$ " " " " " " more than one subsystem

  } describe by hyper graph

  - subsystems are nodes
  - complicating (variables, constraints, objective terms) are hyperedges

  # complicating stuff চেনার উপায়:
  এরা হলো যেসব stuff, যাদেরকে fix করলে all the subsystems become seperable.

Page 2 Conditional Seperability:

- Seperable problem:

$$\begin{pmatrix} \min\ f_1(x_1)+f_2(x_2) \\ \text{s.t. } x_1\in C_1,\ x_2\in C_2 \end{pmatrix} = \min_x \left(f_1(x_1)+I_{C_1}(x_1)\right)+\left(f_2(x_2)+I_{C_2}(x_2)\right) \quad \text{\#Seperable problem}$$

$$= \left(\min_{x_1} f_1(x_1)+I_{C_1}(x_1)\right)+\left(\min_{x_2} f_2(x_2)+I_{C_2}(x_2)\right)$$

$$= \begin{pmatrix} \min_{x_1} f_1(x_1) \\ \text{s.t. } x_1\in C_1 \end{pmatrix} + \begin{pmatrix} \min_{x_2} f_2(x_2) \\ \text{s.t. } x_2\in C_2 \end{pmatrix}$$

- Conditionally seperable:

  Two subsystems: conditionally seperable $\leftrightarrow$ seperable when the complicating variables are fixed

  $\leftrightarrow$ any two subsystems are not connected by a net.

✱Example:

$$\min_{z_1,z_2,x} f_1(z_1,x)+f_2(z_2,x)$$

x: complicating variable, public or interface or boundary variable between the two subproblems

z_1, z_2: local variables

- hypergraph: two nodes connected by an edge

Page 3 Transformation to standard form:

1) $\boxed{\text{Complicating}}$

- complicating inequality constraints $\rightarrow$ (introduce slack variables)
- complicating variables $\rightarrow$ (introduce local copies, these local copies are made public)

2) minimize over private variables (preserves convexity) with the local constraints as indicator function in domain of local objective term.

3) Now we are left with ✱ all variables are public associated with a single node

✱ all constraints are $\underbrace{\text{consistency constraints}}$ (equality with two or more variables)

Example:

$$\begin{pmatrix} \min\limits_{z_1,z_2,x} f_1(z_1,x)+f_2(z_2,x) \end{pmatrix} = \begin{pmatrix} \min\limits_{z_1,z_2,x_1,x_2} f_1(z_1,x_1)+f_2(z_2,x_2) \\ \text{s.t. } x_1=x_2 \end{pmatrix}$$

complicating variables // introduce local copies

// eliminate local variables // by minimizing over // private variables z_1, z_2

$$= \begin{pmatrix} \min\limits_{x_1,x_2} \left(\overbrace{\min_{z_1} f_1(z_1,x_1)}^{\tilde f_1(x_1)} + \overbrace{\min_{z_2} f_2(z_2,x_2)}^{\tilde f_2(x_2)} \right) \\ \text{s.t. } x_1=x_2 \end{pmatrix} = \begin{pmatrix} \min\limits_{x_1,x_2} \tilde f_1(x_1)+\tilde f_2(x_2) \\ \text{s.t. } x_1=x_2 \end{pmatrix}$$

(#Using [Separation principle for coupling variables])

Public variable

consistency constraint

$\downarrow$ in hypergraph general form

$x_1 = Iz$
$x_2 = Iz$ } # z is correspond to a net # with common variable values $z=(z_1,..,z_m)$

Page 4: General form:

each of them are vectors

- n subsystems with $x_1, \dots, x_n$
- m nets with common variable values $z_1, \dots, z_m$
- Problem structure

$$\min \sum_{i=1}^{n} f_i(x_i)$$

these matrices give the netlist

• Problem structure

$$\boxed{\begin{aligned}&\min \sum_{i=1}^{n} f_i(x_i)\\ &\text{s.t.}\\ &\forall_{i\in\{1,\dots,n\}} \quad x_i = E_i z\end{aligned}}$$

$\to$ y_i : associated dual variable

these matrices give the netlist

$\underbrace{E_i}_{\mathbb{R}^{n\times m}}\underbrace{z}_{\mathbb{R}^{m}} = \begin{bmatrix} \bar{e}_1^T z \\ \vdots \\ \bar{e}_n^T z \end{bmatrix}$ where $\bar{e}_j = e_j = (0,\dots,0,\overset{j\text{th place}}{1},\dots,0)$, if x_{ij} is in net p

see the circuit example below:

✱ Optimality conditions (KKT)

$$\left(\begin{aligned}&\min \sum_{i=1}^{n} f_i(x_i)\\ &\text{s.t.}\\ &\forall_{i\in\{1,\dots,n\}} \quad x_i = E_i z\end{aligned}\right)$$: primal problem

$\leftrightarrow$ y_i # dual variable

$$L(x_1,\dots,x_n,z,y_1,\dots,y_n) = \sum_{i=1}^{n} f_i(x_i) + \sum_{i=1}^{n} y_i^T(x_i - E_i z)$$

$$= \sum_{i=1}^{n}\left(f_i(x_i) - y_i^T x_i\right) - \sum_{i=1}^{n} \underbrace{y_i^T(E_i z)}_{(E_i^T y_i)^T z}$$

KKT condition:

Vanishing gradient of the Lagrangian:

$$\nabla_{(x_1,\dots,x_n,z)} L(x_1,\dots,x_n,z,y_1,\dots,y_n) = \begin{bmatrix}\nabla_{x_1} L\\ \nabla_{x_2} L\\ \vdots\\ \nabla_{x_n} L\\ \nabla_z L\end{bmatrix} = \begin{bmatrix}0\\0\\ \vdots\\0\\0\end{bmatrix} \Leftrightarrow \forall_{i\in\{1,\dots,n\}} \underbrace{\nabla_{x_i} L}_{\nabla_{x_i} f_i(x_i) - y_i} = 0 \Leftrightarrow \boxed{\forall_{i\in\{1,\dots,n\}} \nabla_{x_i} f_i(x_i) = y_i}$$

১ম equation shows the relation that is true
for every subsystem

$$\underbrace{\nabla_z L}_{} = 0 \Leftrightarrow \boxed{\sum_{i=1}^{n} E_i^T y = 0}$$

shows the condition of the dual variables
dual variables on each net sum to zero

$$\nabla_z\left(\sum_{i=1}^{n}(E_i^T y_i)^T z\right)$$

$$= \sum_{i=1}^{n} \nabla_z\underbrace{\left[(E_i^T y_i)^T z\right]}_{(E_i^T y)}$$

∇_z linear operator বলে
sum এর মাঝে ঢুকে যেতে পারবে

$$= \sum_{i=1}^{n}(E_i^T y)$$

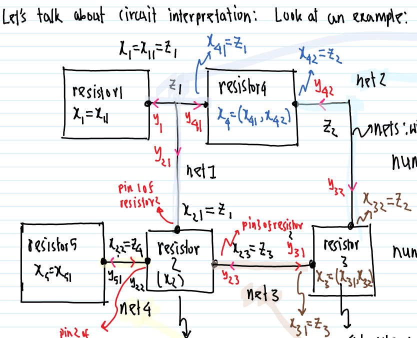
Primal feasibility:

$$\boxed{x_i = E_i z}$$

Primal variables on each net
are the same

✱✱✱

✱ Let's talk about circuit interpretation: Look at an example:



nets: wires

number of wires = $\dim(z) = m$

z_j = the voltage on net j ; in the figure we have 4 nets, so $z = (z_1, z_2, z_3, z_4)$

number of subsystems = number of nonlinear resistors

x_i = vector of pin voltages for resistor i

Subsystem: a subcircuit containing a bunch of nonlinear resistors

e.g.

resistor 2 has 3 pins

the voltages on these 3 pins are

x_{21}, x_{22}, x_{23}, i.e. $x_2 = (x_{21}, x_{22}, x_{23})$

As evident from the figure, $\left.\begin{aligned}x_{21} &= z_1\\ x_{22} &= z_4\\ x_{23} &= z_3\end{aligned}\right\}$ $x_2 = \begin{bmatrix}x_{21}\\x_{22}\\x_{23}\end{bmatrix} = \begin{bmatrix}1&0&0&0\\0&0&0&1\\0&0&1&0\end{bmatrix}\begin{bmatrix}z_1\\z_2\\z_3\\z_4\end{bmatrix} = \underbrace{\begin{bmatrix}e_1^T\\e_4^T\\e_3^T\end{bmatrix}}_{E_2} z$

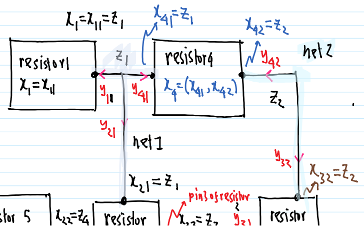
as pin 1 voltage of resistor 2 (x_{21}) is in net 1, the first row of E_2 is e_1^T
" " 2 " " " 2 (x_{22}) " " " 4 " 2nd " " E_2 is e_4^T
" " 3 " " " 2 (x_{23}) " " " 3 " 3rd " " E_2 is e_3^T

Page 4: ✱ interpretation of the dual variables y: vector of current entering resistor i

so the current vector going in the pins of resistor 2, $y_2 = \begin{bmatrix}y_{21}\\y_{22}\\y_{33}\end{bmatrix}$

✱ What would be KCL then?

sum of currents leaving net j is zero infact this is the optimality condition: $\sum_{i=1}^{n} E_i^T y_i = 0$



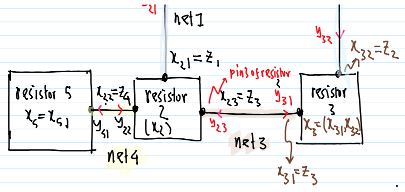
net 1: $y_{11} + y_{41} + y_{21} = 0$

net 2: $y_{42} + y_{32} = 0$

net 3: $y_{23} + y_{31} = 0$

net 4: $y + y = 0$

$\left.\right\}$ $\sum_{i=1}^{n} E_i^T y_i = 0$



net 3: $y_{23}+y_{31}=0$

net 4: $y_{51}+y_{22}=0$

* V-I charecteristic of the resistor

$$y_i = \nabla f_i(x_i)$$
$$\downarrow$$
$$\begin{bmatrix} y_{i1} \\ \vdots \\ y_{in_i} \end{bmatrix} = \begin{bmatrix} \frac{\partial}{\partial x_{i1}} f_i(x_i) \\ \vdots \\ \frac{\partial}{\partial x_{in_i}} f_i(x_i) \end{bmatrix}$$

$f_i(x_i)$: content function of the resistor i
↓
cost function of the ith resistor

* So we have in the circuit interpretation:

x_i = vector of voltages on the pins of nonlinear resistor i $\big|_{i=1}^{n}$

$z=(z_1,\dots,z_m)$ = vector of voltages on the wires (nets) of the circuit

y_i = vector of currents going in different pins of the nonlinear resistor i

The circuit equations are:

- KVL: $\forall_{i\in\{1,\dots,n\}}$ $x_i = E_i z$ ⟷ primal feasibility originally
- KCL: $\sum_{i=1}^{n} E_i^T y_i = 0$ ⟷ optimality condition on the dual variables
- V-I charecteristic of ith resistor: $\forall_{i\in\{1,\dots,n\}}$ $y_i = \nabla_{x_i} f_i(x)$ ⟷ optimality condition for the subsystems

So, circuit equations are the same as the optimality conditions !!!

(*) Convexity of f_i is the incremental passivity of resistor i

$(x_i-\tilde{x}_i)^T(y_i-\tilde{y}_i)\geq 0$, $y_i=\nabla f_i(x_i)$, $\tilde{y}_i=\nabla f_i(x_i)$ # explanation needed later...

* Decomposition methods:

- Solve distributed problems iteratively
- algorithm state maintained in nets

- Each step:
  - parallel update of subsystem primal and dual variables, based on adjacent net states (local blocks optimization)
  - update of the net states, based on adjacent subsystems (central block update)

- Algorithms differ in 1) interface to subsystems
  2) state and update

Page 6: *Primal Decomposition:

repeat:

$\boxed{1}$ Distribute net variables to adjacent subsystems

$x_i := E_i z$ # primal feasibility is always maintained
dual optimal condition is approached in limit.

$\boxed{2}$ Optimize Subsystems seperately

solve subsystems to evaluate $y_i = \nabla f_i(x)$ # Each of these subsystems are voltage controlled
voltage x_i vector is asserted to different pins of resistor i
current through the pins are then determined

$\boxed{3}$ Collect and sum dual variables for each net.

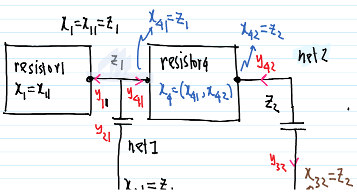
$w := \sum_{i=1}^{n} E_i^T y_i$ # this w corresponds to dual residual at
optimal solution w is 0

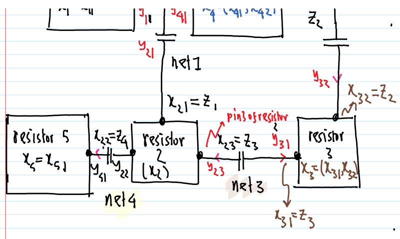
$\boxed{4}$ update net variables

$z := z - \alpha_k w$ # State of the algorithm is the netvoltage vector z
↓
chosen by standard gradient or subgradient rules

*Circuit interpretation:

- Connect capacitor to each net, system relaxes to equilibrium
- forward Euler update is primal decomposition
- incremental passivity implies convergence to equilibrium





*needs more clarifications