· Primal decomposition

$$\min \; f_1(x_1, y) + f_2(x_2, y)$$

fix y and define

subproblem 1: $\underset{x_1}{\text{minimize}} \; f_1(x_1, y)$

subproblem 2: $\underset{x_2}{\text{minimize}} \; f_2(x_2, y)$

and optimal values $\phi_1(y)$ and $\phi_2(y)$

The original problem is equivalent to master problem

$$\min_y \; \phi_1(y) + \phi_2(y)$$

if f is convex in x and y, C is a convex set

$g(x) = \inf_{y \in C} f(x, y)$ is convex

if the original problem is convex, so is the master problem

can solve the master problem using:

- bisection (if y is scalar)
- gradient method (if ϕ is differentiable)
- subgradient, cutting-plane or ellipsoid method

each iteration of master problem requires solving the two subproblems

· Primal decomposition algorithm (using subgradient method)

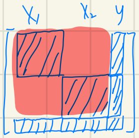
repeat

1. find x_1 that minimizes $f_1(x_1, y)$ and a subgradient $g_1 \in \partial \phi_1(y)$
   find x_2 that minimizes $f_2(x_2, y)$ and a subgradient $g_2 \in \partial \phi_2(y)$
2. update the complicating variable
   $y := y - \alpha_k(g_1 + g_2)$ $\quad \alpha_k$ can be chosen in any of the standard ways

$f_1(x_1, y) + f_2(x_2, y)$

smart linear algebra

the hessian looks like



arrow-shape

the block-diagonal is easily-invertible

if we have a pure quadratic, unconstraind problem, and minimize the set of variables

$$\min_{x_1} \; x^T P x$$

$$\begin{bmatrix} x_1^T & x_2^T \end{bmatrix} \begin{bmatrix} \boxed{P_{11}} & P_{12} \\ P_{21} & P_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

the result is quadratic in the remaing variables

$$(x_1^T P_{11} + x_2^T P_{21}) x_1 + (x_1^T P_{12} + x_2^T P_{22}) x_2$$

$$= x_1^T P_{11} x_1 + x_1^T P_{12} x_2 + x_2^T P_{21} x_1 + x_2^T P_{22} x_2$$

$$= x_1^T P_{11} x_1 + 2 x_1^T P_{12} x_2 + x_2^T P_{22} x_2$$

$$\nabla x_1 = 2 P_{11} x_1 + 2 P_{12} x_2 = 0$$

$$x_1 = - P_{11}^{-1} P_{12} x_2$$

$$\inf_{x_1} x^T P x = x_2^T P_{12}^T P_{11}^{-T} P_{12} x_2^T - x_2^T P_{12} P_{11}^{-1} P_{12} x_2^T + x_2^T P_{22} x_2$$

$$= x_2^T \big(\underbrace{P_{22} - P_{12}^T P_{11}^{-1} P_{12}}_{\text{schur complement}} \big) x_2$$

· Dual decomposition

$$\min \; f_1(x_1, y) + f_2(x_2, y)$$

step 1: introducing new variable

$$\min \; f_1(x_1, y_1) + f_2(x_2, y_2)$$
$$\text{s.t.} \quad y_1 = y_2$$

· y_1, y_2 are local versions of complicating variable

· $y_1 = y_2$ is consistency constraint

step 2: form dual problem

$$L(x_1, y_1, x_2, y_2, v) = f_1(x_1, y_1) + f_2(x_2, y_2) + v^T(y_1 - y_2)$$

$$g(v) = \inf_{x_1, y_1, x_2, y_2} L$$

separable: can minimize over (x_1, y_1) and (x_2, y_2) separately

$$g_1(v) = \inf_{x_1, y_1} f_1(x_1, y_1) + v^T y_1$$

$$g_2(v) = \inf_{x_2, y_2} f_2(x_2, y_2) - v^T y_2$$

think v as price, each dual function uses the price to calculate an "action" x

$\inf_{x_1, y_1} f_1(x_1, y_1) + v^T y_1$: use as much y_1 as you like, but you have to pay for it

dual problem is: $\max \; g(v) = g_1(v) + g_2(v)$

computing $g_i(v)$ are the dual subproblems

can be done in parallel

v becomes a subsidy for $f_2(x_2, y_2) - v^T y_2$

a subgradient of $-g(v)$ is $\underbrace{y_2 - y_1}_{\text{inconsistency}}$ ($g(v)$ is concave $-g(v)$ is convex)

max $g(v)$ finds a right price. if a component of y is increased y_1 will get smaller in solving $g_1(v)$ next round

Dual decomposition algorithm (using subgradient method)

repeat

1. Solve the dual subproblem

   Find x_1 y_1 that minimize $f_1(x_1, y_1) + v^T y_1$

   Find x_2 y_2 that minimize $f_2(x_2, y_2) - v^T y_2$

2. update dual variable (prices)

   $v := v - \alpha_k (y_2 - y_1)$

$$\min \; f_1(x_1\, y_1) + f_2(x_2\, y_2)$$
$$\text{s.t.} \quad y_1 = y_2$$

$$g(v) = \inf_{x_1 y_1\, x_2\, y_2} \{ f_1(x_1\, y_1) + f_2(x_2\, y_2) + v^T(y_1 - y_2) \}$$
$$\le \inf_{\substack{x_1\, x_2\, y_1\, y_2 \\ y_1 = y_2}} f_1(x_1, y_1) + f_2(x_2\, y_2) + v^T(y_1 - y_2)$$
$$= p^*$$

at each step we have a lower bound $g(v)$ on p^*

when consistency is achieved, $y_2 - y_1 = 0$, v maximizes $g(v)$

Finding feasible iterates

reasonable guess of feasible point from $(x_1\, y_1)$ (x_2, y_2)

$(x_1\; \hat{y})$ $(x_2\; \hat{y})$ $\quad \hat{y} = (y_1 + y_2)/2$

projection onto feasible set $y_1 = y_2$

gives an upper bound $p^* \le f_1(x_1, \hat{y}) + f_2(x_2\; \hat{y})$

a better feasible point: replace y_1 y_2 with $\hat{y}$ and solve primal subproblems

$\hat{y} = (y_1 + y_2)/2 \qquad x_1 = \arg\min_{x_1} f_1(x_1, \hat{y}) \qquad x_2 = \arg\min_{x_2} f_2(x_2\; \hat{y})$

gives a better upper bound $p^* \le \min_{x_1} f_1(x_1\; \hat{y}) + \min_{x_2} f_2(x_2, \hat{y})$

Decomposition with constraints

$$\left.\begin{aligned} \min.\quad & f_1(x_1) + f_2(x_2) \\ \text{s.t.}\quad & x_1 \in C_1 \quad x_2 \in C_2 \\ & h_1(x_1) + h_2(x_2) \le 0 \end{aligned}\right\} \quad f_i \;\; h_i \;\; C_i \text{ are convex}$$

$h_1(x) + h_2(x) \le 0$ is a set of complicating / coupling constraints

(shared resource problem)

· primal decomposition with constraints

subproblem 1: $\min.\ f_1(x_1)$

$\text{s.t.}\ x_1 \in C_1 \quad h_1(x_1) \le t$

subproblem 2: $\min.\ f_2(x_2)$

$\text{s.t.}\ x_2 \in C_2 \quad h_2(x_2) \le -t$

master problem: $\min_t.\ \phi_1(t) + \phi_2(t) = \min_t\ p_1^*(t) + p_2^*(t)$

t is the quantity of resources allocated to first subproblem

subproblems can be solved separately when t is fixed

· Primal decomposition algorithm with constraints

$\min.\ f_1(x_1) + f_2(x)$

$\text{s.t.}\ x_1 \in C_1 \quad x_2 \in C_2 \quad h_1(x_1) + h_2(x_2) \le 0$

$\min.\ f_1(x_1)$ $\qquad$ $\min.\ f_2(x)$

$\text{s.t.}\ x_1 \in C \quad h_1(x) \le t$ $\qquad$ $\text{s.t.}\ x_2 \in C \quad h_2(x) \le -t$

$\mathcal{L}_1(x_1, \lambda_1) = f_1(x) + \lambda_1^T (h_1(x) - t)$ $\qquad$ $\mathcal{L}(x_2, \lambda_2) = f_2(x) + \lambda_2^T (h_2(x) + t)$

repeat

1. solve the subproblems

   solve subproblem 1, find $x_1^*(t)$ and $\lambda_1^*(t)$ $\quad$ $-\lambda_1^*(t)$ is a subgradient for $p_1^*(t)$

   solve subproblem 2, find $x_2^*(t)$ and $\lambda_2^*(t)$ $\quad$ $\lambda_2^*(t)$ is a subgradient for $p_2^*(t)$

2. update resource allocation

   $t := t - \alpha_k (\lambda_2 - \lambda_1)$

$\min.\ f_0(x)$

$\text{s.t.}\ f_i(x) \le a_i$

$h_i(x) = b_i$

$\Downarrow$

$p^*(a, b)$

$\mathcal{L}(x, \lambda, v) = f_0(x) + \lambda^T (f(x) - a) + v^T (h(x) - b)$

$= f_0(x) + \lambda^T f(x) + v^T h(x) - a^T \lambda - b^T v$

$\inf_x \mathcal{L}(x, \lambda, v) = \inf_x [f_0(x) + \lambda^T f(x) + v^T h(x)] - a^T \lambda - b^T v$

$\max.\ g(\lambda, v) - \lambda^T a - b^T v$

$p^*(a,b) \ge g(\lambda^*, v^*) - \lambda^{*T} a - v^{*T} b$

$= p^*(0,0) - \lambda^{*T} a - v^{*T} b$

all iterates are feasible

(when subproblems are feasible)

when t is a scalar, f_i's are differentiable

$\lambda_1^* = -\dfrac{\partial p_1^*(0)}{\partial t}$ $\qquad$ $\lambda_2 - \lambda_1 = \dfrac{\partial p_1^*(0)}{\partial t} + \dfrac{\partial p_2^*(0)}{\partial t}$ $\qquad$ $\therefore (-\lambda^*, -v^*)$ is a subgradient of $p^*(0,0)$

$\lambda_2^* = \dfrac{\partial p_2^*(0)}{\partial t}$

it's a gradient descent !!!

Dual decomposition with constraints

$$\min \; f_1(x_1) + f_2(x_2)$$
$$\text{s.t.} \quad x_1 \in C_1 \quad x_2 \in C_2 \quad h_1(x_1) + h_2(x_2) \le 0$$

form separable partial lagrangian

$$L(x_1, x_2, \lambda) = f_1(x_1) + f_2(x_2) + \lambda^T (h_1(x_1) + h_2(x_2))$$
$$= f_1(x_1) + \lambda^T h_1(x_1) + f_2(x_2) + \lambda^T h_2(x_2)$$

fix dual variable and define

subproblem 1: $\min. \; f_1(x_1) + \lambda^T h_1(x_1)$ s.t. $x_1 \in C_1$ $\Rightarrow$ optimal value $g_1(\lambda)$

subproblem 2: $\min. \; f_2(x_2) + \lambda^T h_2(x_2)$ s.t. $x_2 \in C_2$ $\Rightarrow$ optimal value $g_2(\lambda)$

$$\min_x \; f(x) + \lambda^T h(x) \quad \text{s.t.} \; x \in C \qquad x^*$$

$$\min_x \; f(x) + \lambda^T h(x) + \delta\lambda^T h(x) \quad \text{s.t.} \; x \in C \qquad \tilde{x}^*$$

$$= f(\tilde{x}^*) + \lambda^T h(\tilde{x}^*) + \delta\lambda^T h(\tilde{x}^*)$$

$$\ge \left[\min_x \; f(x) + \lambda^T h(x) \quad \text{s.t.} \; x \in C \right] + \delta\lambda^T h(x^*)$$

$\therefore$ $h(x^*)$ is a subgradient of $p^*(\lambda) = \min f(x) + \lambda^T h(x)$ s.t. $x \in C$

Dual decomposition algorithm with constraints

repeat

1. solve the subproblems

   solve subproblem 1, finding an optimal $\tilde{x}_1$

   solve subproblem 2, finding an optimal $\tilde{x}_2$

2. update the dual variable

   $$\lambda := [\lambda + \alpha_k (h_1(x_1) + h_2(x_2))]_+$$

   (projected subgradient method)

   if $h_1(x_1) + h_2(x_2) > 0$
   then they're using too much resources
   update the price to higher (keep the price positive)

· General decomposition structures

multiple subsystems

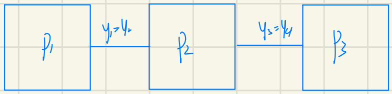
coupling variables / constraints between set of systems

represent as hypergraph with subsystems as vertices, coupling as hyperedge

eg. 3 subsystems

private variables : $x_1 \quad x_2 \quad x_3$

public variables : $y_1 \quad (y_2 \ y_3) \quad y_4$



$$\min. \quad f_1(x_1\ y_1) + f_2(x_2\ y_2\ y_3) + f_3(x_3\ y_4)$$

$$(x_1, y_1) \in C_1 \quad (x_2\ y_2\ y_3) \in C_2 \quad (x_3\ y_4) \in C_3$$

$$y_1 = y_2 \quad y_3 = y_4$$

· General form

$$\min. \ \sum_{i=1}^{k} f_i(x_i, y_i)$$

s.t. $(x_i, y_i) \in C_i \qquad i = 1, \cdots, k$

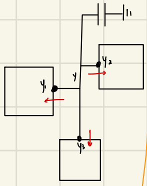
$y_i = E_i z \qquad i = 1, \cdots, k$ consistency constraints

private variables $x_i \in \mathbb{R}^{n_i}$ public variables $y_i \in \mathbb{R}^{p_i}$

hyperedge variables $z \in \mathbb{R}^{N}$; z_i is common value of public variables in hyperedge i

matrix E_i gives the hypergraph, $E_i \in \mathbb{R}^{p_i \times N}$; row k is e_p, the kth entry of y_i is in hyperedge p

· Primal decomposition



$$\min. \ f_1(x_1, y) + f_2(x_2, y) + f_3(x_3, y)$$

$\phi_i(y_i)$ is the optimal value of subproblem

$$\min_{x_i} \ f_i(x_i, y_i)$$

s.t. $(x_i, y_i) \in C_i$

enforce a voltage on this hyperedge (y_1, y_2, y_3)

$$\min_{x_i} \ f_i(x_i, y) \Rightarrow x_i \quad g_i \in \partial\phi_i(y_i)$$

each subsystem optimizes under the voltage, and each optimal reports a current (the Lagrangian multiplier) v_1, v_2, v_3, if $v_1 + v_2 + v_3 = 0$, then

if the currents sums up to positive, the decrease the common voltage, then less current will flow

repeat :

1. distribute net variable to subsystems $y_i := E_i z$
2. optimize subsystem (separately) solve subproblems to find x_i and $g_i \in \partial\phi_i(y_i)$
3. collect and sum up the subgradient for each hyperedge $g := \sum_{i=1}^{k} E_i^T g_i$
4. update hyperedge variable $z := z - \alpha_k g$

Fix voltage and wait currents to equilibrate

· Dual decomposition

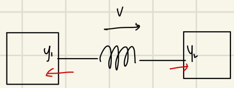
$g_i(v_i)$ is optimal value of subproblem

$$\begin{aligned} \min.\ & f_i(x_i, y_i) + v_i^T y_i \\ \text{s.t.}\ & (x_i, y_i) \in C_i \end{aligned}$$

v_i is the dual variable associated with consistency constraint

repeat:

1. optimize subsystem separately   solve subproblem to obtain x_i y_i
2. compute average value of public variables over each hyperedge $\tilde{z} = (E^T E)^{-1} E^T y$
3. update price on public variables $v := v + \alpha (y - E\tilde{z})$



a negative current $-v$ a positive current v

dual decomposition fixes a current and develop voltages y_1 and y_2

if voltages are equal, then done,

otherwise, y_2 might be bigger than y_1

in a small time later. v will decrease

wait for equilibrium, at equilibrium, we'll get a constant current, which is the optimal lagrangian multiplier

fix current and wait voltage to equilibrate