Revisiting decomposition algorithms

4:11 PM

The optimization process is executed on several computational units

lpu: local processing units

Large problem $\xrightarrow{\text{split}} \cup_{i=1}^{n}$ smaller problem$_i$ → 1: lpu1, i: lpu2, n: lpu3 → Master Unit

Some communication is needed between the LPUs by
- Peer-to-peer basis
- Central coordination unit } Master unit

Simplest decentralized structure := Seperable Problem

$$p^* = \begin{pmatrix} \min_{x_1,\dots,x_v} & \sum_{i=1}^{v} f_{0,i}(x_i) \\ \text{s.t.} & \forall_{i\in\{1,\dots,v\}}\ x_i \in X_i \end{pmatrix} = \min_{x_1,\dots,x_v}\left(\sum_{i=1}^{n} f_{0,i}(x_i) + \sum_{i=1}^{n} I_{X_i}(x_i)\right) = \min_{x_1,\dots,x_v}\left(\sum_{i=1}^{n}\left(f_{0,i}(x_i)+I_{X_i}(x_i)\right)\right) = \sum_{i=1}^{n}\left(\min_{x_i}\left(f_{0,i}(x_i)+I_{X_i}(x_i)\right)\right) = \sum_{i=1}^{n}\underbrace{\begin{pmatrix}\min f_{0,i}(x_i)\\ \text{s.t. } x_i\in X_i\end{pmatrix}}_{= f_{0,i}^*} = \sum_{i=1}^{n} f_{0,i}^*$$

vector blocks $1,\dots,v = n$, $x \in \mathbb{R}^n$

More interesting case:

- Coupled in the objective

$$p^* = \begin{pmatrix} \min_{x_1,\dots,x_v} & f_0(x_1,\dots,x_v) \\ \text{s.t.} & \forall_{i\in\{1,\dots,v\}}\ x_i\in X_i \end{pmatrix}$$

→ Objective is not sum-seperable

→ constraint set seperable

- Coupled in the constraints:

$$p^* = \begin{pmatrix} \min_{x_1,\dots,x_v} & \sum_{i=1}^{n} f_{0,i}(x_i) \\ \text{s.t.} & x\in X \end{pmatrix}$$

→ sum seperable

→ coupled

Decomposition methods:

→ Primal Decomposition: Resource (coupled constraints or imposed limitation) is managed by master unit directly, by assigning individual budgets to the LPUs

→ Dual Decomposition: Master unit manages resources indirectly by assigning resource prices to the subproblems

Dual Decomposition: Problem structure

$$p^* = \begin{pmatrix} \min_{x_1,\dots,x_v} & \sum_{i=1}^{v} f_{0,i}(x_i) \\ \text{s.t.} & \forall_{i\in\{1,\dots,v\}}\ x_i\in X_i \\ & \sum_{i=1}^{v} h_i(x_i)\le c \end{pmatrix}$$

if we look at only this # part the problem is decoupled

$\{$budget limit constraint (coupling one)$\}$

resource consumption of ith block

this is a resource vector constraint, i.e. h_i is a vector function $h_i(x) = \begin{bmatrix} h_{i1}(x)\\ \vdots \\ h_{im}(x)\end{bmatrix}$

e.g. $x_1+x_2+x_3\le 3$, $x_1^2+3x_2+5x_3\le 7$ } in vector form

$$\underbrace{\begin{bmatrix} x_1\\ x_1^2\end{bmatrix}}_{h_1(x_1)}+\underbrace{\begin{bmatrix} x_2\\ 3x_2\end{bmatrix}}_{h_2(x_2)}+\underbrace{\begin{bmatrix} x_3\\ 5x_3\end{bmatrix}}_{h_3(x_3)}\le\begin{bmatrix}3\\7\end{bmatrix}$$

$$= \begin{pmatrix} \min_{x_1,\dots,x_v} & \sum_{i=1}^{v} f_{0,i}(x_i) + \sum_{i=1}^{v} I_{X_i}(x_i) \\ \text{s.t.} & \sum_{i=1}^{v} h_i(x_i)\le c\end{pmatrix} = \begin{pmatrix} \min_{x_1,\dots,x_v} & \sum_{i=1}^{v}\left(f_{0,i}(x_i)+I_{X_i}(x_i)\right) \\ \text{s.t.} & \sum_{i=1}^{v} h_i(x_i)\le c\end{pmatrix}$$

$$L(x,\lambda) = \sum_{i=1}^{v}\left(f_{0,i}(x_i)+I_{X_i}(x_i)\right) + \lambda^T\left(\left(\sum_{i=1}^{v} h_i(x_i)\right)-c\right)\ \#[\lambda\ge 0]\#$$

$$= \sum_{i=1}^{v}\left(f_{0,i}(x_i)+\lambda^T h_i(x_i)+I_{X_i}(x_i)\right) - \lambda^T c$$

$$g(\lambda) = \inf_x L(x,\lambda) = \inf_x\left[\left(\sum_{i=1}^{v}\left(f_{0,i}(x_i)+\lambda^T h_i(x_i)+I_{X_i}(x_i)\right)\right)-\lambda^T c\right] = -\lambda^T c + \inf_x\left(\sum_{i=1}^{v}\left(f_{0,i}(x_i)+\lambda^T h_i(x_i)+I_{X_i}(x_i)\right)\right)$$

↓ seperable in x_i !

$$= -\lambda^T c + \sum_{i=1}^{v}\inf_{x_i}\left(f_{0,i}(x_i)+\lambda^T h_i(x_i)+I_{X_i}(x_i)\right)$$

$$= -\lambda^T c + \sum_{i=1}^{v}\underbrace{\inf_{x_i\in X_i}\left(f_{0,i}(x_i)+\lambda^T h_i(x_i)\right)}_{g_i(\lambda)}\quad \left\{\because \min_x\left(f(x)+I_X(x)\right) = \min_{x\in X}\left(f(x)\right)\right\}$$

$\#\left[g_i(\lambda) = \inf_{x_i\in X_i}\left(f_{0,i}(x_i)+\lambda^T h_i(x_i)\right)\right]\#$

$$= -\lambda^T c + \sum_{i=1}^{v}\underbrace{\left[f_{0,i}(x_i^*(\lambda))+\lambda^T h_i(x_i^*(\lambda))\right]}\quad \#\left[x_i^*(\lambda) = \operatorname*{argmin}_{x_i\in X_i}\left(f_{0,i}(x_i)+\lambda^T h_i(x_i)\right)\right]\# = -\lambda^T c + \sum_{i=1}^{v} g_i(\lambda)$$

these $x_i^*(\lambda)$'s are calculated by ith LPU

This ith LPU is solving the problem $g_i(\lambda) = \inf_{x_i\in X}\left(f_{0,i}(x_i)+\lambda^T h_i(x_i)\right)$, with $g_i(\lambda)$, and $x_i^*(\lambda)$ as returned value.

- Master unit solves the dual problem:

$$d^* = \max_{\lambda\ge 0} g(\lambda) = \max_{\lambda\ge 0}\left(-\lambda^T c+\sum_{i=1}^{v} g_i(\lambda)\right)$$

$$= -\min_{\lambda\ge 0} -\left(-\lambda^T c+\sum_{i=1}^{v} g_i(\lambda)\right)$$

$$= -\underbrace{\min_{\lambda\ge 0}\left(\lambda^T c-\sum_{i=1}^{v} g_i(\lambda)\right)}$$

To solve this we will use projected subgradient method described as:

$$\begin{pmatrix}\min & f(x)\\ \text{s.t.} & x\in X\end{pmatrix} \to x_{k+1} = \left[x_k - s_k g_k\right]_X$$

s_k → suitable step size; g_k → gradient of $f(\cdot)$ at x_k

$$\therefore \begin{pmatrix}\min & \lambda^T c-\sum_{i=1}^{v} g_i(\lambda)\\ \text{s.t.} & \lambda\ge 0\end{pmatrix} \to \lambda_{k+1} = \left[\lambda_k - s_k g_k\right]_{\underset{\mathbb{R}_+^m}{\parallel}X} = \left[\lambda_k - s_k g_k\right]_+$$

$$g_k \in \partial f(\lambda_k) = \partial_{\lambda_k}\left[c^T\lambda_k + \sum_{i=1}^{v} -g_i(\lambda_k)\right]$$

($f(\lambda_k) = \lambda_k^T c - \sum_{i=1}^{v} g_i(\lambda_k)$)

$$= \partial_{\lambda_k}(c^T\lambda_k) + \partial_{\lambda_k}\left(\sum_{i=1}^{v} -g_i(\lambda_k)\right)\ \#\ \partial f_1(x)+\partial f_2(x) = \partial(f_1(x)+f_2(x))\ \#$$

$$= \nabla_{\lambda_k}(c^T\lambda_k) + \sum_{i=1}^{v}\partial_{\lambda_k}(-g_i(\lambda_k))\ \#\ \ "\ \ \#$$

$$= c + \sum_{i=1}^{v}\partial_{\lambda_k}(-g_i(\lambda_k))$$

$\#$ now: $g_i(\lambda) = \inf_{x_i\in X_i}\left(f_{0,i}(x_i)+\lambda^T h_i(x_i)\right) = \inf_{x_i}\left(\underbrace{f_{0,i}(x_i)+I_{X_i}(x_i)+\lambda^T h_i(x_i)}\right)$

this is the Lagrangian function of the primal problem: $\begin{pmatrix}\min & f_{0,i}(x_i)+I_{X_i}(x_i)\\ \text{s.t.} & f_i(x_i)\le 0\end{pmatrix}$ where $f_i(x)\le 0 \Leftrightarrow \begin{bmatrix} f_{i1}(x_i)\\ \vdots\\ f_{im}(x_i)\end{bmatrix}\le\begin{bmatrix}0\\ \vdots\\ 0\end{bmatrix}$

So, by subgradient of dual function rule (4.1.4):

$$-h_i(x_i^*(\lambda)) = \left(-h_{i1}(x_i^*(\lambda)),\dots,-h_{im}(x_i^*(\lambda))\right)\in\partial(-g(\lambda))$$

argmin$(\cdot)$ $\#$

$$g_k = c + \sum_{i=1}^{v}\underbrace{\left(-h_i(x_i^*(\lambda_k))\right)}$$

$\#\left[\because -h_i(x^*(\lambda_k))\in\partial_{\lambda_k}(-g_i(\lambda_k))\right]\#$

$$= c - \sum_{i=1}^{v} h_i(x_i^*(\lambda_k))$$

$$\lambda_{k+1} = \left[\lambda_k - s_k\left(c-\sum_{i=1}^{v} h_i(x_i^*(\lambda_k))\right)\right]_+$$

Master problem is essentially
updating the price

s_k → suitably choosen stepsize

To Paris
↑
22 March (Airfrance 361)

Conceptual scheme of the decomposition problem:

central processes: $\boxed{\lambda_{k+1} = \left[\lambda_k - s_k\left(c-\sum_{i=1}^{v} h_i(x_i^*(\lambda_k))\right)\right]_+\ \#\text{price update}\#;\ k:=k+1 \quad \text{which is solving } \left(\min -g(\lambda)\ \text{s.t. } \lambda\ge 0\right) = -d^* \text{ as } k\to\infty}$

$\lambda_k \downarrow\ \uparrow x_1^*(\lambda_k), h_1(x_1^*(\lambda_k))$ $\quad$ $\lambda_k \downarrow\ \uparrow x_v^*(\lambda_k), h_v(x_v^*(\lambda_k))$

local processes: $\boxed{\min_{x_1\in X_1} f_{0,1}(x_1)+\lambda_k^T h_1(x_1)}$ . . . $\boxed{\min_{x_v\in X_v} f_{0,v}(x_v)+\lambda_k^T h_v(x_v)}$

* (if strong duality holds $\wedge$ $\forall_\lambda\ x_i^*(\lambda)\ \{\,!\,\}$) $\Rightarrow$ dual decomposition provides primal optimal variables $\forall_{i\in\{1,\dots,v\}}\ x_i^*(\lambda)\to x_i^*$

$\updownarrow$

$f_{0,i}(x_i)+\lambda^T h_i(x_i)\ \{\square_{\text{strongly}}\}$

(*) Dual decomposition with coupling variables:

In the problem structure of the previous topic, the coupling was in the constraints (some), objective was seperable

Now consider variable coupling in the objective:

$$p^* = \begin{pmatrix} \min\limits_{x_1,x_2,y} & f_{0,1}(x_1,y) + f_{0,2}(x_2,y) \\ \text{s.t.} & x_1 \in \mathcal{X}_1,\ x_2 \in \mathcal{X}_2,\ y \in \mathcal{Y} \end{pmatrix} = \begin{pmatrix} \min\limits_{x_1,x_2,y} & f_{0,1}(x_1,y) + f_{0,2}(x_2,y) + I_{\mathcal{X}_1}(x) + I_{\mathcal{X}_2}(x) + I_{\mathcal{Y}}(y) \end{pmatrix}$$

y যদি কোথাও না থাকত the problem would have been seperable.

Standard approach: • Introduce slack variables y_1, y_2

• " artificial equality constraint $y_1 = y_2$

$$p^* = \begin{pmatrix} \min & f_{0,1}(x_1,y_1) + f_{0,2}(x_2,y_2) + I_{\mathcal{X}_1}(x_1) + I_{\mathcal{X}_2}(x_2) + I_{\mathcal{Y}}(y_1) + I_{\mathcal{Y}}(y_2) \\ \text{s.t.} & y_1 = y_2 \end{pmatrix}$$

$$L(x_1,x_2,y_1,y_2,\nu) = f_{0,1}(x_1,y_1) + I_{\mathcal{X}_1}(x_1) + f_{0,2}(x_2,y_2) + I_{\mathcal{X}_2}(x_2) + \nu^T(y_1 - y_2) + I_{\mathcal{Y}}(y_1) + I_{\mathcal{Y}}(y_2)$$

$$= \left(f_{0,1}(x_1,y_1) + I_{\mathcal{X}_1}(x_1) + \nu^T y_1 + I_{\mathcal{Y}}(y_1)\right) + \left(f_{0,2}(x_2,y_2) + I_{\mathcal{X}_2}(x_2) - \nu^T y_2 + I_{\mathcal{Y}}(y_2)\right)$$

// now these are seperable
// in (x_1, y_1) and (x_2, y_2)

$$g(\nu) = \inf_{x_1,x_2,y_1,y_2} \left[\left(f_{0,1}(x_1,y_1) + I_{\mathcal{X}_1}(x_1) + \nu^T y_1 + I_{\mathcal{Y}}(y_1)\right) + \left(f_{0,2}(x_2,y_2) + I_{\mathcal{X}_2}(x_2) - \nu^T y_2 + I_{\mathcal{Y}}(y_2)\right)\right]$$

$$= \inf_{x_1,y_1} \left(f_{0,1}(x_1,y_1) + I_{\mathcal{X}_1}(x_1) + \nu^T y_1 + I_{\mathcal{Y}}(y_1)\right) + \inf_{x_2,y_2} \left(f_{0,2}(x_2,y_2) + I_{\mathcal{X}_2}(x_2) - \nu^T y_2 + I_{\mathcal{Y}}(y_2)\right)$$

$$= \underbrace{\inf_{\substack{x_1 \in \mathcal{X}_1 \\ y_1 \in \mathcal{Y}}} \left(f_{0,1}(x_1,y_1) + \nu^T y_1\right)}_{g_1(\nu)} + \underbrace{\inf_{\substack{x_2 \in \mathcal{X}_2 \\ y_2 \in \mathcal{Y}}} \left(f_{0,2}(x_2,y_2) - \nu^T y_2\right)}_{g_2(\nu)}$$

Master Process:

$$\boxed{\begin{array}{l} \max g(\nu) = -\min(-g(\nu)) \\ \nu_{k+1} = \nu_k - s_k g_k \text{ where } g_k = y_1^*(\nu_k) - y_2^*(\nu_k) \end{array}}$$

ν ↓ ↑ $x_1^*(\nu)$, $y_1^*(\nu)$ $\qquad$ $x_2^*(\nu)$, $y_2^*(\nu)$ ↑ ↓ ν

Local Process:

$$\boxed{\begin{array}{ll} \min & f_{0,1}(x_1,y_1) + \nu^T y_1 \\ \text{s.t.} & x_1 \in \mathcal{X}_1 \\ & y_1 \in \mathcal{Y} \end{array}} \qquad \boxed{\begin{array}{ll} \min & f_{0,2}(x_2,y_2) - \nu^T y_2 \\ \text{s.t.} & x_2 \in \mathcal{X}_2 \\ & y_2 \in \mathcal{Y} \end{array}}$$

as the problem is unconstrained, subgradient method will work, s_k is any suitable stepsize, and $g_k \in \partial(-g(\nu)) = \partial(-(g_1(\nu) + g_2(\nu))) = \partial(-g_1(\nu)) + \partial(-g_2(\nu))$ [$\because \partial(f_1(x) + f_2(x)) = \partial f_1(x) + \partial f_2(x)$] (set addition)

It can be shown that, $g_k = y_1^*(\nu_k) - y_2^*(\nu_k) \in \partial(-g(\nu_k))$

Proof: Recall: $\left(-f_i(x)\right)_{i=1}^m \in \partial(-g(\lambda))$, where g is the dual function of

$$\begin{pmatrix} \min & f_0(x) \\ \text{s.t.} & \\ \forall_{i \in \{1,\dots,m\}} & f_i(x) \le 0 \\ & x \in \mathcal{D} \end{pmatrix}$$

now, $g_1(\nu) = \inf\limits_{\substack{x_1 \in \mathcal{X}_1 \\ y_1 \in \mathcal{Y}}} \left(f_{0,1}(x_1,y_1) + \nu^T y_1\right)$ which can be backcalculated as the dual function of

$$\begin{pmatrix} \min & f_{0,1}(x_1,y_1) \\ \text{s.t.} & y_1 \le 0 \end{pmatrix} \quad \therefore y_1 \in \partial(-g_1(\nu))$$

$g_2(\nu) = \inf\limits_{\substack{x_2 \in \mathcal{X}_2 \\ y_2 \in \mathcal{Y}}} \left(f_{0,2}(x_2,y_2) - \nu^T y_2\right)$ is the dual function of

$$\begin{pmatrix} \min & f_{0,2}(x_2,y_2) \\ \text{s.t.} & -y_2 \le 0 \end{pmatrix} \quad \therefore -y_2 \in \partial(-g_2(\nu))$$

$$\therefore y_1 - y_2 \in \partial(-g(\nu))$$

• Primal Decomposition:

Problem:

$$p^* = \begin{pmatrix} \min & \sum_{i=1}^{\nu} f_{0,i}(x_i) \\ \text{s.t.} & \forall_{i \in \{1,\dots,\nu\}}\ x_i \in \mathcal{X}_i \\ & \sum_{i=1}^{\nu} h_i(x_i) \preceq c \end{pmatrix} = \begin{pmatrix} \min & \sum_{i=1}^{\nu} f_{0,i}(x_i) \\ \text{s.t.} & \forall_{i \in \{1,\dots,\nu\}}\ x_i \in \mathcal{X}_i \\ & \sum_{i=1}^{\nu} t_i \preceq c \\ & \forall_{i \in \{1,\dots,\nu\}}\ h_i(x_i) \preceq t_i \end{pmatrix}$$

remember each of the t_i blocks are
vectors with same dimension as of length(c)

$$= \min_{(x,t)} \left(\sum_{i=1}^{\nu} f_{0,i}(x_i) + \sum_{i=1}^{\nu} I_{\mathcal{X}_i}(x_i) + \sum_{i=1}^{\nu} I_{(h_i(\square_i) \preceq \square_i)}(x,t) + I_{(\sum_{i=1}^{\nu} \square_i \preceq c)}(t)\right)$$

$$= \min_{(x,t)} \left[\left(\sum_{i=1}^{\nu} \left(f_{0,i}(x_i) + I_{\mathcal{X}_i}(x_i) + I_{(h_i(\square_i) \preceq \square_i)}(x,t)\right)\right) + I_{(\sum_{i=1}^{\nu} \square_i \preceq c)}(t)\right]$$

$$= \min_t \min_x \left[\left(\sum_{i=1}^{\nu} \left(f_{0,i}(x_i) + I_{\mathcal{X}_i}(x_i) + I_{(h_i(\square_i) \preceq \square_i)}(x,t)\right)\right) + I_{(\sum_{i=1}^{\nu} \square_i \preceq c)}(t)\right]$$

at this the problem
is written in single objective
form, so x ও আলাদা minimization
হিসেবে consider করা

$$= \min_t \left(\min_x \left[\left(\sum_{i=1}^{\nu} \left(f_{0,i}(x_i) + I_{\mathcal{X}_i}(x_i) + I_{(h_i(\square_i) \preceq \square_i)}(x,t)\right)\right) + \underbrace{I_{(\sum_{i=1}^{\nu} \square_i \preceq c)}(t)}_{\text{constant w.r.t } x \text{ so } \min_x \text{ এর বাইরে আনা যাবে}}\right]\right)$$

$$= \min_t \left(\min_x \left[\left(\sum_{i=1}^{\nu} \left(f_{0,i}(x_i) + I_{\mathcal{X}_i}(x_i) + I_{(h_i(\square_i) \preceq \square_i)}(x,t)\right)\right)\right] + I_{(\sum_{i=1}^{\nu} \square_i \preceq c)}(t)\right)$$

$$= \min_t \left(\min_x \left[\left(\sum_{i=1}^{\nu} \left(f_{0,i}(x_i) + I_{\mathcal{X}_i}(x_i) + I_{(h_i(\square_i) \preceq \square_i)}(x,t)\right)\right)\right]\right)$$

$$\text{s.t.} \quad \sum_{i=1}^{\nu} t_i \preceq c$$

Fix the value of $\tilde{t} = (\tilde{t}_1, \dots, \tilde{t}_\nu)$ such that $\sum \tilde{t}_i \preceq c$ then

$\forall_{\tilde{t} : \sum_{i=1}^{\nu} \tilde{t}_i \preceq c}$

$$p^*(\tilde{t}) = \begin{pmatrix} \min & \sum_{i=1}^{\nu} f_{0,i}(x_i) \\ \text{s.t.} & \forall_{i \in \{1,\dots,\nu\}}\ x_i \in \mathcal{X}_i \\ & \forall_{i \in \{1,\dots,\nu\}}\ h_i(x_i) \preceq \tilde{t}_i \end{pmatrix} = \min\left(\sum_{i=1}^{\nu} f_{0,i}(x_i) + I_{\mathcal{X}_1}(x_1) + \dots + I_{\mathcal{X}_\nu}(x_\nu) + I_{\{x_1 : h_1(x_1) \preceq \tilde{t}_1\}}(x_1) + \dots + I_{\{x_\nu : h_\nu(x_\nu) \preceq \tilde{t}_\nu\}}(x_\nu)\right)$$

$$= \min\left(\sum_{i=1}^{\nu} f_{0,i}(x_i) + \sum_{i=1}^{\nu} I_{\mathcal{X}_i}(x_i) + \sum_{i=1}^{\nu} I_{\{x_i : h_i(x_i) \preceq \tilde{t}_i\}}(x_i)\right)$$

$$= \inf\left(\sum_{i=1}^{v} f_{0,i}(x_i) + \sum_{i=1}^{v} I_{\mathcal{X}_i}(x_i) + \sum_{i=1}^{v} I_{\{x_i : h_i(x_i) \le \tilde{t}_i\}}(x_i)\right)$$

$$= \inf\left(\sum_{i=1}^{v}\left(f_{0,i}(x_i) + I_{\mathcal{X}_i}(x_i) + I_{\{x_i : h_i(x_i)\le \tilde{t}_i\}}(x_i)\right)\right)$$

// each of the summands are seperable as they are

// not coupled

$$= \sum_{i=1}^{v}\left(\inf\left(f_{0,i}(x_i) + I_{\mathcal{X}_i}(x_i) + I_{\{x_i : h_i(x_i)\le \tilde{t}_i\}}(x_i)\right)\right)$$

$$= \sum_{i=1}^{v}\left(\inf\left(f_{0,i}(x_i) + I_{\mathcal{X}_i}(x_i) + I_{\{x_i : h_i(x_i)\le \tilde{t}_i\}}(x_i)\right)\right)$$

$$= \sum_{i=1}^{v} \underbrace{\begin{pmatrix} \inf & f_{0,i}(x_i) \\ \text{s.t.} & x_i \in \mathcal{X}_i \\ & h_i(x_i) \le \tilde{t}_i \end{pmatrix}}_{p_i^*(\tilde{t}_i)} = \sum_{i=1}^{v} p_i^*(\tilde{t}_i)$$

Each LPDs will solve them locally

now each $p_i^*(t_i)$ is convex in t_i

proof: $p_i^*(\tilde{t}_i) = \inf_{x_i \in \mathcal{X}_i} \Big(\underbrace{f_{0,i}(x_i) + I}_{\mathcal{D}_{x_i} = \mathcal{D}_{(x_i;t_i)}} \underbrace{{}_{\{(x,t) : h_i(x_i) - t_i \le 0\}}(x_i, t_i)}_{\mathcal{D}_{(x_i;t_i)}}\Big)$

$\mathcal{D}_{(x_i;t_i)}$

$\left(\inf_{x_i \in \mathcal{X}_i} \mathcal{D}_{(x_i;t_i)}\right) = \mathcal{D}_{t_i}$

the master process then solves:

$$\begin{pmatrix} \inf\limits_{t_1,\dots,t_v} & \sum_{i=1}^{v} p_i^*(t_i) \\ \text{s.t.} & \sum_{i=1}^{v} t_i \le c \end{pmatrix}$$

Now this can be solved using projected gradient method:

$$\underbrace{t_{k+1}}_{(t_{k+1,1},\dots,t_{k+1,v})} = \Big[t_k - \underbrace{s_k}_{\text{suitable step size}}\, g_{t_k}\Big]_{\{t : \sum_{i=1}^{v} t_i \le c\}}$$

subgradient $g_{t_k} \in \partial\left(\sum_{i=1}^{v} p_i^*(t_{k,i})\right) = \sum_{i=1}^{v} \partial p_i^*(t_{k,i})$

so, if $g_{t_{k,i}} \in \partial p_i^*(t_{k,i})$, then $g_{t_k} = \sum_{i=1}^{v} g_{t_{k,i}}$

Now to find the optimum using subgradient we need a subgradient of $p_i^*(t_i)$, g_{t_i}, such that:

$\forall \tilde{t}_i \quad p_i^*(\tilde{t}_i) \ge p_i^*(t_i) + g_{t_i}^T(\tilde{t}_i - t_i)$

Now, (eg: subgradient of optimal value function)

~~one~~ ~~such~~ : one such

$g_{t_i} = -\lambda_{t_i}$ wher λ_{t_i} is the optimal dual variable of the problem $p_i^*(t_i) = \begin{pmatrix} \inf & f_{0,i}(x_i) \\ \text{s.t.} & x_i \in \mathcal{X}_i \\ & h_i(x_i) \le t_i \to \lambda_{t_i} \end{pmatrix}$

so the decomposition scheme looks like:

master

$$\lambda_{k+1} = [\lambda_k - s_k g_k]_{\{t : \sum_{i=1}^{v} t_i \le c\}}$$

$$g_k = \sum_{i=1}^{v} -\lambda_{t_{k,i}} \quad \text{where } \lambda_{t_{k,i}} = \text{optimal\_dual\_variable}\begin{pmatrix} \inf & f_{0,i}(x_i) \\ \text{s.t.} & x_i \in \mathcal{X}_i \\ & h_i(x_i) \le \lambda_{t_{k,i}} \end{pmatrix}$$

As a result this will solve $\begin{pmatrix} \inf\limits_{t_1,\dots,t_v} & \sum_{i=1}^{v} p_i^*(t_i) \\ \text{s.t.} & \sum_{i=1}^{v} t_i \le c \end{pmatrix}$

t_1 ; $f_{0,1}(x_1^*(t_1)), \lambda_{t_1}$; t_v ; $f_{0,v}(x_v^*(t_v)), \lambda_{t_v}$

$$p_1^*(t_1) = \begin{pmatrix} \inf & f_{0,1}(x_1) \\ \text{s.t.} & x_1 \in \mathcal{X}_1 \\ & h_1(x_1) \le t_1 \end{pmatrix} \quad \cdots \quad p_v^*(t_v) = \begin{pmatrix} \inf & f_{0,v}(x_v) \\ \text{s.t.} & x_v \in \mathcal{X}_v \\ & h_v(x_v) \le t_v \end{pmatrix}$$

This is absolutely amazing! The subsystems donot even need to send the primal variable, only the dual variable, and the function value

* Primal decomposition with coupling variables:

$$p^* = \begin{pmatrix} \inf & f_{0,1}(x_1, y) + f_{0,2}(x_2, y) \\ \text{s.t.} & x_1 \in \mathcal{X}_1,\ x_2 \in \mathcal{X}_2,\ y \in \mathcal{Y} \end{pmatrix} \quad // \text{ say } (x_1^*, x_2^*, y^*) \text{ be the optimal solution}$$

$$= \begin{pmatrix} \inf\limits_{x_1,x_2,y} & \left(f_{0,1}(x_1,y) + I_{\mathcal{X}_1}(x_1)\right) + \left(f_{0,2}(x_2,y) + I_{\mathcal{X}_2}(x_2)\right) \\ \text{s.t.} & y \in \mathcal{Y} \end{pmatrix};$$

First we show: $\begin{pmatrix} \inf & f_{0,1}(x_1,y) + f_{0,2}(x_2,y) \\ \text{s.t.} & x_1 \in \mathcal{X}_1,\ x_2 \in \mathcal{X}_2,\ y \in \mathcal{Y} \end{pmatrix} = \inf\limits_{y \in \mathcal{Y}} \left(\begin{pmatrix} \inf\limits_{x_1} & f_{0,1}(x_1,y) \\ \text{s.t.} & x_1 \in \mathcal{X}_1 \end{pmatrix} + \begin{pmatrix} \inf\limits_{x_2} & f_{0,2}(x_2,y) \\ \text{s.t.} & x_2 \in \mathcal{X}_2 \end{pmatrix}\right)$ [Separation principle for coupling variables]

By definition:

$$\forall_{x_1, x_2, y \in \mathcal{Y}} \quad f_{0,1}(x_1^*, y^*) + \underbrace{I_{\mathcal{X}_1}(x_1^*)}_{0} + f_{0,2}(x_2^*, y^*) + \underbrace{I_{\mathcal{X}_2}(x_2^*)}_{0} \le f_{0,1}(x_1, y) + I_{\mathcal{X}_1}(x_1) + f_{0,2}(x_2, y) + I_{\mathcal{X}_2}(x_2)$$

$$\to \forall_{y \in \mathcal{Y}} \quad f_{0,1}(x_1^*, y^*) + f_{0,2}(x_2^*, y^*) \le \min_{x_1}\left(f_{0,1}(x_1,y) + I_{\mathcal{X}_1}(x_1)\right) + \min_{x_2}\left(f_{0,2}(x_2,y) + I_{\mathcal{X}_2}(x_2)\right)$$

$$= \underbrace{\begin{pmatrix} \inf & f_{0,1}(x_1,y) \\ \text{s.t.} & x_1 \in \mathcal{X}_1 \end{pmatrix}}_{p_1^*(y)} + \underbrace{\begin{pmatrix} \inf & f_{0,2}(x_2,y) \\ \text{s.t.} & x_2 \in \mathcal{X}_2 \end{pmatrix}}_{p_2^*(y)}$$

$$\to f_{0,1}(x_1^*, y^*) + f_{0,2}(x_2^*, y^*) \le \min_{y \in \mathcal{Y}}\left(p_1^*(y) + p_2^*(y)\right)$$

Again,

$$\min_{y \in \mathcal{Y}} p_1^*(y) + p_2^*(y) \le p_1^*(y^*) + p_2^*(y^*) \quad \text{[By def]}$$

Again,

$$\min_{y \in \mathcal{Y}} P_1^*(y) + P_2^*(y) \leq P_1^*(y^*) + P_2^*(y^*) \quad \text{[By def]}$$

$$= \min_{x_1} \left(f_{0,1}(x_1, y^*) + I_{\mathcal{X}_1}(x_1) \right) + \min_{x_2} \left(f_{0,2}(x_2, y^*) + I_{\mathcal{X}_2}(x_2) \right)$$

$$= \min_{x_1 \in \mathcal{X}_1} \left(f_{0,1}(x_1, y^*) \right) + \min_{x_2 \in \mathcal{X}_2} \left(f_{0,2}(x_2, y^*) \right) \leq f_{0,1}(x_1^*, y^*) + f_{0,2}(x_2^*, y^*) \quad \text{[By defn]}$$

$$\therefore f_{0,1}(x_1^*, y_1^*) + f_{0,2}(x_2^*, y_2^*) = \min_{y \in \mathcal{Y}} \Big(\underbrace{P_1^*(y)}_{\begin{pmatrix} \text{minimize}_{x_1} & f_{0,1}(x_1, y) \\ \text{subject to} & x_1 \in \mathcal{X}_1 \end{pmatrix}} + \overbrace{P_2^*(y)}^{\begin{pmatrix} \text{minimize}_{x_2} & f_{0,2}(x_2, y) \\ \text{subject to} & x_2 \in \mathcal{X}_2 \end{pmatrix}} \Big) = \begin{pmatrix} \text{minimize}_{y} & P_1^*(y) + P_2^*(y) \\ \text{subject to} & y \in \mathcal{Y} \end{pmatrix}$$

// to solve, using projected subgradient: $y_{k+1} = [y_k - s_k g_k]_{\mathcal{Y}}$

suitable stepsize

$$g_k \in \partial \left(P_1^*(y) + P_2^*(y) \right) = \partial P_1^*(y) + \partial P_2^*(y)$$

$$\begin{pmatrix} \text{minimize}_{y} & P_1^*(y) + P_2^*(y) \\ \text{subject to} & y \in \mathcal{Y} \end{pmatrix}$$

now $P_1^*(y) = \begin{pmatrix} \text{minimize}_{x_1} & f_{0,1}(x_1, y) \\ \text{subject to} & x_1 \in \mathcal{X}_1 \end{pmatrix}$, $P_2^*(y) = \begin{pmatrix} \text{minimize}_{x_2} & f_{0,2}(x_2, y) \\ \text{subject to} & x_2 \in \mathcal{X}_2 \end{pmatrix}$

for a particular problem $\mathcal{X}_1, \mathcal{X}_2$ এর structure জানা থাকবে.
so আমরা $P_1^*(y)$, $P_2^*(y)$ is optimal value function, so

(eq: subgradient of optimal value function)

থেকে subgradient
$g_{k,1} \in \partial P_1^*(y)$, $g_{k,2} \in \partial P_2^*(y)$ বের করা যাবে

$\therefore g_k = g_{k,1} + g_{k,2}$

The decomposition scheme is as follows:

master problem

sub problem