

Nested Benders Decomposition

Prof. Alexandre Salles da Cunha

Universidade Federal de Minas Gerais
Departamento de Ciência da Computação
Belo Horizonte, Brasil

`acunha@dcc.ufmg.br`

2021/2



UNIVERSIDADE FEDERAL
DE MINAS GERAIS



Relativas à esta apresentação

- 1 J. R. Birge & F. Louveaux. Introduction to Stochastic Programming, Springer, 2011, 2a Edição [Seção 6.1]

Variáveis de decisão contínuas

- H estágios
- x^1 : variável de decisão de 1o. estágio.
- $x^t(\omega^t)$: $t = 2, \dots, H$ variável de decisão do estágio t , dependendo da realização ω^t .
- $\xi^t(\omega)^T = (c^t(\omega)^T, h^t(\omega)^T, T_1^{t-1}(\omega), \dots, T_{m_t}^{t-1}(\omega))$ é um vetor N_t dimensional. $\xi^t(\omega)$ é independente de $\xi^{t-1}(\omega), \dots, \xi^1(\omega)$.
- $T^{t-1}(\omega^t)$ é uma matriz $m_t \times n_{t-1}$ dimensional.
- W^t é uma matriz $m_t \times n_t$ dimensional, fixa, independented de ξ .
- As decisões x dependem da história até o instante t , que indicamos por ω^t .
- Ξ^t é o suporte finito de ξ^t .
- No modelo seguinte, a notação transposta foi intencionalmente suprimida.

$$\min c^1 x^1 + \mathbb{E}_{\xi^2} \left[\min c^2(\omega^2) x^2(\omega^2) + \dots + \mathbb{E}_{\xi^H} \left[\min c^H(\omega^H) x^H(\omega^H) \right] \dots \right]$$

$$h^1 = W^1 x^1$$

$$h^2(\omega^2) = T^1(\omega^2) x^1 + W^2 x^2(\omega^2)$$

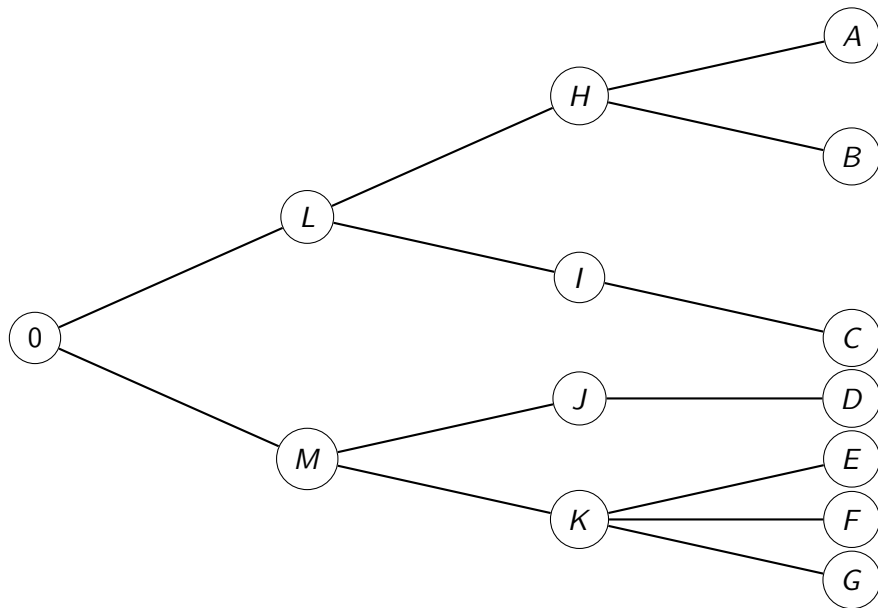
$$\dots \vdots$$

$$h^H(\omega^H) = T^{H-1}(\omega^H) x^{H-1} + W^H x^H(\omega^H)$$

$$x^1 \geq 0$$

$$x^t(\omega^t) \geq 0, \quad t = 2, \dots, H.$$

Árvore de cenários com 4 estágios e 7 cenários



$t = H$:

$$\begin{aligned}Q^H(x^{H-1}, \xi^H(\omega)) &= \min c^H(\omega) x^H(\omega) \\ W^H x^H(\omega) &= h^H(\omega) - T^{H-1}(\omega) x^{H-1} \\ x^H(\omega) &\geq 0\end{aligned}$$

Para $t = 1, \dots, H-1$

$$\begin{aligned}Q^t(x^{t-1}, \xi^t(\omega)) &= \min c^t(\omega) x^t(\omega) + Q^{t+1}(x^t) \\ W^t x^t(\omega) &= h^t(\omega) - T^{t-1}(\omega) x^{t-1} \\ x^t(\omega) &\geq 0\end{aligned}$$

onde $Q^{t+1}(x^t) = \mathbb{E}_{\xi^{t+1}} [Q^{t+1}(x^t, \xi^{t+1}(\omega))]$

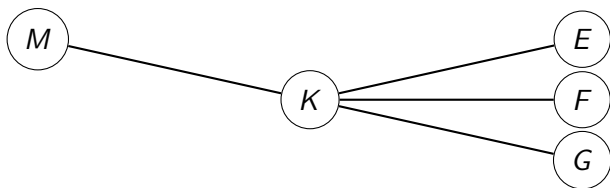
Problema estocástico linear multi-estágio

$$\min z = \min c^1 x^1 + Q^2(x^1) \quad (1)$$

$$W^1 x^1(\omega) = h^1$$

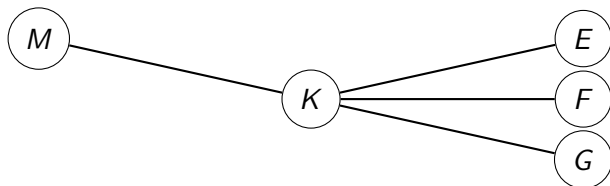
$$x^1 \geq 0$$

- $\mathcal{K}^t$: conjunto de cenários em $t \in \{1, \dots, H\}$.
- $p_k^t > 0$: probabilidade de ocorrência do cenário $k \in \mathcal{K}^t$.
 $\sum_{k \in \mathcal{K}^t} p_k^t = 1$.
- Para um par k, t , $a(k) \in \mathcal{K}^{t-1}$ denota o antecessor de k, t na árvore de cenários.
- para j no período t , $\mathcal{D}^t(j)$ denota os cenários sucessores de j no período $t + 1$.
- NLSD(t, k): subproblema a ser resolvido em no tempo t e cenário $k \in \mathcal{K}^t$.



- $a(K) = M$
- $\mathcal{D}^{H-1}(K) = \{E, F, G\}$ - descendentes do cenário K (que é um cenário em $H - 1$) no estágio imediatamente posterior H .
- Para $t = H$ temos $\mathcal{K}^H = \{A, B, C, D, E, F, G\}$

- Como tratamos do caso de suporte finito, vamos formular um problema de otimização NLSD(t, k) para cada nó da árvore de cenários.
- Formular uma outer approximation para cada nó da árvore, substituindo a função recurso em cada nó $k \in \mathcal{K}^t$, por uma aproximação θ_k^t e inserindo cortes de Benders.
- Em cada problema NLSD(t, k):
 - Uma realização (ω_k^t) da incerteza em t “aconteceu”
 - Decisão do antecessor na árvore: $x_{a(k)}^{t-1} \geq 0$.
 - Decisão a ser tomada: $x_k^t \geq 0$.
- NLSD(t, k) serão alterados ao longo do algoritmo, através da introdução de planos de corte.

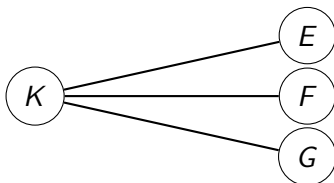


- 1 Deduzimos os cortes gerados para os subproblems $\text{NLSD}(H-1,k)$ na camada $t = H - 1$, a partir dos multiplicadores em $\text{NLSD}(H,j)$, para os sucessores j de k , pois em $t = H$ não há recurso a ser aproximado.
- 2 Sabemos conduzir a tarefa acima, pois na camada $H - 1$, temos um problema de dois estágios...
- 3 Então, generalizamos a ideia para os demais, pois os cortes em $\text{NLSD}(t,k)$ envolverão multiplicadores associados a cortes de Benders presentes em $\text{NLSD}(t+1,j)$ para os j sucessores de k na árvore de enumeração.

NLSD(H,k)

$$\begin{aligned} \min \quad & (c_k^H)^T x_k^H \\ \text{s.t.} \quad & W^H x_k^H = h_k^H - T_k^H x_{a(k)}^{H-1} \\ & x_k^H \geq 0 \end{aligned}$$

- Em $t = H$, não há função recurso a ser aproximada.
- Para qualquer $x_{a(k)}^{H-1}$ proposta pelo antecessor de H, k :
 - 1 Se $x_{a(k)}^{H-1}$ é factível em H, k , usamos os multiplicadores duais ótimos de NLSD(H,k) para formular um corte de otimalidade no antecessor NLSD(H-1,a(k)) e verificar sua eventual violação.
 - 2 Caso contrário, usamos os multiplicadores duais para formular um corte de viabilidade em NLSD(t-1,a(k)).



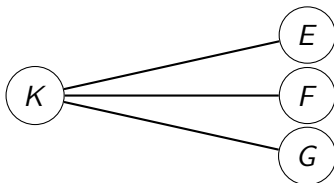
- Aproximação que empregamos é agregada, pois os cortes de otimalidades que serão gerados são agregados:

$p_k^t \theta_k^t$ aproxima a função $\sum_{j \in \mathcal{D}^t(k)} p_j^{t+1} Q(x_k^t, \xi(\omega_k^{t+1}))$

- Para compreender a dedução dos cortes podemos desagregar:

$$p_k^t \theta_k^t = \sum_{j \in \mathcal{D}^t(k)} p_j^{t+1} \theta_{k,j}^t$$

onde $\theta_{k,j}^t$ aproxima $Q(x_k^t, \xi(\omega_k^{t+1}))$.



$p_k^t \theta_k^t$ aproxima a função $\sum_{j \in \mathcal{D}^t(k)} p_j^{t+1} Q(x_k^t, \xi(\omega_k^{t+1}))$

NLSD(t, k)

$$\begin{aligned}
 \min \quad & (c_k^t)^T x_k^t + \theta_k^t \\
 \text{s.t.} \quad & W^t x_k^t = h_k^t - T_k^t x_{a(k)}^{t-1} \\
 & D_{k,j}^t x_k^t \geq d_{k,j}^t & j = 1, \dots, r_k^t \\
 & \theta_k^t + E_{k,j}^t x_k^t \geq e_{k,j}^t & j = 1, \dots, s_k^t \\
 & x_k^t \geq 0
 \end{aligned}$$

NLSD(t, k)

$$\begin{aligned}
 \min \quad & (c_k^t)^T x_k^t + \theta_k^t \\
 \text{s.t.} \quad & W^t x_k^t = h_k^t - T_k^t x_{a(k)}^{t-1} \\
 & D_{k,j}^t x_k^t \geq d_{k,j}^t & j = 1, \dots, r_k^t \\
 & \theta_k^t + E_{k,j}^t x_k^t \geq e_{k,j}^t & j = 1, \dots, s_k^t \\
 & x_k^t \geq 0
 \end{aligned}$$

Ideias essenciais:

- ($\rightarrow$) NLSD(t, k) propõe uma decisão x_k^t para seus sucessores, todos no estágio posterior $t + 1$ da árvore de cenários.
- ($\leftarrow$) NLSD(t, k) recebe, de seus sucessores, multiplicadores duais para formular cortes de Benders e aprimorar a aproximação.
- ($\leftrightarrow$) O algoritmo *atravessa a árvore de cenários*, inicialmente da raiz para as folhas, e posteriormente, nas duas direções.

NLSD(H,k)

$$\begin{aligned}
 &\min (c_k^H)^T x_k^H \\
 &s.t. \quad W^H x_k^H = h_k^H - T_k^H \bar{x}_{a(k)}^{H-1} \quad \rightarrow \quad \text{raio extremo } \bar{\pi}_k^H \\
 &\quad x_k^H \geq 0
 \end{aligned}$$

- Se é **inviável** para $\bar{x}_{a(k)}^{H-1}$ formulamos o corte:

$$(\bar{\pi}_k^H)^T h_k^H \leq (\bar{\pi}_k^H)^T T_k^H x_{a(k)}^{H-1}$$

- Usamos o contador de cortes de viabilidade $r_{a(k)}^{H-1}$
- Definimos $d_{a(k), r_{a(k)}}^{H-1} = (\bar{\pi}_k^H)^T h_k^H$ e $D_{a(k), r_{a(k)}}^{H-1} = (\bar{\pi}_k^H)^T T_k^H$.
- O corte $d_{a(k), r_{a(k)}}^{H-1} \leq D_{a(k), r_{a(k)}}^{H-1} x_{a(k)}^{H-1}$ pode ser enviado backwards, para NLSD(H-1,a(k)).

NLSD(H,k)

$$\begin{aligned}
 & \min (c_k^H)^T x_k^H \\
 & \text{s.t. } W^H x_k^H = h_k^H - T_k^H \bar{x}_{a(k)}^{H-1} \quad \rightarrow \quad \text{ponto extremo } \bar{\pi}_k^H \\
 & \quad x_k^H \geq 0
 \end{aligned}$$

- Se viável para $\bar{x}_{a(k)}^{H-1}$ formulamos o corte desagregado:

$$(\bar{\pi}_k^H)^T h_k^H \leq (\bar{\pi}_k^H)^T T_k^H x_{a(k)}^{H-1} + \theta_{a(k),k}^{H-1}$$

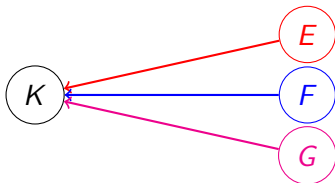
ou, multiplicando por p_k^H :

$$(p_k^H \bar{\pi}_k^H)^T h_k^H \leq (p_k^H \bar{\pi}_k^H)^T T_k^H x_{a(k)}^{H-1} + p_k^H \theta_{a(k),k}^{H-1}$$

que equivale a minimizar $(p_k^H c_k^H)^T x_k^H$ e re-escalar os multiplicadores duais pelo fator p_k^H .

Supondo que E, F, G são viáveis com a solução proposta $\bar{x}_K^{H-1}$

Particularizando....



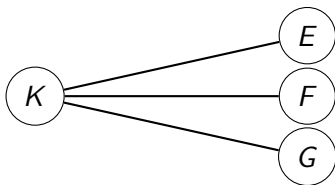
$$(p_E^H \bar{\pi}_E^H)^T h_E^H \leq (p_E^H \bar{\pi}_E^H)^T T_E^H x_K^{H-1} + p_E^H \theta_{K,E}^{H-1}$$

$$(p_F^H \bar{\pi}_F^H)^T h_F^H \leq (p_F^H \bar{\pi}_F^H)^T T_F^H x_K^{H-1} + p_F^H \theta_{K,F}^{H-1}$$

$$(p_G^H \bar{\pi}_G^H)^T h_G^H \leq (p_G^H \bar{\pi}_G^H)^T T_G^H x_K^{H-1} + p_G^H \theta_{K,G}^{H-1}$$

Somando, substituindo $p_K^{H-1} \theta_K^{H-1}$ e dividindo por p_K^{H-1} :

$$\sum_{u \in \{E, F, G\}} \left(\frac{p_u^H}{p_K^{H-1}} (\bar{\pi}_u^H)^T \right) h_u^H \leq \sum_{u \in \{E, F, G\}} \left(\frac{p_u^H}{p_K^{H-1}} (\bar{\pi}_u^H)^T T_u^H \right) x_K^{H-1} + \theta_K^{H-1}$$



$$\sum_{u \in \{E, F, G\}} \left(\frac{p_u^H}{p_K^{H-1}} (\bar{\pi}_u^H)^T \right) h_u^H \leq \sum_{u \in \{E, F, G\}} \left(\frac{p_u^H}{p_K^{H-1}} (\bar{\pi}_u^H)^T T_u^H \right) x_K^{H-1} + \theta_K^{H-1}$$

$$\sum_{u \in \mathcal{D}^{H-1}(K)} \left(\frac{p_u^H}{p_K^{H-1}} (\bar{\pi}_u^H)^T \right) h_u^H \leq \sum_{u \in \mathcal{D}^{H-1}(K)} \left(\frac{p_u^H}{p_K^{H-1}} (\bar{\pi}_u^H)^T T_u^H \right) x_K^{H-1} + \theta_K^{H-1}$$

que, para algum indexador j de cortes pode ser escrita de forma compacta como:

$$e_{K,j} \leq E_{K,j} x_K^{H-1} + \theta_K^{H-1}$$

NLSD(t,k)

$$\min (c_k^t)^T x_k^t + \theta_k^t \quad (2)$$

$$\text{s.t. } W^t x_k^t = h_k^t - T_k^t x_{a(k)}^{t-1} \quad \rightarrow \pi_k^t \quad (3)$$

$$D_{k,j}^t x_k^t \geq d_{k,j}^t \quad j = 1, \dots, r_k^t \quad \rightarrow \rho_{k,j}^t \geq 0 \quad (4)$$

$$\theta_k^t + E_{k,j}^t x_k^t \geq e_{k,j}^t \quad j = 1, \dots, s_k^t \quad \rightarrow \sigma_{k,j}^t \geq 0 \quad (5)$$

$$x_k^t \geq 0 \quad (6)$$

Dual de NLSD(t,k) - DNLSD(t,k)

$$\max (\pi_k^t)^T (h_k^t - T_k^t x_{a(k)}^{t-1}) + \sum_j (\rho_{k,j}^t)^T d_{k,j}^t + \sum_j (\sigma_{k,j}^t)^T e_{k,j}^t$$

$$\text{s.t. } (\pi_k^t)^T W^t + \sum_j (\rho_{k,j}^t)^T D_{k,j}^t + \sum_j (\sigma_{k,j}^t)^T E_{k,j}^t \leq (c_k^t)^T$$

$$\sum_j \sigma_{k,j}^t \leq 1$$

$$\rho_{k,j}^t \geq 0 \quad j = 1, \dots, r_k^t$$

$$\sigma_{k,j}^t \geq 0 \quad j = 1, \dots, s_k^t$$

No slide anterior quando somamos em $\rho_{k,j}^t$ e $\sigma_{k,j}^t$

$$\sum_j \rho_{k,j}^t = \sum_{j=1}^{r_k^t} \rho_{k,j}^t$$

$$\sum_j \sigma_{k,j}^t = \sum_{j=1}^{s_k^t} \sigma_{k,j}^t$$

Observação para simplificar o corte de viabilidade

Observe que se existir x_k^t que satisfaça as restrições (3) e (4), haverá um θ_k^t tal que (x_k^t, θ_k^t) satisfaz (5) (basta tomar θ_k^t suficientemente grande)

NLSD(t,k)

$$\begin{aligned} \min \quad & (c_k^t)^T x_k^t + \theta_k^t \\ \text{s.t.} \quad & W^t x_k^t = h_k^t - T_k^t x_{a(k)}^{t-1} \end{aligned} \tag{3}$$

$$D_{k,j}^t x_k^t \geq d_{k,j}^t \quad j = 1, \dots, r_k^t \tag{4}$$

$$\theta_k^t + E_{k,j}^t x_k^t \geq e_{k,j}^t \quad j = 1, \dots, s_k^t \tag{5}$$

$$x_k^t \geq 0$$

Então, se $\text{NLSD}(t,k)$ é inviável temos para algum vetor dual-viável $(\bar{\pi}_k^t, \{\bar{\rho}_{k,j}^t : j = 1, \dots, r_k^t\})$:

$$(\bar{\pi}_k^t)^T W^t + \sum_j (\bar{\rho}_{k,j}^t)^T D_{k,j}^t \leq 0 \quad (7)$$

e

$$(\bar{\pi}_k^t)^T (h_k^t - T_k^t \bar{x}_{a(k)}^{t-1}) + \sum_j (\bar{\rho}_{k,j}^t)^T d_{k,j}^t > 0 \quad (8)$$

caracterizando a ilimitação de $\text{DNLSD}(t,k)$.

O corte de viabilidade

$$(\bar{\pi}_k^t)^T (h_k^t - T_k^t x_{a(k)}^{t-1}) + \sum_j (\bar{\rho}_{k,j}^t)^T d_{k,j}^t \leq 0 \quad (9)$$

pode ser inserido em $\text{NLSD}(t-1,a(k))$.

O corte de viabilidade

$$(\bar{\pi}_k^t)^T (h_k^t - T_k^t x_{a(k)}^{t-1}) + \sum_j (\bar{\rho}_{k,j}^t)^T d_{k,j}^t \leq 0 \quad (9)$$

pode ser escrito na forma

$$D_{a(k), r_{a(k)}^{t-1}}^{t-1} x_{a(k)}^{t-1} \geq d_{a(k), r_{a(k)}^{t-1}}^{t-1} \quad (10)$$

para:

$$D_{a(k), r_{a(k)}^{t-1}}^{t-1} = (\bar{\pi}_k^t)^T T_k^t \text{ e}$$

$$d_{a(k), r_{a(k)}^{t-1}}^{t-1} = (\bar{\pi}_k^t)^T h_k^t + \sum_j (\bar{\rho}_{k,j}^t)^T d_{k,j}^t$$

- Com a solução primal-dual ótima do par NLSD(t,k) e DNLSD(t,k): $\bar{x}_k^t, \bar{\pi}_k^t, \{\bar{\rho}_{k,j}^t : j = 1, \dots, r_k^t\}, \{\bar{\sigma}_{k,j}^t : j = 1, \dots, s_k^t\}$
- Formulamos um corte (desagregado, multiplicado por p_k^t)

$$p_k^t \theta_{a(k),k}^{t-1} \geq (p_k^t \bar{\pi}_k^t)^T (h_k^t - T_k^t x_{a(k)}^{t-1}) + \sum_j (p_k^t \bar{\rho}_{k,j}^t)^T d_{k,j}^t + \sum_j (p_k^t \bar{\sigma}_{k,j}^t)^T e_{k,j}^t$$

- Caso todos os sucessores de $a(k)$, $t-1$ em t sejam viáveis, podemos verificar se o corte de otimalidade agregado (11) é violado pela solução de NLSD(t-1,a(k))

$$\theta_{a(k)}^{t-1} \geq \frac{1}{p_{a(k)}^{t-1}} \sum_{k \in \mathcal{D}^{t-1}(a(k))} \left[(p_k^t \bar{\pi}_k^t)^T (h_k^t - T_k^t x_{a(k)}^{t-1}) + \sum_j (p_k^t \bar{\rho}_{k,j}^t)^T d_{k,j}^t + \sum_j (p_k^t \bar{\sigma}_{k,j}^t)^T e_{k,j}^t \right] \quad (11)$$

ou, reorganizando de forma compacta:

$$\theta_{a(k)}^{t-1} + E_{a(k),r_{a(k)}^{t-1}}^{t-1} \geq e_{a(k),r_{a(k)}^{t-1}}^{t-1} \quad (12)$$

- ❶ Atravessa a árvore de cenários, possivelmente visitando cada par (k, t) muitas vezes.
- ❷ Dependendo da solução de $NLSD(t, k)$, percorre a árvore de cenários em uma de duas direções:
 - forward: propaga soluções viáveis do nó (k, t) para os nós vizinhos no estágio seguinte.
 - backward: retorna um ou mais cortes para o cenário pai, no estágio anterior.
- ❸ Respeitando as observações acima, há flexibilidade para conjugar as direções forward e backward, em um algoritmo correto.
- ❹ Vamos discutir uma implementação particular de como atravessar a árvore de cenários (fast-forward-fast-backward). Há outras.

Direção forward

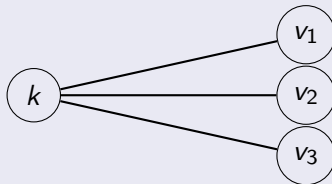
NLSD(t, k) é resolvido.

- Se é viável e atinge ótimo em $\bar{x}_k^t, \bar{\theta}_k^t$, a solução ótima

$$\bar{x}_k^t$$

pode ser propagada ($\rightarrow$) para todos seus descententes em $\mathcal{D}^t(k)$.

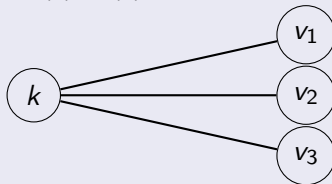
- Os multiplicadores duais ótimos associados à solução primal $\bar{x}_k^t, \bar{\theta}_k^t$ são armazenados para (em um passe backward, $\leftarrow$) formular cortes de otimalidade (agregados) para $(t-1, a(k))$.



Direção backward

NLSD(t, k) é resolvido.

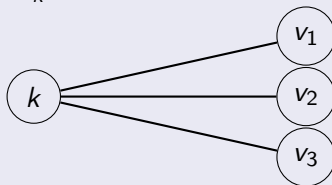
- Se NLSD(t, k) é inviável, não há solução a ser propagada.
Imediatamente usa os multiplicadores duais para fornecer cortes de viabilidade para NLSD($t-1, a(k)$), e então propor uma nova solução tentativa $\bar{x}_{a(k)}^{t-1}, \bar{\theta}_{a(k)}^{t-1}$.



Direção backward

NLSD(t, k) é resolvido.

- Se NLSD(t, k) é viável, esperamos todos os subproblemas NLSD(t, z): $z \in \mathcal{K}^t : a(k) = a(z)$ serem resolvidos e constatados viáveis, para formular um corte de otimalidade agregado para NLSD($t-1, a(k)$).
- Poderíamos exportar um corte de otimalidade desagregado imediatamente (sem esperar os demais irmãos de k), mas o algoritmo que apresentamos emprega a versão agregada, por meio de variáveis θ_k^t agregadas.



Para ajudar a responder a questão, identificamos quando o algoritmo não pode parar:

- Sempre que um algum $NLSD(t,k)$ é inviável, há necessidade de mover-se backward, inserir um corte de viabilidade em $NLSD(t-1,a(k))$ e reotimizar $NLSD(t-1,a(k))$, para propagar nova solução tentativa.
- Sempre que, por meio de um movimento backward, retornamos a $NLSD(t,k)$ e verificamos que o corte de otimalidade agregado proposto por seus descententes em $t+1$ é violado, precisamos gerar nova solução tentativa que deve ser propagada.

- Retornarmos ao nó raiz ($t = 1, k = 1$) com a informação de que os subproblemas descententes foram viáveis com as soluções propostas desde a última mudança da direção de forward para backward e
- A solução atual $(\bar{x}_1^1, \bar{\theta}_1^1)$ não subestima o recurso $Q^2(\bar{x}_1^1)$, isto é, nenhum corte de otimalidade foi inserido em $t = 1$, desde a última vez que, a partir de $t = 1, k = 1$, implementamos a direção forward.

Passo 0:

❶ Inicialização

- ❶ Para todo $t \in \{1, \dots, H\}$ e $k \in \mathcal{K}^t$, faça $r_k^t = s_k^t = 0$ e insira a restrição $\theta_k^t = 0$ (será eliminada posteriormente).
 - ❷ $t \leftarrow 1, k \leftarrow 1$
 - ❸ Faça $\text{dir} = \text{FOR}$.
- ❷ Vá para o Passo 1.

Passo 1:

- ① Resolva o problema NLSD(t, k) corrente.
- ② Se NLSD(t, k) é inviável e $t == 1$, páre.
- ③ Caso contrário, se NLSD(t, k) é inviável e $t > 1$, faça:
 - ① $r_{a(k)}^{t-1} = r_{a(k)}^{t-1} + 1$.
 - ② $\text{dir} = \text{BACK}$.
 - ③ Com os multiplicadores $\pi_k^t, \{\rho_{k,j}^t \geq 0 : j = 1, \dots, r_k^t\}$ que garantem (7) e (8), **formule e insira um corte de viabilidade** (10) em NLSD($t-1, a(k)$).
 - ④ Faça $t = t - 1$, $k = a(k)$ e retorne ao Passo 1.
- ④ Se NLSD(t, k) admite ótimo finito:
 - ① Atualize x_k^t, θ_k^t e **se $t > 1$, armazene os multiplicadores** $\pi_k^t, \{\rho_{k,j}^t \geq 0 : j = 1, \dots, r_k^t\}, \{\sigma_{k,j}^t \geq 0 : j = 1, \dots, s_k^t\}$ duais ótimos.
 - ② Se $k < |\mathcal{K}^t|$, faça $k = k + 1$ e retorne ao Passo 1.
 - Caso contrário (i.e., $k == |\mathcal{K}^t|$):
 - ① Se $t == 1$, faça $\text{dir} = \text{FOR}$.
 - ② Caso contrário: a) Se ($t < H$ e $\text{dir} = \text{FOR}$), faça $t = t + 1$, $k = 1$, retorne ao Passo 1.
b) Caso contrário, se $t == H$, faça $\text{dir} = \text{BACK}$.
 - ③ Vá para o Passo 2.

Passo 2:

- ① Se $t == 1$, faça $t = t + 1, k = 1$ e vá para o Passo 1.
- ② Caso contrário, para todo cenário $j \in \mathcal{K}^{t-1}$:
 - ① Calcule $E_j^{t-1}, e_j^{t-1}, \bar{\theta}_j^{t-1} = e_j^{t-1} - E_j^{t-1}x_j^{t-1}$ (x_j^{t-1} foi armazenado antes).
 - ② Se a restrição $\theta_j^{t-1} = 0$ ainda está presente em NLSD(t-1,j), remova-a, faça $s_j^{t-1} = 1$ e insira um corte de otimalidade (12) com E_j^{t-1}, e_j^{t-1} em NLSD(t-1,j). Caso contrário, (i.e., restrição foi eliminada):
 - ① Se $\bar{\theta}_j^{t-1} > \theta_j^{t-1}$ (θ_j^{t-1} também foi armazenada antes), faça: $s_j^{t-1} = s_j^{t-1} + 1$, insira um corte (12) com E_j^{t-1}, e_j^{t-1} em NLSD(t-1,j).
 - ③ Se $t = 2$ e nenhuma restrição foi inserida em NLSD(1,1), páre, x_1^1 é a solução ótima.
 - ④ Caso contrário, faça $t = t - 1, k = 1$.
 - ⑤ Se $t == 1$, faça dir = FOR.
 - ⑥ Vá para o Passo 1.

- 1 Deseja-se planejar a produção de aparelhos de ar condicionado, por três períodos.
- 2 Tomada a decisão de produção, o produto torna-se disponível após 1 mês.
- 3 A demanda ao final do primeiro mês é conhecida: 100 unidades. Ao final dos dois meses seguintes pode ser 100 ou 300, com igual probabilidade.
- 4 Capacidade de produção nominal (sem hora extra): 200 unidades por mês, ao custo de \$100. Pagando hora extra, é possível produzir ao custo de \$300, sem limite de capacidade.
- 5 Pode-se estocar produto de um mês para o seguinte ao custo de \$50 por unidade.
- 6 A demanda precisa ser atendida.
- 7 Ao final dos 3 meses, não há valor de recuperação do estoque residual.

Parâmetros

- $H = \{1, 2, 3\}$, $\mathcal{K}^1 = \{1\}$, $\mathcal{K}^2 = \{1, 2\}$, $\mathcal{K}^3 = \{1, 2, 3, 4\}$
- d_k^t denota a demanda no cenário $k \in \mathcal{K}^t$.
 $d_1^2 = 1$, $d_2^2 = 3$, $d^3 = (1, 3, 1, 3)^T$.

Variáveis

- $x_k^t \geq 0$: quantidade a produzir em t , no cenário k , sem hora extra.
- $w_k^t \geq 0$: quantidade a produzir em t , no cenário k , com hora extra.
- $y_k^t \geq 0$: quantidade estocada de t para $t + 1$.

$$\min x_1^1 + 3w_1^1 + 0.5y_1^1 + \sum_{k=1}^2 p_k^2(x_k^2 + 3w_k^2 + 0.5y_k^2) +$$

$$\sum_{k=1}^4 p_k^3(x_k^3 + 3w_k^3)$$

$$\text{s.t. } x_1^1 \leq 2$$

$$x_1^1 + w_1^1 - y_1^1 = 1$$

$$y_1^1 + x_k^2 + w_k^2 - y_k^2 = d_k^2 \quad k = 1, 2$$

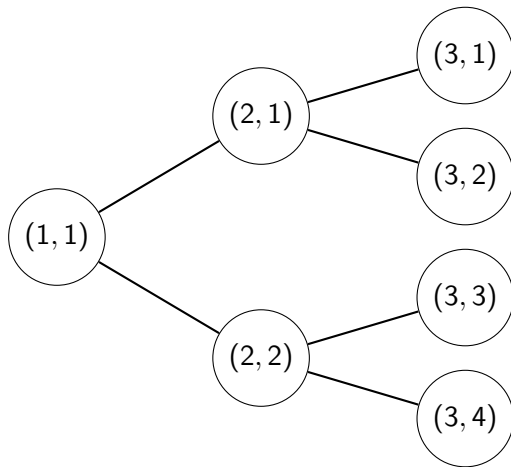
$$x_k^2 \leq 2 \quad k = 1, 2$$

$$y_{a(k)}^2 + x_k^2 + w_k^3 - y_k^3 = d_k^3 \quad k = 1, \dots, 4$$

$$x_k^3 \leq 2 \quad k = 1, \dots, 4$$

$$y_k^t, w_k^t, x_k^t \geq 0 \quad t = 1, \dots, 3, k \in \mathcal{K}^t$$

Vamos contar as iterações do algoritmo quando a dir mudar.



Passo 1: *NLSD*(1, 1)

$$\begin{aligned} \min \quad & x_1^1 + 3w_1^1 + 0.5y_1^1 + \theta_1^1 \\ \text{s.t.} \quad & x_1^1 \leq 2 \\ & x_1^1 + w_1^1 - y_1^1 = 1 \\ & \theta_1^1 = 0 \\ & x_1^1, y_1^1, w_1^1 \geq 0 \end{aligned}$$

Solução ótima $x_1^1 = 1$, todas demais primais nulas.

Passo 1: $NLSD(2, 1)$

$$\begin{aligned} \min \quad & x_1^2 + 3w_1^2 + 0.5y_1^2 + \theta_1^2 \\ \text{s.t.} \quad & x_1^2 \leq 2 \\ & x_1^2 + w_1^2 - y_1^2 = 1 \\ & \theta_1^2 = 0 \\ & x_1^2, y_1^2, w_1^2 \geq 0 \end{aligned}$$

Solução ótima $x_1^2 = 1$, todas demais primais nulas.

Passo 1: *NLSD*(2, 2) $\min x_2^2 + 3w_2^2 + 0.5y_2^2 + \theta_2^2$
 $s.t. \ x_2^2 \leq 2$
 $x_2^2 + w_2^2 - y_2^2 = 3$
 $\theta_2^2 = 0$
 $x_2^2, y_2^2, w_2^2 \geq 0$

Solução ótima $x_2^2 = 2, w_2^2 = 1$, todas demais primais nulas.

$$\begin{aligned}\text{Passo 1: } NLSD(3, 1) \quad & \min \quad x_1^3 + 3w_1^3 + \theta_1^3 \\ & \text{s.t.} \quad x_1^3 \leq 2 \\ & \quad \quad x_1^3 + w_1^3 - y_1^3 = 1 \\ & \quad \quad \theta_1^3 = 0 \\ & \quad \quad x_1^3, w_1^3, y_1^3 \geq 0\end{aligned}$$

Solução ótima $x_1^3 = 1$, todas demais primais nulas, $\pi_1^3 = (0, 1)^T$.

$$\begin{aligned}\text{Passo 1: } NLSD(3, 2) : \quad & \min \quad x_2^3 + 3w_2^3 + \theta_2^3 \\ & \text{s.t.} \quad x_2^3 \leq 2 \\ & \quad \quad x_2^3 + w_2^3 - y_2^3 = 3 \\ & \quad \quad \theta_2^3 = 0 \\ & \quad \quad x_2^3, w_2^3, y_2^3 \geq 0\end{aligned}$$

Solução ótima $x_2^3 = 2, w_2^3 = 1$, todas demais primais nulas, $\pi_2^3 = (-2, 3)^T$.

$$\begin{aligned}\text{Passo 1: } NLSD(3, 3) \quad & \min \quad x_3^3 + 3w_3^3 + \theta_3^3 \\ & \text{s.t. } x_3^3 \leq 2 \\ & \quad x_3^3 + w_3^3 - y_3^3 = 1 \\ & \quad \theta_3^3 = 0 \\ & \quad x_3^3, w_3^3, y_3^3 \geq 0\end{aligned}$$

Solução ótima $x_3^3 = 1$, todas demais primais nulas, $\pi_3^3 = (0, 1)^T$.

$$\begin{aligned}\text{Passo 1: } NLSD(3, 4) : \quad & \min \quad x_4^3 + 3w_4^3 + \theta_4^3 \\ & \text{s.t. } x_4^3 \leq 2 \\ & \quad x_4^3 + w_4^3 - y_4^3 = 3 \\ & \quad \theta_4^3 = 0 \\ & \quad x_4^3, w_4^3, y_4^3 \geq 0\end{aligned}$$

Solução ótima $x_4^3 = 2, w_4^3 = 1$, todas demais primais nulas, $\pi_3^3 = (-2, 3)^T$.

dir = BACK

Passo 2: Para cenário $k = 1, t - 1 = 2$, temos:

$$\begin{aligned} E_{11}^2 &= \frac{0.25}{0.5} ((\pi_1^3)^T T_1^3 + (\pi_2^3)^T T_2^3) \\ &= (0.5) \left(\begin{pmatrix} 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \begin{pmatrix} -2 & 3 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \right) \\ &= \begin{pmatrix} 0 & 0 & 2 \end{pmatrix} \end{aligned}$$

$$\begin{aligned} e_{11}^2 &= \frac{0.25}{0.5} ((\pi_1^3)^T h_1^3 + (\pi_2^3)^T h_2^3) \\ &= (0.5) \left(\begin{pmatrix} 0 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \end{pmatrix} + \begin{pmatrix} -2 & 3 \end{pmatrix} \begin{pmatrix} 2 \\ 3 \end{pmatrix} \right) \\ &= 3 \end{aligned}$$

Corte violado (em 3 unidades) a ser inserido em NLS(2,1):

$$\theta_1^2 + 2y_1^2 \geq 3$$

Passo 2: Para cenário $k = 2, t - 1 = 2$, temos (ainda) condições simétricas ao caso $k = 1, t - 1 = 2$.

$$E_{21}^2 = (\begin{array}{ccc} 0 & 0 & 2 \end{array}), \quad e_{21}^3 = 3$$

Inserimos o corte $2y_2^2 + \theta_2^2 \geq 3$ em NLSD(2,2).

Passo 1: *NLSD*(2, 1) $\min x_1^2 + 3w_1^2 + 0.5y_1^2 + \theta_1^2$

$$\text{s.t. } x_1^2 \leq 2$$

$$x_1^2 + w_1^2 - y_1^2 = 1$$

$$2y_1^2 + \theta_1^2 \geq 3$$

$$x_1^2, y_1^2, w_1^2 \geq 0$$

Solução ótima $x_1^2 = 2, y_1^2 = 1, \theta_1^2 = 1, w_1^2 = 0, \pi_1^2 = (-0.5 \ 1.5)^T, \sigma_{11}^2 = 1$.

Passo 1: *NLSD*(2, 2) $\min x_2^2 + 3w_2^2 + 0.5y_2^2 + \theta_2^2$
 $s.t. \ x_2^2 \leq 2$
 $x_2^2 + w_2^2 - y_2^2 = 3$
 $2y_2^2 + \theta_2^2 \geq 3$
 $x_2^2, y_2^2, w_2^2 \geq 0$

Solução ótima $x_2^2 = 2, y_2^2 = 0, \theta_2^2 = 3, w_2^2 = 1, \pi_2^2 = (-2 \ 3)^T, \sigma_{21}^2 = 1$.

Continuamos com $\text{dir} = \text{BACK}$, para Passo 2.

Passo 2: Para cenário $k = 1, t - 1 = 1$, temos:

$$\begin{aligned}
 E_{11}^1 &= \frac{0.25}{0.5} ((\pi_1^2)^T T_1^2 + (\pi_2^2)^T T_2^2) \\
 &= (0.5) \left(\begin{pmatrix} -0.5 & 1.5 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \begin{pmatrix} -2 & 3 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \right) \\
 &= \begin{pmatrix} 0 & 0 & 2.25 \end{pmatrix} \\
 e_{11}^2 &= \frac{0.25}{0.5} ((\pi_1^2)^T h_1^2 + (\pi_2^2)^T h_2^2 + \sigma_{11}^2 e_{11}^2 + \sigma_{21}^2 e_{21}^2) \\
 &= (0.5) \left(\begin{pmatrix} -0.5 & 1.5 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \end{pmatrix} + \begin{pmatrix} -2 & 3 \end{pmatrix} \begin{pmatrix} 2 \\ 3 \end{pmatrix} + (1)(3) + (1)(3) \right) \\
 &= 5.75
 \end{aligned}$$

produzindo o corte de otimalidade

$$\theta_1^1 + 2.25y_1^1 \geq 5.75$$

a ser adicionado em NLSD(1,1).

Passo 1: $NLSD(1,1)$

$$\begin{aligned} \min \quad & x_1^1 + 3w_1^1 + 0.5y_1^1 + \theta_1^1 \\ \text{s.t.} \quad & x_1^1 \leq 2 \\ & x_1^1 + w_1^1 - y_1^1 = 1 \\ & \theta_1^1 + 2.25y_1^1 \geq 5.75 \\ & x_1^1, y_1^1, w_1^1 \geq 0 \end{aligned}$$

Solução ótima $x_1^1 = 2, y_1^1 = 1, w_1^1 = 0, \theta_1^1 = 3.5$. $\text{dir} = \text{FOR}$, incrementamos o contador das iterações.

- O processo persiste por um total de 6 iterações, correspondendo a 3 passes na direção FOR e 3 na direção BACK.
- Ao final, NLSD(1,1) é resolvido sem a necessidade de introdução de cortes adicionais, com a seguinte solução ótima:
 $\theta_1^1 = \bar{\theta}_1^1 = 3.75$, $(x_1^1, y_1^1)^* = (2, 1)$ com custo ótimo
 $z^* = 2.5 + 3.75 = 6.25$.

1 Primeiro corte gerado:

$$\theta_1^1 + 2.25y_1^1 \geq 5.75$$

Substituindo $y_1^1 = x_1^1 + w_1^1 - 1$ e definindo $prod^1 = x_1^1 + w_1^1$, obtemos

$$\theta_1^1 \geq 8 - 2.25(x_1^1 + w_1^1) = 8 - 2.25prod^1$$

- 2 A função objetivo é $z_1^1 = x_1^1 + 3w_1^1 + 0.5y_1^1 + \theta_1^1$. Substituindo para θ_1^1 , y_1^1 , reescrevendo em função de $prod^1$, no intervalo $x_1^1 \in [1, 2]$ ($prod^1 \geq 1$) temos:

$$z_1^1 = 7.5 + 1.5 \min\{prod^1, 2\} + 3.5 \max\{prod^1 - 2, 0\} - 2.25prod^1$$

que pode ser re-escrita como:

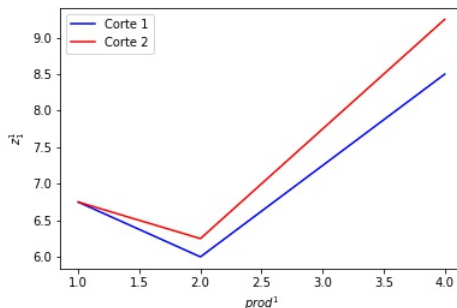
$$z_1^1(prod^1) = \begin{cases} 7.5 & - & 0.75prod^1 & \text{Se } prod^1 \leq 2 \\ 3.5 & + & 1.25prod^1 & \text{Se } prod^1 > 2 \end{cases}$$

Segundo corte para NLSD(1,1), gerado na iteração 4, (Ver Exerc. 2):

$$\theta_1^1 + 2x_1^1 + 2w_1^1 \geq 7.75$$

- Procedendo da mesma forma, obtemos:

$$z_1^1(prod^1) = \begin{cases} 7.25 - 0.5prod^1 & \text{Se } prod^1 \leq 2 \\ 3.25 + 1.5prod^1 & \text{Se } prod^1 > 2 \end{cases}$$



- ❶ O exemplo mostra que os cortes de otimalidade não necessariamente suportam a função recurso.
- ❷ Degeneração nas soluções ótimas de NLSD(t,k) agravam o problema. Múltiplas bases são ótimas para NLSD(t,k).
- ❸ Estratégia multi-cortes ajudam a melhorar a convergência.