Primal-dual methods for nonconvex problems

Claudia Sagastizábal
IMECC-UNICAMP Brazil
joint with M. Cordova and W. de Oliveira

CMO - BIRS
Oaxaca, September 23rd, 2019

Motivation: UC problems

Two power plants

$$y_T \in \mathcal{S}_T$$
$$\langle \mathcal{F}, x \rangle + f_T(y_T)$$
$$x \in \{0,1\} \text{ and } y_T \leq x\, y^{up}$$

$$y_H \in \mathcal{S}_H$$
$$f_H(y_H)$$

$$y_T + y_H = d$$

Motivation: UC problems

Two power plants (**three** in fact)

$y_T \in \mathcal{S}_T$

$\langle \mathcal{F}, x \rangle + f_T(y_T)$

$x \in \{0, 1\}$ and $y_T \leq x\, y^{up}$

$y_H \in \mathcal{S}_H$

$f_H(y_H)$

$$y_T + y_H = d$$

net demand, substracting the WIND

Stylized Brazilian UC problem (2020)

$$
\left\{
\begin{array}{ll}
\min & \langle \mathcal{F}, x \rangle + f(y) \\
\text{s.t.} & x \in \{0,1\}, y \geq 0 \\
 & y \in \boldsymbol{\mathcal{S}} \\
 & By = d \\
 & Cy \leq e \\
 & x\, y_{low} \leq y \leq x\, y^{up}
\end{array}
\right.
$$

{ water balance

{ demand, flow limits

{ generation only if switched on

- f is convex, linear o quadratic

Real-life UC problem

Dessem com UCT: Formulação matemática

Eletrobras Cepel

$$\min \sum_{t=1}^{T}\sum_{i=1}^{NUT} c_i(gt_i^t) + S_i^t + \alpha^T(V^T)$$

s.a.

Demanda

$$\sum_{j\in\Theta_i} gt_j^t + \sum_{j\in\Psi_i} gh_j^t + \sum_{j\in\Omega_i}\left(Int_{j\to i}^t - Int_{i\to j}^t\right) = D_i^t \qquad i = 1,\ldots,NS,\ t = 1,\ldots,T$$

$$Int_{i\to j}^t \le \overline{Int_{i\to j}} \qquad i, j = 1,\ldots,NS,\ t = 1,\ldots,T$$

E

Conservação da água

$$V_i^t = V_i^{t-1} + I_i^t - (Q_i^t + S_i^t) + \sum_{j\in M_i}(Q_j^t + S_j^t)$$

FPHA

$$gh_i^t = FPH(V_i^t, Q_i^t, S_i^t)$$

Restrições operativas

$$\underline{V_i^t} \le V_i^t \le \overline{V_i^t},\quad \underline{Q_i^t} \le Q_i^t \le \overline{Q_i^t},\quad \underline{gh_i^t} \le gh_i^t \le \overline{gh_i^t} \qquad i = 1,\ldots,NH,\ t = 1,\ldots,T$$

H

Restrições Térmicas:

Unit Commitment

$$\underline{gt_i}\cdot u_i^t \le gt_i^t \le \overline{gt_i}\cdot u_i^t \qquad \left|gt_i^t - gt_i^{t+1}\right| \le R$$

$$\sum_{k=t}^{t+Ton_i-1} u_i^k \ge Ton_i\cdot(u_i^t - u_i^{t-1}) \qquad Cs_i\left(u_i^{t-1} - u_i^t\right) \le S_i^t$$

$$\sum_{k=t}^{t+Toff_i-1} (1-u_i^k) \ge Toff_i\cdot(u_i^{t-1} - u_i^t) \qquad u_i^t \in \{0,1\}$$

$$i = 1,\ldots,NUT,\ t = 1,\ldots,T,$$

T

Real-life UC problem

	COM UCT				SEM UCT	
	VAR.	VAR. INT.	REST.	REST. INT.	VAR.	REST.
1 DIA	327649	83121	433120	163520	327649	269600
2 DIAS	349413	80979	434843	159142	349413	275701
3 DIAS	375835	92831	490513	182545	375835	307968
4 DIAS	402171	86753	481152	170721	402171	310431
5 DIAS	429323	93913	512149	184730	429323	327419
6 DIAS	453110	111214	589245	218656	453110	370589
7 DIAS	476453	109717	590785	215853	476453	374932
GERAL	400897	93586	501438	184063	400897	317375

TABELA: Quantidade média de variáveis e restrições

Real-life UC problem

	COM UCT				SEM UCT	
	VAR.	VAR. INT.	REST.	REST. INT.	VAR.	REST.
1 DIA	327649	83121	433120	163520	327649	269600
2 DIAS	349413	80979	434843	159142	349413	275701
3 DIAS	375835	92831	490513	182545	375835	307968
4 DIAS	402171	86753	481152	170721	402171	310431
5 DIAS	429323	93913	512149	184730	429323	327419
6 DIAS	453110	111214	589245	218656	453110	370589
7 DIAS	476453	109717	590785	215853	476453	374932
GERAL	400897	93586	501438	184063	400897	317375

TABELA: Quantidade média de variáveis e restrições

for the deterministic formulation...

Stylized UC problem: stochastic version

$$
\left\{
\begin{array}{lll}
\min & \langle \mathcal{F}, x \rangle + \mathbb{E}[f(y(\omega))] & \\
\text{s.t.} & x \in \{0,1\}, \qquad y(\omega) \geq 0 & \\
 & y(\omega) \in \mathcal{S}(\omega) & \{ \text{water balance} \\
 & By(\omega) = d(\omega) & \text{demand} \\
 & Cy(\omega) \leq e(\omega) & \text{flow limits} \\
 & x\, y_{low} \leq y(\omega) \leq x\, y^{up} & \text{generation only if switched on, ramps}
\end{array}
\right.
$$

- f is convex, linear o quadratic
- d and e have uncertain components (wind and inflows)

Stylized UC problem: stochastic version

$$\begin{cases} \min & \langle \mathcal{F}, x \rangle + \mathbb{E}[f(y(\omega))] \\ \text{s.t.} & x \in \{0,1\}, \qquad y(\omega) \geq 0 \\ & y(\omega) \in \mathcal{S}(\omega) \\ & By(\omega) = d(\omega) \\ & Cy(\omega) \leq e(\omega) \\ & Tx + Wy(\omega) \leq h \end{cases}$$

{ water balance

{ demand
flow limits

{ generation only
if switched on
ramps

- f is convex, linear o quadratic
- d and e have uncertain components (wind and inflows)

Stylized UC problem: stochastic version

$$
\begin{cases}
\min & \langle \mathcal{F}, x \rangle + \mathbb{E}[f(y(\omega))] \\
\text{s.t.} & x \in \{0,1\}, \qquad y(\omega) \geq 0 \\
 & y(\omega) \in \mathcal{S}(\omega) & \{ \text{ water balance} \\
 & By(\omega) = d(\omega) & \text{demand} \\
 & Cy(\omega) \leq e(\omega) & \text{flow limits} \\
 & Tx + Wy(\omega) \leq h & \text{generation only if switched on, ramps}
\end{cases}
$$

- f is convex, linear o quadratic
- d and e have uncertain components (wind and inflows)

Interested in preserving separability along technologies: Lagrangian relaxation

Lagrangian relaxation 101

$$\left\{\begin{array}{rl} \min & f_T(x, y_T) + f_H(y_H) \\ \text{s.t.} & x \in \{0,1\}, y_T \in \mathcal{S}_T \\ & y_T \leq x\, y^{up} \\ & y_H \in \mathcal{S}_H \\ & y_T + y_H = d \qquad \leftrightarrow \lambda \end{array}\right.$$

$$\Longrightarrow$$

$$\begin{aligned} L(x, y, \lambda) \quad = \quad & f_T(x, y_T) + f_H(y_H) \\ & + \langle \lambda, d - y_T - y_H \rangle \end{aligned}$$

Lagrangian relaxation 101

$$\left\{\begin{array}{rl} \min & f_T(x, y_T) + f_H(y_H) \\ \text{s.t.} & x \in \{0,1\}, y_T \in \mathcal{S}_T \\ & y_T \leq x\, y^{up} \\ & y_H \in \mathcal{S}_H \\ & y_T + y_H = d \qquad \leftrightarrow \lambda \end{array}\right.$$

$$\Longrightarrow$$

$$\begin{aligned} L(x, y, \lambda) &= f_T(x, y_T) + f_H(y_H) \\ &\quad + \langle \lambda, d - y_T - y_H \rangle \\ &= L_T(x, y_T, \lambda) \\ &\qquad + L_H(y_H, \lambda) \\ &\qquad + \langle \lambda, d \rangle \end{aligned}$$

$$\Updownarrow$$

Lagrangian relaxation 101

$$\left\{\begin{array}{rl} \min & f_T(x,y_T)+f_H(y_H) \\ \text{s.t.} & x\in\{0,1\}, y_T\in\mathcal{S}_T \\ & y_T \le x\, y^{up} \\ & y_H \in \mathcal{S}_H \\ & y_T + y_H = d \qquad \leftrightarrow \lambda \end{array}\right.$$

$$\Longrightarrow$$

$$\begin{aligned} L(x,y,\lambda) &= f_T(x,y_T)+f_H(y_H) \\ &\quad +\langle \lambda, d-y_T-y_H\rangle \\ &= L_T(x,y_T,\lambda) \\ &\qquad +L_H(y_H,\lambda) \\ &\qquad +\langle \lambda, d\rangle \end{aligned}$$

$$\Updownarrow$$

$$\left\{\begin{array}{rl} \min\limits_{x,y}\max\limits_{\lambda} & L(x,y,\lambda) \\ \text{s.t.} & x\in\{0,1\}, y_T\in\mathcal{S}_T \\ & y_T \le x\, y^{up} \\ & y_H \in \mathcal{S}_H \end{array}\right.$$

Lagrangian relaxation 101

$$\left\{\begin{array}{rl} \min & f_T(x,y_T)+f_H(y_H) \\ \text{s.t.} & x\in\{0,1\}, y_T\in\mathcal{S}_T \\ & y_T \le x\,y^{up} \\ & y_H\in\mathcal{S}_H \\ & y_T+y_H=d \qquad \leftrightarrow \lambda \end{array}\right.$$

$$\Longrightarrow$$

$$\begin{array}{rcl} L(x,y,\lambda) & = & f_T(x,y_T)+f_H(y_H) \\ & & +\langle\lambda, d-y_T-y_H\rangle \\ & = & L_T(x,y_T,\lambda) \\ & & \quad +L_H(y_H,\lambda) \\ & & \quad +\langle\lambda,d\rangle \end{array}$$

DUAL: maxmin replaces minmax

$$\left\{\begin{array}{rl} \max\limits_{\lambda}\min\limits_{x,y} & L(x,y,\lambda) \\ \text{s.t.} & x\in\{0,1\}, y_T\in\mathcal{S}_T \\ & y_T \le x\,y^{up} \\ & y_H\in\mathcal{S}_H \end{array}\right.$$

Lagrangian relaxation 101

$$\left\{\begin{array}{ll} \min & f_T(x,y_T)+f_H(y_H) \\ \text{s.t.} & x\in\{0,1\}, y_T\in\mathcal{S}_T \\ & y_T\le x\,y^{up} \\ & y_H\in\mathcal{S}_H \\ & y_T+y_H=d \qquad \leftrightarrow \lambda \end{array}\right. \Longrightarrow \begin{array}{rcl} L(x,y,\lambda) & = & f_T(x,y_T)+f_H(y_H) \\ & & +\langle\lambda, d-y_T-y_H\rangle \\ & = & L_T(x,y_T,\lambda) \\ & & \quad +L_H(y_H,\lambda) \\ & & \quad +\langle\lambda,d\rangle \end{array}$$

DUAL: maxmin replaces minmax

$$\left\{\begin{array}{ll} \max_{\lambda}\min_{x,y} & L(x,y,\lambda) \\ \text{s.t.} & x\in\{0,1\}, y_T\in\mathcal{S}_T \\ & y_T\le x\,y^{up} \\ & y_H\in\mathcal{S}_H \end{array}\right. \Longrightarrow \max_{\lambda} \quad \psi_T(\lambda)+\psi_H(\lambda)+\langle\lambda,d\rangle$$

Lagrangian relaxation 101

$$\left\{\begin{array}{ll} \min & f_T(x,y_T)+f_H(y_H) \\ \text{s.t.} & x\in\{0,1\}, y_T\in\mathcal{S}_T \\ & y_T\le x\,y^{up} \\ & y_H\in\mathcal{S}_H \\ & y_T+y_H=d \qquad \leftrightarrow \lambda \end{array}\right. \Longrightarrow \begin{array}{rcl} L(x,y,\lambda) & = & f_T(x,y_T)+f_H(y_H) \\ & & +\langle\lambda, d-y_T-y_H\rangle \\ & = & L_T(x,y_T,\lambda) \\ & & \quad +L_H(y_H,\lambda) \\ & & \quad +\langle\lambda,d\rangle \end{array}$$

DUAL: maxmin replaces minmax

$$\left\{\begin{array}{ll} \max\limits_{\lambda}\min\limits_{x,y} & L(x,y,\lambda) \\ \quad\text{s.t.} & x\in\{0,1\}, y_T\in\mathcal{S}_T \\ & y_T\le x\,y^{up} \\ & y_H\in\mathcal{S}_H \end{array}\right. \Longrightarrow \begin{array}{ll} \max\limits_{\lambda} & \psi_T(\lambda)+\psi_H(\lambda)+\langle\lambda,d\rangle \\ \text{for} & \psi_T(\lambda):=\left\{\begin{array}{ll} \min & L_T(x,y_T,\lambda) \\ & x\in\{0,1\} \\ & y_T\in\mathcal{S}_T \\ & y_T\le x\,y^{up} \end{array}\right. \end{array}$$

Lagrangian relaxation 101

$$\begin{cases} \min & f_T(x,y_T) + f_H(y_H) \\ \text{s.t.} & x \in \{0,1\}, y_T \in \mathcal{S}_T \\ & y_T \leq x\, y^{up} \\ & y_H \in \mathcal{S}_H \\ & y_T + y_H = d \qquad \leftrightarrow \lambda \end{cases} \implies \begin{aligned} L(x,y,\lambda) &= f_T(x,y_T) + f_H(y_H) \\ &\quad + \langle \lambda, d - y_T - y_H \rangle \\ &= L_T(x,y_T,\lambda) \\ &\quad + L_H(y_H,\lambda) \\ &\quad + \langle \lambda, d \rangle \end{aligned}$$

DUAL : maxmin replaces minmax

(duality gap)

$$\begin{cases} \max\limits_{\lambda} \min\limits_{x,y} & L(x,y,\lambda) \\ \text{s.t.} & x \in \{0,1\}, y_T \in \mathcal{S}_T \\ & y_T \leq x\, y^{up} \\ & y_H \in \mathcal{S}_H \end{cases} \implies \begin{aligned} &\max_{\lambda} \quad \psi_T(\lambda) + \psi_H(\lambda) + \langle \lambda, d \rangle \\ &\text{for} \quad \psi_T(\lambda) := \begin{cases} \min & L_T(x,y_T,\lambda) \\ & x \in \{0,1\} \\ & y_T \in \mathcal{S}_T \\ & y_T \leq x\, y^{up} \end{cases} \end{aligned}$$

Lagrangian relaxation - 2nd version

$$\left\{\begin{array}{ll} \min & f_T(x, z_T) + f_H(y_H) \\ \text{s.t.} & x \in \{0,1\}, z_T \in \mathcal{S}_T \\ & z_T \leq x\, y^{up} \\ & y_H \in \mathcal{S}_H \\ & y_T + y_H = d \\ & y_T - z_T = 0 \qquad \leftrightarrow \lambda \end{array}\right.$$

Lagrangian relaxation - 2nd version

$$\left\{\begin{array}{ll} \min & f_T(x,z_T)+f_H(y_H) \\ \text{s.t.} & x\in\{0,1\}, z_T\in\mathcal{S}_T \\ & z_T \le x\, y^{up} \\ & y_H \in \mathcal{S}_H \\ & y_T + y_H = d \\ & y_T - z_T = 0 \quad \leftrightarrow \lambda \end{array}\right. \Longrightarrow$$

$$\begin{aligned} L(x,y,z,\lambda) &= f_T(x,z_T)+f_H(y_H) \\ &\quad +\langle \lambda, y_T - z_T\rangle \\ &= L_T(x,z_T,\lambda) \\ &\quad + L_{HT}(y_T,y_H,\lambda) \\ &\quad +\langle \lambda, d\rangle \end{aligned}$$

Lagrangian relaxation - 2nd version

$$
\left\{
\begin{array}{ll}
\min & f_T(x, z_T) + f_H(y_H) \\
\text{s.t.} & x \in \{0,1\}, z_T \in \mathcal{S}_T \\
& z_T \leq x\, y^{up} \\
& y_H \in \mathcal{S}_H \\
& y_T + y_H = d \\
& y_T - z_T = 0 \quad \leftrightarrow \lambda
\end{array}
\right.
\Longrightarrow
\begin{array}{rcl}
L(x,y,z,\lambda) & = & f_T(x,z_T) + f_H(y_H) \\
& & \quad + \langle \lambda, y_T - z_T \rangle \\
& = & L_T(x, z_T, \lambda) \\
& & \quad + L_{HT}(y_T, y_H, \lambda) \\
& & \quad + \langle \lambda, d \rangle
\end{array}
$$

DUAL: maxmin replaces minmax

$$
\left\{
\begin{array}{ll}
\max\limits_{\lambda} \min\limits_{x,y,z} & L(x,y,z,\lambda) \\
\text{s.t.} & x \in \{0,1\}, z_T \in \mathcal{S}_T \\
& z_T \leq x\, y^{up} \\
& y_H \in \mathcal{S}_H
\end{array}
\right.
$$

Lagrangian relaxation - 2nd version

$$\left\{\begin{array}{rl} \min & f_T(x,z_T)+f_H(y_H) \\ \text{s.t.} & x\in\{0,1\}, z_T\in\mathcal{S}_T \\ & z_T \le x\, y^{up} \\ & y_H\in\mathcal{S}_H \\ & y_T+y_H=d \\ & y_T-z_T=0 \quad \leftrightarrow \lambda \end{array}\right. \Longrightarrow \begin{array}{rcl} L(x,y,z,\lambda) & = & f_T(x,z_T)+f_H(y_H) \\ & & +\langle\lambda, y_T-z_T\rangle \\ & = & L_T(x,z_T,\lambda) \\ & & +L_{HT}(y_T,y_H,\lambda) \\ & & +\langle\lambda,d\rangle \end{array}$$

DUAL: maxmin replaces minmax

$$\left\{\begin{array}{rl} \max\limits_{\lambda}\min\limits_{x,y,z} & L(x,y,z,\lambda) \\ \text{s.t.} & x\in\{0,1\}, z_T\in\mathcal{S}_T \\ & z_T\le x\, y^{up} \\ & y_H\in\mathcal{S}_H \end{array}\right. \Longrightarrow \max_{\lambda} \quad \psi_T(\lambda)+\psi_{HT}(\lambda)+\langle\lambda,d\rangle$$

Lagrangian relaxation - 2nd version

$$\left\{\begin{array}{ll} \min & f_T(x,z_T)+f_H(y_H) \\ \text{s.t.} & x\in\{0,1\}, z_T\in\mathcal{S}_T \\ & z_T\le x\,y^{up} \\ & y_H\in\mathcal{S}_H \\ & y_T+y_H=d \\ & y_T-z_T=0 \quad \leftrightarrow \lambda \end{array}\right. \Longrightarrow \begin{array}{rcl} L(x,y,z,\lambda) & = & f_T(x,z_T)+f_H(y_H) \\ & & +\langle\lambda, y_T-z_T\rangle \\ & = & L_T(x,z_T,\lambda) \\ & & +L_{HT}(y_T,y_H,\lambda) \\ & & +\langle\lambda,d\rangle \end{array}$$

DUAL: maxmin replaces minmax

$$\left\{\begin{array}{ll} \max\limits_{\lambda}\min\limits_{x,y,z} & L(x,y,z,\lambda) \\ \text{s.t.} & x\in\{0,1\}, z_T\in\mathcal{S}_T \\ & z_T\le x\,y^{up} \\ & y_H\in\mathcal{S}_H \end{array}\right. \Longrightarrow \begin{array}{ll} \max\limits_{\lambda} & \psi_T(\lambda)+\psi_{HT}(\lambda)+\langle\lambda,d\rangle \\ \text{for} & \psi_T(\lambda):=\left\{\begin{array}{ll} \min & L_T(x,z_T,\lambda) \\ & x\in\{0,1\} \\ & z_T\in\mathcal{S}_T \\ & z_T\le x\,y^{up} \end{array}\right. \end{array}$$

Lagrangian relaxation - 2nd version

$$\left\{\begin{array}{ll} \min & f_T(x,z_T)+f_H(y_H) \\ \text{s.t.} & x\in\{0,1\}\,,z_T\in\mathcal{S}_T \\ & z_T\le x\,y^{up} \\ & y_H\in\mathcal{S}_H \\ & y_T+y_H=d \\ & y_T-z_T=0 \qquad \leftrightarrow \lambda \end{array}\right. \Longrightarrow \begin{array}{rcl} L(x,y,z,\lambda) & = & f_T(x,z_T)+f_H(y_H) \\ & & +\langle\lambda,y_T-z_T\rangle \\ & = & L_T(x,z_T,\lambda) \\ & & +L_{HT}(y_T,y_H,\lambda) \\ & & +\langle\lambda,d\rangle \end{array}$$

DUAL: maxmin replaces minmax

$$\left\{\begin{array}{ll} \max\limits_{\lambda}\min\limits_{x,y,z} & L(x,y,z,\lambda) \\ \text{s.t.} & x\in\{0,1\}\,,z_T\in\mathcal{S}_T \\ & z_T\le x\,y^{up} \\ & y_H\in\mathcal{S}_H \end{array}\right. \Longrightarrow \max_{\lambda} \quad \psi_T(\lambda)+\psi_{HT}(\lambda)+\langle\lambda,d\rangle$$

Lagrangian relaxation - 2nd version

$$\left\{\begin{array}{ll} \min & f_T(x,z_T)+f_H(y_H) \\ \text{s.t.} & x\in\{0,1\}, z_T\in\mathcal{S}_T \\ & z_T\le x\,y^{up} \\ & y_H\in\mathcal{S}_H \\ & y_T+y_H=d \\ & y_T-z_T=0 \quad \leftrightarrow \lambda \end{array}\right. \Longrightarrow \begin{array}{rcl} L(x,y,z,\lambda) & = & f_T(x,z_T)+f_H(y_H) \\ & & +\langle\lambda, y_T-z_T\rangle \\ & = & L_T(x,z_T,\lambda) \\ & & +L_{HT}(y_T,y_H,\lambda) \\ & & +\langle\lambda,d\rangle \end{array}$$

DUAL: maxmin replaces minmax

$$\left\{\begin{array}{ll} \max\limits_{\lambda}\min\limits_{x,y,z} & L(x,y,z,\lambda) \\ \text{s.t.} & x\in\{0,1\}, z_T\in\mathcal{S}_T \\ & z_T\le x\,y^{up} \\ & y_H\in\mathcal{S}_H \end{array}\right. \Longrightarrow \begin{array}{ll} \max\limits_{\lambda} & \psi_T(\lambda)+\psi_{HT}(\lambda)+\langle\lambda,d\rangle \\ \text{for} & \psi_{HT}(\lambda):=\left\{\begin{array}{ll} \min & L_{HT}(y_T,y_H,\lambda) \\ & y_H\in\mathcal{S}_H \\ & y_T+y_H=d \end{array}\right. \end{array}$$

Lagrangian relaxation - 2nd version

$$\left\{\begin{array}{rl} \min & f_T(x,z_T)+f_H(y_H) \\ \text{s.t.} & x\in\{0,1\}, z_T\in\mathcal{S}_T \\ & z_T\leq x\, y^{up} \\ & y_H\in\mathcal{S}_H \\ & y_T+y_H=d \\ & y_T-z_T=0 \quad \leftrightarrow \lambda \end{array}\right. \Longrightarrow \begin{array}{rcl} L(x,y,z,\lambda) & = & f_T(x,z_T)+f_H(y_H) \\ & & \quad +\langle\lambda, y_T-z_T\rangle \\ & = & L_T(x,z_T,\lambda) \\ & & \quad +L_{HT}(y_T,y_H,\lambda) \\ & & \quad +\langle\lambda, d\rangle \end{array}$$

DUAL : maxmin replaces minmax

(duality gap)

$$\left\{\begin{array}{rl} \max\limits_{\lambda}\min\limits_{x,y,z} & L(x,y,z,\lambda) \\ \text{s.t.} & x\in\{0,1\}, z_T\in\mathcal{S}_T \\ & z_T\leq x\, y^{up} \\ & y_H\in\mathcal{S}_H \end{array}\right. \Longrightarrow \begin{array}{rl} \max\limits_{\lambda} & \psi_T(\lambda)+\psi_{HT}(\lambda)+\langle\lambda, d\rangle \\ \text{for} & \psi_{HT}(\lambda) := \left\{\begin{array}{rl} \min & L_{HT}(y_T,y_H,\lambda) \\ & y_H\in\mathcal{S}_H \\ & y_T+y_H=d \end{array}\right. \end{array}$$

Lagrangian relaxation - 3rd version

$$\begin{cases} \min & f_T(x, z_T) + f_H(y_H) \\ \text{s.t.} & x \in \{0,1\}, z_T \in \mathcal{S}_T \\ & z_T \leq x\, y^{up} \\ & y_H \in \mathcal{S}_H \\ & y_T + y_H = d \\ & y_T - z_T = 0 \qquad \leftrightarrow \lambda \end{cases}$$

p=(x,y,z)

$$\text{either} \begin{cases} \min & f(p) \\ \text{s.t.} & p \in \mathcal{S} \\ & F(p) = 0 \quad \leftrightarrow \lambda \end{cases}$$

Lagrangian relaxation - 3rd version

p=(x,y,z)

$$\text{either} \begin{cases} \min & f(p) \\ \text{s.t.} & p \in \mathcal{S} \\ & F(p) = 0 \quad \leftrightarrow \lambda \end{cases}$$

Lagrangian relaxation - 3rd version

p=(x,y,z)

$$\text{either}\left\{\begin{array}{ll}\min & f(p)\\ \text{s.t.} & p\in\mathcal{S}\\ & F(p)=0 \quad \leftrightarrow \lambda\end{array}\right.$$

$$\text{or}\left\{\begin{array}{ll}\min & f(p)\\ \text{s.t.} & p\in\mathcal{S}\\ & F(p)=0 \quad \leftrightarrow \lambda\\ & \sigma(F(p))\leq 0 \quad \leftrightarrow \rho\in\mathbb{R}\end{array}\right.$$

$\sigma(t)=\|t\|$ or $\frac{1}{2}\|t\|^2$

$\sigma(0)=0$

Lagrangian relaxation - 3rd version

relation?

p=(x,y,z)

either
$$\begin{cases} \min & f(p) \\ \text{s.t.} & p \in \mathcal{S} \\ & F(p) = 0 \quad \leftrightarrow \lambda \end{cases}$$

or
$$\begin{cases} \min & f(p) \\ \text{s.t.} & p \in \mathcal{S} \\ & F(p) = 0 \quad \leftrightarrow \lambda \\ & \sigma(F(p)) \leq 0 \quad \leftrightarrow \rho \end{cases}$$

$\sigma(t) = \|t\|$ or $\frac{1}{2}\|t\|^2$

$\sigma(0) = 0$

Lagrangian relaxation - 3rd version

p=(x,y,z)

OC1: $\exists(\bar{p},\bar{\lambda}) \in \mathcal{S} \times \mathbb{R}^m :$

$$F(\bar{p}) = 0$$

and

$$0 \in f'(\bar{p}) + N_P(\bar{p}) + F'(\bar{p})\bar{\lambda}$$

either $\begin{cases} \min & f(p) \\ \text{s.t.} & p \in \mathcal{S} \\ & F(p) = 0 \end{cases}$ $\leftrightarrow \lambda$

Lagrangian relaxation - 3rd version

OC1: $\exists(\bar{p},\bar{\lambda}) \in \mathcal{S} \times \mathbb{R}^m$:

$$F(\bar{p}) = 0$$

and

$$0 \in f'(\bar{p}) + N_P(\bar{p}) + F'(\bar{p})\bar{\lambda}$$

p=(x,y,z)

either
$$\left\{\begin{array}{ll} \text{min} & f(p) \\ \text{s.t.} & p \in \mathcal{S} \\ & F(p) = 0 \quad \leftrightarrow \lambda \end{array}\right.$$

or
$$\left\{\begin{array}{ll} \text{min} & f(p) \\ \text{s.t.} & p \in \mathcal{S} \\ & F(p) = 0 \quad \leftrightarrow \lambda \\ & \sigma(F(p)) \le 0 \quad \leftrightarrow \rho \end{array}\right.$$

$\sigma(t) = \|t\|$ or $\frac{1}{2}\|t\|^2$

$\sigma(0) = 0$

Lagrangian relaxation - 3rd version

OC1: $\exists(\bar{p},\bar{\lambda}) \in \mathcal{S} \times \mathbb{R}^m$:

$$F(\bar{p}) = 0$$

and

$$0 \in f'(\bar{p}) + N_P(\bar{p}) + F'(\bar{p})\bar{\lambda}$$

p=(x,y,z)

either
$$\begin{cases} \min & f(p) \\ \text{s.t.} & p \in \mathcal{S} \\ & F(p) = 0 \quad \leftrightarrow \lambda \end{cases}$$

OC2: $\exists(\bar{p},\bar{\lambda},\bar{\rho}) \in \mathcal{S} \times \mathbb{R}^m \times \mathbb{R}_+$:

$$F(\bar{p}) = 0, \quad \sigma(F(\bar{p})) \leq 0$$

and

$$0 \in f'(\bar{p}) + N_P(\bar{p}) + F'(\bar{p})\bar{\lambda} + \bar{\rho} F'(\bar{p})\sigma(F(\bar{p}))$$

or
$$\begin{cases} \min & f(p) \\ \text{s.t.} & p \in \mathcal{S} \\ & F(p) = 0 \quad \leftrightarrow \lambda \\ & \sigma(F(p)) \leq 0 \quad \leftrightarrow \rho \end{cases}$$

$\sigma(t) = \|t\|$ or $\frac{1}{2}\|t\|^2$

$\sigma(0) = 0$

Lagrangian relaxation - 3rd version

p=(x,y,z)

OC1: $\exists(\bar{p},\bar{\lambda}) \in \mathcal{S} \times \mathbb{R}^m$:

$$F(\bar{p}) = 0$$

and

$$0 \in f'(\bar{p}) + N_P(\bar{p}) + F'(\bar{p})\bar{\lambda}$$

either

$$\begin{cases} \min & f(p) \\ \text{s.t.} & p \in \mathcal{S} \\ & F(p) = 0 \quad \leftrightarrow \lambda \end{cases}$$

OC2: $\exists(\bar{p},\bar{\lambda},\bar{\rho}) \in \mathcal{S} \times \mathbb{R}^m \times \mathbb{R}_+$:

$$F(\bar{p}) = 0\,, \quad \sigma(F(\bar{p})) \leq 0$$

and

$$0 \in f'(\bar{p}) + N_P(\bar{p}) + F'(\bar{p})\bar{\lambda}$$

$$+\bar{\rho}\,F'(\bar{p})\sigma(F(\bar{p}))$$

or

$$\begin{cases} \min & f(p) \\ \text{s.t.} & p \in \mathcal{S} \\ & F(p) = 0 \quad \leftrightarrow \lambda \\ & \sigma(F(p)) \leq 0 \quad \leftrightarrow \rho \end{cases}$$

$\sigma(t) = \|t\|$ or $\frac{1}{2}\|t\|^2$

$\sigma(0) = 0$

Lagrangian relaxation - 3rd version

$$\begin{cases} \min & f(p) \\ \text{s.t.} & p \in \mathcal{S} \\ & F(p) = 0 \qquad \leftrightarrow \lambda \\ & \sigma(F(p)) \leq 0 \qquad \leftrightarrow \rho \end{cases}$$

Lagrangian relaxation - 3rd version

$$\begin{cases} \min & f(p) \\ \text{s.t.} & p \in \mathcal{S} \\ & F(p) = 0 & \leftrightarrow \lambda \\ & \sigma(F(p)) \leq 0 & \leftrightarrow \rho \end{cases}$$

$$\Longrightarrow$$

$$\begin{aligned} L(p, \lambda, \rho) &= f(p) + \langle \lambda, F(p) \rangle \\ &\quad + \rho \sigma(F(p)) \end{aligned}$$

Lagrangian relaxation - 3rd version

$$\begin{cases} \min & f(p) \\ \text{s.t.} & p \in \mathcal{S} \\ & F(p) = 0 & \leftrightarrow \lambda \\ & \sigma(F(p)) \leq 0 & \leftrightarrow \rho \end{cases} \Longrightarrow \quad \begin{aligned} L(p,\lambda,\rho) \quad = \quad & f(p) + \langle \lambda, F(p) \rangle \\ & + \rho\sigma(F(p)) \end{aligned}$$

DUAL: maxmin replaces minmax

$$\begin{cases} \max\limits_{\lambda,\rho} \min\limits_{p} & L(p,\lambda,\rho) \\ \text{s.t.} & p \in \mathcal{S} \end{cases}$$

Lagrangian relaxation - 3rd version

$$\left\{\begin{array}{ll} \min & f(p) \\ \text{s.t.} & p \in \mathcal{S} \\ & F(p) = 0 \qquad \leftrightarrow \lambda \\ & \sigma(F(p)) \leq 0 \qquad \leftrightarrow \rho \end{array}\right. \Longrightarrow \quad L(p,\lambda,\rho) = f(p) + \langle \lambda, F(p) \rangle + \rho\,\sigma(F(p))$$

DUAL: maxmin replaces minmax

NO duality gap

$$\left\{\begin{array}{ll} \max\limits_{\lambda,\rho} \min\limits_{p} & L(p,\lambda,\rho) \\ \text{s.t.} & p \in \mathcal{S} \end{array}\right. \Longrightarrow \quad \begin{array}{ll} \max\limits_{\lambda} & \psi(\lambda,\rho) \\ \text{for} & \psi(\lambda,\rho) := \left\{\begin{array}{ll} \min & L(p,\lambda,\rho) \\ & p \in \mathcal{S} \end{array}\right. \end{array}$$

Lagrangian relaxation - 3rd version

$$\begin{cases} \min & f(p) \\ \text{s.t.} & p \in \mathcal{S} \\ & F(p) = 0 & \leftrightarrow \lambda \\ & \sigma(F(p)) \leq 0 & \leftrightarrow \rho \end{cases} \Longrightarrow \begin{aligned} L(p,\lambda,\rho) &= f(p) + \langle \lambda, F(p) \rangle \\ &\quad + \rho\sigma(F(p)) \end{aligned}$$

DUAL: maxmin replaces minmax

NO duality gap

$$\begin{cases} \max\limits_{\lambda,\rho} \min\limits_{p} & L(p,\lambda,\rho) \\ \text{s.t.} & p \in \mathcal{S} \end{cases} \Longrightarrow \begin{array}{ll} \max\limits_{\lambda} & \psi(\lambda,\rho) \\ \text{for} & \psi(\lambda,\rho) := \begin{cases} \min & L(p,\lambda,\rho) \\ & p \in \mathcal{S} \end{cases} \end{array}$$

Sharp/Proximal Lagrangian $\quad \sigma(t) = \|t\|$ or $\frac{1}{2}\|t\|^2$

Lagrangian relaxation - 3rd version

$$\begin{cases} \min & f(p) \\ \text{s.t.} & p \in \mathcal{S} \\ & F(p) = 0 & \leftrightarrow \lambda \\ & \sigma(F(p)) \le 0 & \leftrightarrow \rho \end{cases} \Longrightarrow \quad L(p,\lambda,\rho) = f(p) + \langle \lambda, F(p) \rangle + \rho\sigma(F(p))$$

DUAL: maxmin replaces minmax

NO duality gap

$$\begin{cases} \max\limits_{\lambda,\rho} \min\limits_{p} & L(p,\lambda,\rho) \\ \text{s.t.} & p \in \mathcal{S} \end{cases} \Longrightarrow \quad \max_{\lambda} \ \psi(\lambda,\rho) \quad \text{for} \quad \psi(\lambda,\rho) := \begin{cases} \min & L(p,\lambda,\rho) \\ & p \in \mathcal{S} \end{cases}$$

Sharp/Proximal Lagrangian

$$\sigma(t) = \|t\| \text{ or } \frac{1}{2}\|t\|^2$$

NOTE: ρ is a dual variable

GAL: Generalizad Augmented Lagrangians

Chapter 11K [RW99], on Nonconvex Duality

$\mathcal{D}(p,v,w) := f(p) : F(p) = v\,, \sigma(F(p)) \leq w$ is level-bounded in $p \in \mathcal{S}$ locally uniformly in (v,w)

Theorem 11.59 (duality without convexity)

- Optimal solutions to the primal and dual problems, resp. $\bar{p}$ and $(\bar{\lambda},\bar{\rho})$, are saddle points of the Lagrangian

$$\left.\begin{array}{ll} \bar{p} & \text{minimizes the primal} \\ (\bar{\lambda},\bar{\rho}) & \text{maximizes the dual} \end{array}\right\} \iff \inf_p L(p,\bar{\lambda},\bar{\rho}) = L(\bar{p},\bar{\lambda},\bar{\rho}) = \sup_{\lambda,\rho} L(\bar{p},\lambda,\rho)$$

- Exact penalty $\iff \exists(\bar{\lambda},\bar{\rho}) \in \arg\max \psi(\lambda,\rho)$

$\Longrightarrow$ no need to drive $\rho \to \infty$, as in the augmented Lagrangian

GAL: algorithmic scheme

$$\left.\begin{array}{ll}\bar{p} & \text{minimizes the primal} \\ (\bar{\lambda},\bar{\rho}) & \text{maximizes the dual}\end{array}\right\} \iff \inf_p L(p,\bar{\lambda},\bar{\rho}) = L(\bar{p},\bar{\lambda},\bar{\rho}) = \sup_{\lambda,\rho} L(\bar{p},\lambda,\rho)$$

- **Primal Step:** Given (λ^k,ρ^k) current multipliers,
  p^k solves $\min_p L(p,\lambda^k,\rho^k)$ $\quad (\Longrightarrow \psi(\lambda^k,\rho^k) = L(p^k,\lambda^k,\rho^k))$

- **Dual Step:** $(\lambda^{k+1},\rho^{k+1})$ one iteration to solve $\max_{\lambda,\rho} \psi(\lambda,\rho)$

GAL: algorithmic scheme

$$\left.\begin{array}{ll} \bar{p} & \text{minimizes the primal} \\ (\bar{\lambda},\bar{\rho}) & \text{maximizes the dual} \end{array}\right\} \iff \inf_p L(p,\bar{\lambda},\bar{\rho}) = L(\bar{p},\bar{\lambda},\bar{\rho}) = \sup_{\lambda,\rho} L(\bar{p},\lambda,\rho)$$

- **Primal Step:** Given (λ^k,ρ^k) current multipliers,
  p^k solves $\min_p L(p,\lambda^k,\rho^k)$ $(\Longrightarrow \psi(\lambda^k,\rho^k) = L(p^k,\lambda^k,\rho^k))$
  $p^k := \mathcal{A}^{ex}(\lambda^k,\rho^k)$

- **Dual Step:** $(\lambda^{k+1},\rho^{k+1})$ one iteration to solve $\max_{\lambda,\rho} \psi(\lambda,\rho)$

GAL: algorithmic scheme

$$\left.\begin{array}{ll}\bar{p} & \text{minimizes the primal} \\ (\bar{\lambda},\bar{\rho}) & \text{maximizes the dual}\end{array}\right\} \iff \inf_p L(p,\bar{\lambda},\bar{\rho}) = L(\bar{p},\bar{\lambda},\bar{\rho}) = \sup_{\lambda,\rho} L(\bar{p},\lambda,\rho)$$

- **Primal Step:** Given (λ^k,ρ^k) current multipliers,
  p^k solves $\min_p L(p,\lambda^k,\rho^k)$ $\quad (\Longrightarrow \psi(\lambda^k,\rho^k) = L(p^k,\lambda^k,\rho^k))$
  $p^k := \mathcal{A}^{ex}(\lambda^k,\rho^k)$

- **Dual Step:** $(\lambda^{k+1},\rho^{k+1})$ one iteration to solve $\max_{\lambda,\rho} \psi(\lambda,\rho)$

How easy is to find p^k?

GAL: algorithmic scheme

$\bar{p}$ minimizes the primal
$(\bar{\lambda}, \bar{\rho})$ maximizes the dual $\Bigg\} \iff \inf_p L(p, \bar{\lambda}, \bar{\rho}) = L(\bar{p}, \bar{\lambda}, \bar{\rho}) = \sup_{\lambda,\rho} L(\bar{p}, \lambda, \rho)$

- **Primal Step:** Given (λ^k, ρ^k) current multipliers,
  p^k solves $\min_p L(p, \lambda^k, \rho^k)$ $\quad (\Longrightarrow \psi(\lambda^k, \rho^k) = L(p^k, \lambda^k, \rho^k))$
  $p^k := \mathcal{A}^{ex}(\lambda^k, \rho^k)$

- **Dual Step:** $(\lambda^{k+1}, \rho^{k+1})$ one iteration to solve $\max_{\lambda,\rho} \psi(\lambda, \rho)$

How easy is to find p^k?

$$L(p, \lambda, \rho) = L_T(x, z_T, \lambda) + L_{HT}(y_T, y_H, \lambda) + \langle \lambda, d \rangle + \rho \|y_T - z_T\|$$

GAL: algorithmic scheme

$$\left.\begin{array}{ll}\bar{p} & \text{minimizes the primal}\\ (\bar{\lambda},\bar{\rho}) & \text{maximizes the dual}\end{array}\right\} \iff \inf_p L(p,\bar{\lambda},\bar{\rho}) = L(\bar{p},\bar{\lambda},\bar{\rho}) = \sup_{\lambda,\rho} L(\bar{p},\lambda,\rho)$$

- **Primal Step:** Given (λ^k,ρ^k) current multipliers,
  p^k solves $\min_p L(p,\lambda^k,\rho^k)$ $\quad(\Longrightarrow \psi(\lambda^k,\rho^k) = L(p^k,\lambda^k,\rho^k))$
  $p^k := \mathcal{A}^{ex}(\lambda^k,\rho^k)$
- **Dual Step:** $(\lambda^{k+1},\rho^{k+1})$ one iteration to solve $\max_{\lambda,\rho} \psi(\lambda,\rho)$

How easy is to find p^k?

$$L(p,\lambda,\rho) = L_T(x,z_T,\lambda) + L_{HT}(y_T,y_H,\lambda) + \langle\lambda,d\rangle + \rho\|y_T - z_T\|$$

Separability is lost!

GALs: difficulties

- **Primal Step:** Given (λ^k, ρ^k) current multipliers,
  Instead of p^k solves $\min_p L(p, \lambda^k, \rho^k)$ $(\Longrightarrow \psi(\lambda^k, \rho^k) = L(p^k, \lambda^k, \rho^k))$
  $$p^k := \mathcal{A}^{ex}(\lambda^k, \rho^k)$$

GALs: difficulties

- **Primal Step:** Given (λ^k, ρ^k) current multipliers,
  Instead of p^k solves $\min_p L(p, \lambda^k, \rho^k)$ $(\Longrightarrow \psi(\lambda^k, \rho^k) = L(p^k, \lambda^k, \rho^k))$
  $$p^k := \mathcal{A}^{ex}(\lambda^k, \rho^k)$$
  $$p^k := \mathcal{A}^{inex}(\lambda^k, \rho^k) \approx \min_p L(p, \lambda^k, \rho^k)$$
  with an error $E^k := L(p^k, \lambda^k, \rho^k) - \psi(\lambda^k, \rho^k) \in [0, \eta]$
  Error is unknown, η bounds approximation inaccuracy

GALs: difficulties

- **Primal Step:** Given (λ^k, ρ^k) current multipliers,
  Instead of p^k solves $\min_p L(p, \lambda^k, \rho^k)$ $(\Longrightarrow \psi(\lambda^k, \rho^k) = L(p^k, \lambda^k, \rho^k))$
  $$p^k := \mathcal{A}^{ex}(\lambda^k, \rho^k)$$
  $$p^k := \mathcal{A}^{inex}(\lambda^k, \rho^k) \approx \min_p L(p, \lambda^k, \rho^k)$$
  with an error $E^k := L(p^k, \lambda^k, \rho^k) - \psi(\lambda^k, \rho^k) \in [0, \eta]$
  Error is unknown, η bounds approximation inaccuracy

- **Dual Step:** $(\lambda^{k+1}, \rho^{k+1})$ one iteration to solve $\max_{\lambda,\rho} \psi(\lambda, \rho)$
  Subgradient (Uzawa), Cutting-planes (DW), **Bundle**

GALs: difficulties

- **Primal Step:** Given (λ^k, ρ^k) current multipliers,
  Instead of p^k solves $\min_p L(p, \lambda^k, \rho^k)$ $(\Longrightarrow \psi(\lambda^k, \rho^k) = L(p^k, \lambda^k, \rho^k))$
  $$p^k := \mathcal{A}^{ex}(\lambda^k, \rho^k)$$
  $$p^k := \mathcal{A}^{inex}(\lambda^k, \rho^k) \approx \min_p L(p, \lambda^k, \rho^k)$$
  with an error $E^k := L(p^k, \lambda^k, \rho^k) - \psi(\lambda^k, \rho^k) \in [0, \eta]$
  Error is unknown, η bounds approximation inaccuracy

- **Dual Step:** $(\lambda^{k+1}, \rho^{k+1})$ one iteration to solve $\max_{\lambda,\rho} \psi(\lambda, \rho)$
  Subgradient (Uzawa), Cutting-planes (DW), **Bundle**

When is $\mathcal{A}^{inex}$ sufficiently good?

GALs: difficulties

- **Primal Step:** Given (λ^k, ρ^k) current multipliers,
  Instead of p^k solves $\min_p L(p, \lambda^k, \rho^k)$ $(\Longrightarrow \psi(\lambda^k, \rho^k) = L(p^k, \lambda^k, \rho^k))$
  $$p^k := \mathcal{A}^{ex}(\lambda^k, \rho^k)$$
  $$p^k := \mathcal{A}^{inex}(\lambda^k, \rho^k) \approx \min_p L(p, \lambda^k, \rho^k)$$
  with an error $E^k := L(p^k, \lambda^k, \rho^k) - \psi(\lambda^k, \rho^k) \in [0, \eta]$
  Error is unknown, η bounds approximation inaccuracy

- **Dual Step:** $(\lambda^{k+1}, \rho^{k+1})$ one iteration to solve $\max_{\lambda,\rho} \psi(\lambda, \rho)$
  Subgradient (Uzawa), Cutting-planes (DW), **Bundle**

When is $\mathcal{A}^{inex}$ sufficiently good?

Primal points may not be sufficiently good

GALs: difficulties

- **Primal Step:** Given (λ^k, ρ^k) current multipliers,
  Instead of p^k solves $\min_p L(p, \lambda^k, \rho^k)$ $\quad (\Longrightarrow \psi(\lambda^k, \rho^k) = L(p^k, \lambda^k, \rho^k))$
  $$p^k := \mathcal{A}^{ex}(\lambda^k, \rho^k)$$
  $$p^k := \mathcal{A}^{inex}(\lambda^k, \rho^k) \approx \min_p L(p, \lambda^k, \rho^k)$$
  with an error $E^k := L(p^k, \lambda^k, \rho^k) - \psi(\lambda^k, \rho^k) \in [0, \eta]$
  Error is unknown, η bounds approximation inaccuracy

- **Dual Step:** $(\lambda^{k+1}, \rho^{k+1})$ one iteration to solve $\max_{\lambda,\rho} \psi(\lambda, \rho)$
  Subgradient (Uzawa), Cutting-planes (DW), **Bundle**

When is $\mathcal{A}^{inex}$ sufficiently good?

Primal points may not be sufficiently good

When to stop?

GALs: primal-dual bundle scheme

- **Primal Step:** Given $(\hat{\lambda}^k, \hat{\rho}^k)$ current good multipliers,
  $p^k := \mathcal{A}^{inex}(\hat{\lambda}^k, \hat{\rho}^k) \approx \min_p L(p, \hat{\lambda}^k, \hat{\rho}^k) = \psi(\hat{\lambda}^k, \hat{\rho}^k)$
- **Dual Step:** $(\lambda^{k+1}, \rho^{k+1})$ one iteration to solve $\max_{\lambda,\rho} \psi(\lambda, \rho)$

  Bundle of p^i's: defines a simple QP, with
  - a matrix Γ^F with $|B_k|$ columns $F(p^i) \in \mathbb{R}^m$
  - Γ^σ with $|B_k|$ columns $\sigma(F(p^i)) \in \mathbb{R}$

  $$\min\left\{\frac{1}{2t_k^F}\left\langle \Gamma^F\alpha, \Gamma^F\alpha\right\rangle + \frac{1}{2t_k^\sigma}\left\langle \Gamma^\sigma\alpha, \Gamma^\sigma\alpha\right\rangle + \langle q, \alpha\rangle : \alpha \in \Delta\right\}$$

  gives $\lambda^{k+1} = \hat{\lambda}^k + t_k^F \Gamma^F \alpha^k$ and $\rho^{k+1} = \hat{\rho}^k + t_k^\sigma \Gamma^\sigma \alpha^k$
- **Goodness:** New primal point $p^{k+1} := \mathcal{A}^{inex}(\lambda^{k+1}, \rho^{k+1})$ is good if it gives a larger Lagrangian value than $p^k = \mathcal{A}^{inex}(\hat{\lambda}^k, \hat{\rho}^k)$
- **Stopping test:** checks, up to η,
  - primal feasibility
  - optimality gap

GALs: primal-dual bundle scheme

Convergence within the error η of $\mathcal{A}^{inex}$

Theorem (Primal-dual convergence)

For the subsequence $\{p^k, \hat{\lambda}^k, \hat{\rho}^k\}$

- all cluster points (if any) of $\{\hat{\lambda}^k, \hat{\rho}^k\}$ are dual η-solutions
- all cluster points of p^k are primal η-solutions
- the optimality gap eventually vanishes

GALs: primal-dual bundle scheme

Convergence within the error η of $\mathcal{A}^{inex}$

Theorem (Primal-dual convergence)

For the subsequence $\{p^k, \hat{\lambda}^k, \hat{\rho}^k\}$

- all cluster points (if any) of $\{\hat{\lambda}^k, \hat{\rho}^k\}$ are dual η-solutions
- all cluster points of p^k are primal η-solutions
- the optimality gap eventually vanishes

Versatile

- DC programming
- Sparse optimization
- UC

DC programming: $\mathcal{A}^{ex}$

Global solution to

$$\begin{cases} \min & \frac{1}{2}\langle p, Mp\rangle + \langle q, p\rangle - \max_{1\le i\le N}\{\langle \alpha_i, p\rangle + \beta_i\} \\ \text{s.t.} & Ap = b \end{cases}$$

found by solving N QP's

$$\begin{cases} \min & \frac{1}{2}\langle p, Mp\rangle + \langle q, p\rangle - \langle \alpha_i, p\rangle - \beta_i \\ \text{s.t.} & Ap = b\,, \end{cases}$$

Alternative: Proximal Lagrangian with $\sigma(F(p)) = \frac{1}{2}\|Ap - b\|^2$, with

$$\psi(\lambda, \rho) = \min_{1\le i\le N}\left(\min \frac{1}{2}\langle p, Mp\rangle + \langle q - \alpha_i, p\rangle - \beta_i + \langle \lambda, Ap - b\rangle + \frac{\rho}{2}\|Ap - b\|^2\right)$$

- **Primal Step:** $p^k := \mathcal{A}^{ex}(\hat{\lambda}^k, \hat{\rho}^k) = \min_p L(p, \hat{\lambda}^k, \hat{\rho}^k)$

  Solve N LP's: $0 = Mp + q - \alpha_i + A^\top\lambda + \rho A^\top(Ap - b)$ for $1 \le i \le N$
- **Dual Step:** Solve 1 QP to find $(\lambda^{k+1}, \rho^{k+1})$

Unit Commitment: $\mathcal{A}^{inex}$

Sharp Lagrangian

$L(p,\lambda,\rho) = L_T(x,z_T,\lambda) + L_{HT}(y_T,y_H,\lambda) + \langle\lambda,d\rangle + \rho\|y_T - z_T\|$

- **Primal Step:** ADMM-like inexact minimization

$$\mathcal{A}^{inex}(\hat{\lambda}^k,\hat{\rho}^k)\begin{cases} \min\limits_{x,z_T} L_T(x,z_T,\hat{\lambda}^k) + \dfrac{\hat{\rho}^k}{2}\|y_T^{k-1} - z_T\| \\ \min\limits_{y_T,y_H} L_{HT}(y_T,y_H,\hat{\lambda}^k) + \dfrac{\hat{\rho}^k}{2}\|y_T - z_T^{k-1}\| \\ + \left\langle \hat{\lambda}^k, d \right\rangle \end{cases}$$

- **Dual Step:** $(\lambda^{k+1},\rho^{k+1})$ defined using QP solution α_k

Numerical Assessment

- UniToy OK!
- On progress for real-life instances