

Programação Estocástica Inteira

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 [Seções 7.1, 7.2, 7.4 e 7.6]

(SIP) - Stochastic Integer Program

$$\min_{x \in X} c^T x + \min_y \mathbb{E}_{\xi} \left\{ q(\omega)^T y : W(\omega)y = h(\omega) - T(\omega)x, y \in Y \right\} \quad (1)$$

$$s.t. \quad Ax = b \quad (2)$$

Nos problemas que vamos discutir:

- O conjunto Y sempre envolverá restrições de integralidade.
- O conjunto X poderá ou não envolver restrições de integralidade.

Como de costume, a função objetivo de segundo estágio

$$Q(x, \xi) = \min_y \left\{ q(\omega)^T y : W(\omega)y = h(\omega) - T(\omega)x, y \in Y \right\}$$

e sua expectativa

$$\mathcal{Q}(x) = \mathbb{E}_{\xi} Q(x, \xi)$$

podem ser usadas para formular o problema determinístico equivalente (DEP)

$$\begin{aligned} & \min_{x \in X} c^T x + \mathcal{Q}(x) \\ & s.t. \quad Ax = b \end{aligned}$$

- Diferentemente do caso linear contínuo que foi estudado, $\mathcal{Q}(x)$ não apresenta nenhuma propriedade que *facilite* o desenvolvimento de algoritmos de solução.
- A forma de dualidade que empregamos no caso linear contínuo, por exemplo, não se aplica no caso inteiro.
- É bastante difícil encontrar um par dual forte para o problema de segundo estágio, para uma dada realização de $\hat{x} \in X$ e $\xi \in \Xi$.
- Tudo isso se traduz em dificuldades para escrever cortes fortes.
- Para um dado par \hat{x}, ξ , avaliar $Q(\hat{x}, \xi)$ requer a resolução de um programa inteiro, via de regra, um problema computacionalmente difícil. **Pouca ou nenhuma possibilidade de warm-start existe.**

Definimos:

- \bar{X} e \bar{Y} como as relaxações contínuas de X e de Y , respectivamente.
- A relaxação contínua do problema de segundo estágio $Q(x, \xi)$:

$$C(x, \xi) = \min_y \left\{ q(\omega)^T y : W(\omega)y = h(\omega) - T(\omega)x, y \in \bar{Y} \right\}$$

- e a expectativa da relaxação contínua do problema de 2o. estágio:

$$C(x) = \mathbb{E}_\xi C(x, \xi)$$

Observação:

Os cortes de otimalidade obtidos para a relaxação contínua de SIP são válidos para SIP.

- Para qualquer x, ξ , temos $C(x, \xi) \leq Q(x, \xi)$, uma vez que $X \subseteq \bar{X}, Y \subseteq \bar{Y}$.
- Logo, tomando as expectativas, $C(x) \leq Q(x)$.

- Muitos algoritmos dedicados à resolução exata de SIP **primeiro resolvem sua relaxação contínua** através do método de Decomposição de Benders (L-Shaped).
- Na sequência, **adotam alguma outra estratégia para tratar a integralidade:**
 - ① *Branching* (enumeração do conjunto X) **e/ou**
 - ② Reformulação dos cortes de optimalidade, para **considerar a integralidade** na dedução dos cortes.

Vamos assumir que a relaxação contínua tenha sido resolvida e disponhamos do problema atual (CP):

CP

$$\min c^T x + \theta$$

$$s.t. \quad Ax = b$$

$$D_I x \geq d_I \quad l = 1, \dots, r$$

$$E_I x + \theta \geq e_I \quad l = 1, \dots, s$$

$$x \geq 0$$

$$\theta \in \mathbb{R}$$

As estratégias branching e reformulação requerem resolver uma sequência de CPs, em um processo iterativo cuja finitude é garantida pela integralidade das variáveis.

- (Branch-and-bound) Particionamos o domínio X e resolvemos uma sequência de CPs, cada um definido para um destes subespaços do particionamento.
- Na reformulação, os cortes de otimalidade são re-escritos para se considerar a integralidade das variáveis x .
- Também aplicamos reformulação para melhor descrever a envoltória convexa de soluções viáveis de segundo estágio (para cada cenário).
- Enumeração dos problemas de segundo estágio também pode ser empregada.

- Assumimos existir um limite inferior L para o recurso

$$L \leq \min_x \{Q(x) : Ax = b, x \in X\}.$$

Este limite não precisa ser apertado, mas quanto mais apertado L , melhor o corte.

- Vamos assumir que $S(\hat{x}) \subseteq \{1, \dots, n_1\}$ denota o conjunto dos índices de variáveis de primeiro estágio que assumem valor 1 em uma solução de primeiro estágio, i.e., $S(\hat{x}) = \{i : \hat{x}_i = 1\}$ e $q_{S(\hat{x})} = Q(\hat{x})$, o valor correspondente da função recurso.
- Definimos a diferença simétrica de x relativa a $S(\hat{x})$ como:

$$\delta(x, S(\hat{x})) = \sum_{i \in S(\hat{x})} x_i - \sum_{i \notin S(\hat{x})} x_i.$$

Proposição

O corte de *otimalidade L-shaped inteiro* (3) é válido para SIP.

$$\theta \geq (q_{S(\hat{x})} - L) \left(\sum_{i \in S(\hat{x})} x_i - \sum_{i \notin S(\hat{x})} x_i \right) - (q_{S(\hat{x})} - L)(|S(\hat{x})| - 1) + L \quad (3)$$

Demonstração:

- Observe que $\delta(x, S(\hat{x})) \leq |S(\hat{x})|$. Se $x = \hat{x}$, $\delta(x, S(\hat{x})) = |S(\hat{x})|$ e, caso contrário, i.e., $x \neq \hat{x}$, $\delta(x, S(\hat{x})) < |S(\hat{x})|$.
- Se $x = \hat{x}$, o lado direito de (3) é $q_{S(\hat{x})}$.
- Se $x \neq \hat{x}$, o lado direito de (3) é L ou menor que L , que limita inferiormente $\mathcal{Q}(x)$. □

Reescrevendo (3)

$$\theta \geq (q_{S(\hat{x})} - L) (\delta(x, S(\hat{x})) - (|S(\hat{x})| - 1)) + L$$

Problema de segundo estágio

$$\begin{aligned} \min \quad & -2y_1 - 3y_2 \\ \text{s.t. } & 2y_1 + 2y_2 \leq \xi_1 - x_1 \\ & y_1 \leq \xi_2 - x_2 \\ & y \in \mathbb{Z}_+^2 \end{aligned}$$

Assumimos dois cenários $\xi^1 = (2, 2)^T, \xi^2 = (4, 3)$ equiprováveis.

Deseja-se:

- Encontrar um limite inferior para $\mathcal{Q}(x)$
- Deduzir um corte de otimalidade inteiro para $\hat{x} = (0, 1)^T$.

Devemos identificar o caso que amplia o domínio de $Q(x, \xi)$ para todo ξ

- Para limitar inferiormente $Q(x)$ basta perceber que o lado direito das restrições de segundo estágio devem ser tão grandes quanto possível, seja em uma relaxação contínua ou inteira para o problema de segundo estágio. Isto pode ser obtido para o ponto $\hat{x} = (0, 0)^T$.
- Possibilidade 1: Resolvemos a relaxação contínua do problema de segundo estágio $\hat{x} = (0, 0)^T$, para cada um dos dois cenários possíveis e ponderamos os valores de $Q(\hat{x}, \xi)$ pelas probabilidades. Neste caso, obtemos $L = -5.75$.
- Possibilidade 2: Resolvemos o problema inteiro para cada um dos dois cenários possíveis (e mesmo $\hat{x} = (0, 0)^T$), obtendo $L = -5.5$.

Escrevendo o corte de otimalidade inteiro

Para $\hat{x} = (0, 1)^T$, $\delta(x, S(\hat{x})) = x_2 - x_1$, $|S(\hat{x})| = 1$.

$Q(\hat{x}, \xi^1)$, $\xi^1 = (2, 2)^T$

$$\min -2y_1 - 3y_2$$

$$2y_1 + 2y_2 \leq 2$$

$$y_1 \leq 1$$

$$y \in \mathbb{Z}_+^2$$

$$y^* = (0, 1), Q(\hat{x}, \xi^1) = -3$$

$Q(\hat{x}, \xi^2)$, $\xi^2 = (4, 3)^T$

$$\min -2y_1 - 3y_2$$

$$2y_1 + 2y_2 \leq 4$$

$$y_1 \leq 2$$

$$y \in \mathbb{Z}_+^2$$

$$y^* = (2, 1), Q(\hat{x}, \xi^2) = -7$$

$\rightarrow q_{S(\hat{x})} = -5$ e o corte de otimalidade é escrito (usando a possibilidade 1) como:

$$\begin{aligned}\theta &\geq (-5 + 5.75)(x_2 - x_1) - 5.75 \\ &= 0.75(x_2 - x_1) - 5.75\end{aligned}$$

- Proposto por Laporte e Loveaux (1993).
- A versão apresentada aqui assume, de início, recurso relativamente completo. Logo, nesta versão não serão necessários cortes de viabilidade, apenas cortes de otimalidade.
- Posteriormente, discutiremos como cortes de viabilidade podem ser formulados, para o caso de variáveis de segundo estágio inteiras.
- É um algoritmo do tipo Branch-and-cut, que resolve uma sequência de CPs, e separa cortes de otimalidade.

Passo 0 - Inicialização

- $s \leftarrow 0, v \leftarrow 0, \bar{z} = +\infty$. v é o contador de subproblemas, s de cortes de optimidade.
- Inicialmente, atribui-se um valor qualquer a θ , inicialmente desconsiderado nos cálculos pois nenhum corte de optimidade foi gerado.
- Criamos uma lista \mathcal{L} de subproblemas.
- Esta lista \mathcal{L} inicialmente contém apenas o subproblema inicial (relaxação contínua do SIP), sem aplicação de particionamento algum.
- Obtenha algum L inicial para iniciar o algoritmo.

Passo 1 - Seleção de subproblema de \mathcal{L}

- Se $\mathcal{L} = \emptyset$, páre. O algoritmo foi concluído.
- Caso contrário, $v \leftarrow v + 1$. Escolha um subproblema \mathcal{P} da lista, como o problema corrente CP.

Passo 2 - Resolução do problema corrente

- Se o subproblema corrente \mathcal{P} é inviável, pode \mathcal{P} e retorne para o passo 1.
- Se \mathcal{P} é factível, seja (x^ν, θ^ν) uma solução ótima para o problema corrente.

Passo 3 - Verificação de poda por limites

- Se $c^T x^\nu + \theta^\nu > \bar{z}$, pode o problema corrente \mathcal{P} por otimalidade (limites) e volte ao passo 1.

Passo 4 - Ramificação ou branching

- Se $x^\nu \notin \mathbb{Z}^{n_1}$, ramifique da forma usual em algoritmos Branch-and-bound, criando dois novos subproblemas, que são adicionados à lista \mathcal{L} .

Passo 5 - Resolução do recurso para x^v

- Calcule $\mathcal{Q}(x^v)$ e defina $z^v = c^T x^v + \mathcal{Q}(x^v)$.
- Se $z^v < \bar{z}$, atualize $\bar{z} \leftarrow z^v$.

Observe que este passo só é executado para x^v satisfazendo a restrição de integralidade.

Passo 6 - Identificação de cortes de otimalidade L-shaped inteiros

- Se $\theta^v \geq \mathcal{Q}(x^v)$, retorne para o passo 1, podando o problema corrente.
- Caso contrário, formule um corte de otimalidade (3), para $q_{S(x^v)}$, faça $s \leftarrow s + 1$, atualizando a formulação para o problema corrente \mathcal{P} que deverá ser reotimizado, isto é, volte ao Passo 2, após ter inserido este corte de otimalidade em \mathcal{P} .

Ideia básica

Consiste em identificar desigualdades válidas (preferencialmente facetas) para a envoltória convexa das soluções inteiras do problema, violadas pelas soluções básicas da formulação empregada.

Isto pode ser feito tanto para o problema de primeiro estágio quanto para os de segundo estágio, na medida em que envolvam restrições de integralidade. Dificuldades encontradas no caso de SIP, relativas ao segundo estágio:

- As desigualdades válidas identificadas dependem da solução x^v de primeiro estágio fornecidas e da realização de cada estágio.
- Há um esforço adicional de caracterização e separação destas desigualdades, para cada realização de segundo estágio.

- Para ilustrar, vamos considerar o caso em que as variáveis de segundo estágio sejam binárias.
- Especializando as definições anteriores:
 - $Q(x, \xi^k) = \min_y \{q_k^T y : Wy = h_k - T_k x, y \in \{0, 1\}^{n_2}\}$
 - $C(x, \xi^k) = \min_y \{q_k^T y : Wy = h_k - T_k x, y \in [0, 1]^{n_2}\}$
- A ideia é modificar a formulação do domínio $\{Wy = h_k - T_k x, y \in [0, 1]^{n_2}\}$ da relaxação contínua do problema de segundo estágio, relativa a cada cenário ou realização $\xi^k : k = 1, \dots, K$, adicionando desigualdades válidas (ou cortes) que são violadas pelas soluções básicas do modelo.

Problema de segundo estágio

$$\begin{aligned} Q(x, \xi) = \min \quad & 3y_1 + 7y_2 + 9y_3 + 6y_4 \\ \text{s.t.} \quad & 2y_1 + 4y_2 + 5y_3 + 3y_4 \geq h - Tx \\ & y \in \{0, 1\}^4 \end{aligned}$$

Vamos considerar duas realizações (cenários):

- Para ξ^1 , temos: $h - Tx = 10 - 2x_1 - 4x_2$, com probabilidade 0.25
- Para ξ^2 , temos: $h - Tx = 11 - 4x_1 - 3x_2$, com probabilidade 0.75

Vamos considerar também a seguinte proposta de solução

$$x^v = (0.3, 0.6)^T$$
 de primeiro estágio.

$$h - Tx = 10 - 2x_1 - 4x_2, \quad x^\nu = (0.3, 0.6)^T.$$

Problema de segundo estágio

$$\begin{aligned} Q(x^\nu, \xi^1) = \min \quad & 3y_1 + 7y_2 + 9y_3 + 6y_4 \\ \text{s.t.} \quad & 2y_1 + 4y_2 + 5y_3 + 3y_4 \geq 7 \\ & y \in \{0, 1\}^4 \end{aligned}$$

- Solução da relaxação linear $y^* = (1, 1, 0.2, 0)^T$.
- Observe que a desigualdade de cobertura $y_3 + y_4 \geq 1$ é válida e corta y^* .
- Inserindo esta restrição e reotimizando, obtemos uma relaxação linear inteira $y^* = (1, 0, 1, 0)^T$.

$$h - Tx = 11 - 4x_1 - 3x_2, \quad x^\nu = (0.3, 0.6)^T.$$

Problema de segundo estágio

$$\begin{aligned} Q(x^\nu, \xi^2) = \min \quad & 3y_1 + 7y_2 + 9y_3 + 6y_4 \\ \text{s.t.} \quad & 2y_1 + 4y_2 + 5y_3 + 3y_4 \geq 8 \\ & y \in \{0, 1\}^4 \end{aligned}$$

- Solução da primeira relaxação linear: $y^* = (1, 1, 0.4, 0)^T$, que viola a restrição de cobertura $y_2 + y_3 + y_4 \geq 2$.
- Após reotimização, uma restrição adicional, $y_1 + y_3 \geq 1$, é necessária para obtenção de uma relaxação linear que satisfaça integralidade. A solução final é $y^* = (0, 0, 1, 1)^T$.

- Para cada realização ξ^k , é necessário um processo de identificação de desigualdades.
- Para um ξ^k fixo, as desigualdades geradas para um x^ν não podem garantidamente ser reutilizadas para o mesmo ξ^k e outro x^ν , a menos que sejam escritas em função de x .
- A reformulação vista permite resolver melhor cada subproblema, mas não permite melhor caracterizar a função recurso no primeiro estágio.
- Vamos empregar cortes disjuntivos para tentar tratar estas questões.

Desigualdades válidas para união de poliedros

Se $P^i = \{x \in \mathbb{R}_+^n : A^i x \geq b^i\}$ para $i = 0, 1$ são dois poliedros não vazios, então $\pi^T x \geq \pi_0$ é uma desigualdade válida para a envoltória convexa da união de P^0 e P^1 , $\text{co}(P^0 \cup P^1)$, se e somente se existirem $u^0, u^1 \geq 0$ tais que $\pi^T \geq (u^i)^T A^i$ e $\pi_0 \leq (u^i)^T b^i$ para $i = 0, 1$.

Demonstração

- Tomando multiplicadores $u^i \geq 0$, obtemos uma desigualdade agregada $(u^i)^T A x \geq (u^i)^T b^i$ válida para P^i .
- Podemos gerar uma desigualdade $\pi^T x \geq \pi_0$ válida para $\text{co}(P^0 \cup P^1)$, garantindo que $\pi^T \geq (u^i)^T A$ e $\pi_0 \leq (u^i)^T b^i$ para $i = 0, 1$.
- Como $x \geq 0$, $u^i \geq 0$ temos que

$$\pi^T x \geq (u^i)^T A x \geq (u^i)^T b^i \geq \pi_0.$$



- Suponha que a relaxação linear de uma formulação $P = \{y \in \mathbb{B}_+^{n_2} : Wy \geq d\}$ de um programa inteiro tenha sido resolvida e que uma variável y_j , binária no modelo, seja fracionária na solução ótima y^* da relaxação linear. Isto é, $y_j^* \in (0, 1)$.
- Definimos $P^0 = P \cap \{y \in \mathbb{R}_+^{n_2} : y_j = 0\}$ e $P^1 = P \cap \{y \in \mathbb{R}_+^{n_2} : y_j = 1\}$ e, ao invés de realizar branching em y_j , identificamos desigualdades válidas para $\text{co}(P^0 \cup P^1)$, violadas por y^* .
- Estas desigualdades válidas são inseridas em uma nova formulação do problema, cuja relaxação linear é reotimizada. Repete-se o processo até que a solução seja inteira ou não seja possível encontrar desigualdades válidas violadas.

- Resolvemos o problema de segundo estágio para uma solução de primeiro estágio e um cenário fixo.
- Desejamos uma desigualdade válida para $\text{conv}(P^0 \cup P^1)$, violada pela solução fracionária y^* . Identificamos uma componente j fracionária, por exemplo, a mais fracionária.
- Resolvemos o problema de separação, visando encontrar uma desigualdade válida, que maximize a violação:

$$\begin{aligned} & \max \pi_0 - (y^*)^T \pi \\ & \pi \geq (u^0)^T W - v^0 + w^0 e_j \\ & \pi \geq (u^1)^T W - v^1 + w^1 e_j \\ & \pi_0 \leq (u^0)^T d - e^T v^0 \\ & \pi_0 \leq (u^1)^T d - e^T v^1 + w^1 \end{aligned}$$

$$u^i, v^i, w^i \geq 0, i = 0, 1$$

$|\pi_0| \leq 1$ rest. normalização

Common cut coefficient technique

A ideia é utilizar cortes disjuntivos para obter uma família de desigualdades da forma

$$\pi^T y \geq \pi_0^k, \quad k = 1, \dots, K$$

que emprega o mesmo vetor de coeficientes π para todo cenário $k = 1, \dots, K$, mas com termos independentes π_0^k especializados para cada cenário k .

Assumimos que, para pelo menos um cenário ξ^k , a componente y_j^k da solução ótima da relaxação linear do problema de segundo estágio obtido com x^ν seja fracionária.

Proposição

As desigualdades $\pi^T y \geq \pi_0^k$, $k = 1, \dots, K$ são válidas se e somente se existirem $u^i, v^i, w^i \geq 0$ para $i = 0, 1$, tais que:

$$\pi \geq (u^0)^T W - v^0 - w^0 e_j$$

$$\pi \geq (u^1)^T W - v^1 - w^1 e_j$$

$$\pi_0^k \leq (u^0)^T d^k - e^T v^0 \quad k = 1, \dots, K$$

$$\pi_0^k \leq (u^1)^T d^k - e^T v^1 + w_1 \quad k = 1, \dots, K$$

onde $d^k = h^k - T^k x^\nu$.

Problema de segundo estágio

$$\begin{aligned}
 Q(x, \xi) = \min \quad & 3y_1 + 7y_2 + 9y_3 + 6y_4 \\
 \text{s.t.} \quad & 2y_1 + 4y_2 + 5y_3 + 3y_4 \geq h - Tx \\
 & y \in \{0, 1\}^4
 \end{aligned}$$

- $x^\nu = (0.3, 0.6)^T$
- Para ξ^1 , temos: $h - Tx = 10 - 2x_1 - 4x_2$, com probabilidade 0.25. Solução de segundo estágio: $y = (1, 1, 0.2, 0)^T$.
- Para ξ^2 , temos: $h - Tx = 11 - 4x_1 - 3x_2$, com probabilidade 0.75. Solução de segundo estágio: $y = (1, 1, 0.4, 0)^T$.
- $j = 3$, uma vez que $y_3 \notin \{0, 1\}$ nos dois cenários.
- Vamos buscar duas desigualdades disjuntivas, válidas para seus respectivos cenários, resolvendo um problema de separação só, maximizando a expectativa de violação $\sum_{k=1,2} p_k (\pi_0^k - \pi^T y^k)$.

Exemplo: problema de separação

$$\max z = 0.25\pi_0^1 + 0.75\pi_0^2 - \pi_1 - \pi_2 - 0.35\pi_3$$

$$s.t. \quad \pi_1 \geq 2u^0 - v_1^0 \quad \pi_1 \geq 2u^1 - v_1^1$$

$$\pi_2 \geq 4u^0 - v_2^0 \quad \pi_2 \geq 4u^1 - v_2^1$$

$$\pi_3 \geq 5u^0 - v_3^0 - w^0 \quad \pi_3 \geq 5u^1 - v_3^1 + w^1$$

$$\pi_4 \geq 3u^0 - v_4^0 \quad \pi_4 \geq 3u^1 - v_4^1$$

$$\pi_0^1 \leq 7u^0 - v_1^0 - v_2^0 - v_3^0 - v_4^0$$

$$\pi_0^1 \leq 7u^1 - v_1^1 - v_2^1 - v_3^1 - v_4^1 + w^1$$

$$\pi_0^2 \leq 8u^0 - v_1^0 - v_2^0 - v_3^0 - v_4^0$$

$$\pi_0^2 \leq 8u^1 - v_1^1 - v_2^1 - v_3^1 - v_4^1 + w^1$$

$$-e \leq \pi \leq e \quad -1 \leq \pi_0^1, \pi_0^2 \leq 1$$

$$u, v, w \geq 0$$

$$z = 0.35$$

$$\pi = (0, 0, \frac{2}{3}, 1)^T$$

$$u^0 = \frac{1}{3}$$

$$v^0 = (\frac{2}{3}, \frac{4}{3}, 0, 0)^T$$

$$u^1 = 0$$

$$v^1 = (0, 0, 0, 0)^T$$

$$w^0 = 1$$

$$w^1 = \frac{2}{3}$$

$$\pi_0^1 = \frac{1}{3}$$

$$\pi_0^2 = \frac{2}{3}$$

Os dois cortes gerados, para ξ^1, ξ^2 são respectivamente:

$$\frac{2}{3}y_3 + y_4 \geq \frac{1}{3}$$

$$\frac{2}{3}y_3 + y_4 \geq \frac{2}{3}$$

- Nosso objetivo agora é gerar cortes disjuntivos para cada programa de segundo estágio onde seja explícita a dependência em x . Este corte será válido para qualquer x e não apenas para o x^v fornecido.
- O corte (C^3) $\pi^T y \geq \pi_0^k$ que deduzimos é válido para o estágio k desde que:

$$\pi_0^k \leq (u^0)^T d^k - e^T v^0$$

$$\pi_0^k \leq (u^1)^T d^k - e^T v^1 + w^1$$

onde $d^k = h^k - T^k x^v$ é uma quantidade calculada para $x = x^v$, fixo.

- Explicitando a dependência de π_0^k em x :

$$\pi_0^k(x) \leq (u^0)^T (h^k - T^k x) - e^T v^0$$

$$\pi_0^k(x) \leq (u^1)^T (h^k - T^k x) - e^T v^1 + w^1$$

$$\pi_0^k(x) \leq (u^0)^T(h^k - T^k x) - e^T v^0$$

$$\pi_0^k(x) \leq (u^1)^T(h^k - T^k x) - e^T v^1 + w^1$$

- Mantendo π, u, v e w constantes, como foram calculados no procedimento que produziu o corte C^3 , basta tomarmos $\pi_0^k(x)$ suficientemente pequenos para produzir um corte válido.
- Para simplificar a notação e para π, u, v e w constantes, definimos:
 - $\alpha^{0,k} = (u^0)^T h^k - e^T v^0$ e $\alpha^{1,k} = (u^1)^T h^k - e^T v^1 + w^1$
 - $\beta^{i,k} = (u^i)^T T^k$ para $i = 0, 1$.
- Buscamos então um corte da forma:

$$\pi^T y \geq \min\{\alpha^{0,k} - \beta^{0,k} x, \alpha^{1,k} - \beta^{1,k} x\}$$

$$\pi^T y \geq \min\{\alpha^{0,k} - \beta^{0,k}x, \alpha^{1,k} - \beta^{1,k}x\} \quad (4)$$

- O operador min é não linear e não convexo (em x)
- Linearizando e convexificando, por meio de duas disjunções:
 - $P^0 = \{x \in \mathbb{R}_+^n : Ax \geq b, \gamma \geq 0, \gamma \geq \alpha^{0,k} - \beta^{0,k}x\}$
 - $P^1 = \{x \in \mathbb{R}_+^n : Ax \geq b, \gamma \geq 0, \gamma \geq \alpha^{1,k} - \beta^{1,k}x\}$
- Observe que se

$$\gamma \geq \alpha^{0,k} - \beta^{0,k}x$$

ou

$$\gamma \geq \alpha^{1,k} - \beta^{1,k}x,$$

isto é se $(x, \gamma) \in P^0 \cup P^1$, $\pi^T y \geq \gamma$ garante a satisfação de (4).

- Ou seja, vamos buscar uma desigualdade da forma

$$1\gamma + \rho_k^T x \geq \rho_{0,k}$$

válida para $\text{conv}(P^0 \cup P^1)$ e usa-la para escrever um corte da forma geral que desejamos:

$$\pi^T y \geq \rho_{0,k} - \rho_k^T x.$$

- Observe que procuramos uma desigualdade já **normalizada**, uma vez que impomos que o coeficiente de γ é +1.
- A desigualdade será dependente de x , mas será gerada visando maximizar a violação para x^v .
- Resolvemos um problema RHS(k) para cada cenário ξ^k .
- RHS(k) consiste em encontrar $r^i, s^i \geq 0$ para $i = 0, 1$ e $(\rho_k, \rho_{0,k})$ que:

$$\begin{aligned} RHS(k) \quad & \max \rho_{0,k} - \rho_k^T x^v \\ \text{s.t. } & \rho_k \geq (r^0)^T A + \beta^{0,k} s^0 \\ & \rho_k \geq (r^1)^T A + \beta^{1,k} s^1 \\ & \rho_{0,k} \leq (r^0)^T b + \alpha^{0,k} s^0 \\ & \rho_{0,k} \leq (r^1)^T b + \alpha^{1,k} s^1 \\ & \underline{1 \geq s^0} \\ & \underline{1 \geq s^1} \end{aligned}$$

- Vamos assumir que o primeiro estágio envolve a restrição $4x_1 + 6x_2 \leq 5$.
- Ponto a ser separado $x^v = (0.3, 0.6)^T$.
- Solução do C^3 : $u^0 = \frac{1}{3}$, $v^0 = (\frac{2}{3}, \frac{4}{3}, 0, 0)^T$,
 $u^1 = 0$, $v^1 = (0, 0, 0, 0)^T$, $w^0 = 1$, $w^1 = \frac{2}{3}$.

RHS(1)

- Vamos ilustrar para o caso $k = 1$, $h - Tx = 10 - 2x_1 - 4x_2$.
- $\alpha^{0,1} = \frac{1}{3}(10) - \frac{2}{3} - \frac{4}{3} = \frac{4}{3}$, $\beta^{0,1} = \frac{1}{3}(2, 4) = (\frac{2}{3}, \frac{4}{3})^T$.
 $\alpha^{1,1} = (u^1)^T 10 - e^T v^1 + w^1 = 0 + 0 + \frac{2}{3} = \frac{2}{3}$, $\beta^{0,1} = (0, 0)^T$.
- RHS(1) convexifica $\min\{\frac{4}{3} - \frac{2}{3}x_1 - \frac{4}{3}x_2, \frac{2}{3}\}$ para
 $x_1, x_2 \geq 0$, $4x_1 + 6x_2 \leq 5$

RHS(1)

$$z = \max \quad \rho_{0,1} - 0.3\rho_{1,1} - 0.6\rho_{1,2}$$

$$s.t. \quad \rho_{1,1} \geq -4r^0 + \frac{2}{3}s^0 \quad \rho_{1,1} \geq -4r^1$$

$$\rho_{1,2} \geq -6r^0 + \frac{4}{3}s^0 \quad \rho_{1,2} \geq -6r^1$$

$$\rho_{0,1} \leq -5r^0 + \frac{4}{3}s^0 \quad \rho_{0,1} \leq -5r^1 + \frac{2}{3}s^1$$

$$0 \leq r^0, r^1 \quad 0 \leq s^0, s^1 \leq 1$$

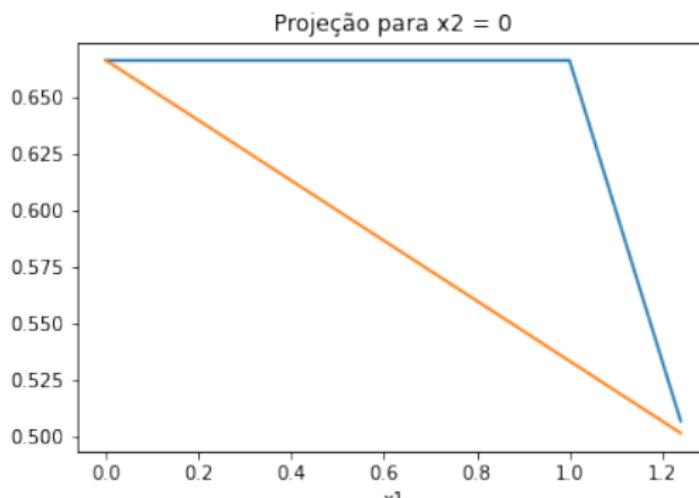
Solução ótima: $z = \frac{0.92}{3}$, $\rho_{0,1} = \frac{2}{3}$, $\rho_{1,1} = \frac{0.4}{3}$, $\rho_{1,2} = \frac{1.6}{3}$, que fornece o corte

$$\pi^T y = \frac{2}{3}y_1 + y_4 \geq \frac{2}{3} - \frac{0.4}{3}x_1 - \frac{1.6}{3}x_2$$

Exemplo: efeito convexificador da disjunção

Na figura, indicamos duas curvas, projetadas para $x_2 = 0, x_1 \leq 1.25$.

- $\pi^T y \geq \min\left\{\frac{4}{3} - \frac{2}{3}x_1 - \frac{4}{3}x_2, \frac{2}{3}\right\}$ é não linear, não convexa.
- $\pi^T y \geq \frac{2}{3} - \frac{0.4}{3}x_1 - \frac{1.6}{3}x_2$, a convexificação da anterior, é linear, convexa.
- No intervalo considerado, a função $\frac{2}{3} - \frac{0.4}{3}x_1 - \frac{1.6}{3}x_2$ limita inferiormente $\min\left\{\frac{4}{3} - \frac{2}{3}x_1 - \frac{4}{3}x_2, \frac{2}{3}\right\}$, podendo por esta razão ser empregada para dedução do corte.



Os passos abaixo são implementados na raiz dos problemas de segundo estágio, antes de fazer branching em algum deles.

- ① Requer a identificação de uma variável fracionária y_j^k para pelo menos um cenário ξ^k . Se isso não ocorrer, podemos gerar um corte L-shaped inteiro.
- ② Requer a separação do C^3 , common cut coefficient (para esta variável y_j^k fracionária). Os coeficientes π do corte C^3 produzido são anexados à matriz W^k daquele estágio.
- ③ Com a matriz W^k obtida com a anexação de π , resolvemos um problema RHS(k) para todo cenário ξ^k . Cada um dos cortes obtidos com RHS(k), anexamos $\rho_{0,k}$ ao vetor h^k e ρ_k^T à matriz T^k .
- ④ Resolvemos a relaxação linear fortalecida do segundo estágio, considerando os cortes RSH(k) introduzidos acima.
- ⑤ Usamos os multiplicadores duais associados à formulação fortalecida, para gerar cortes L-shaped, que serão exportados para o primeiro estágio.

- $K_2(\xi) = \{x : \exists y \in \mathbb{Z}_+^{n_2}, Wy = h(\omega) - T(\omega)x\}$ e
 $C_2(\xi) = \{x : \exists y \in \mathbb{R}_+^{n_2}, Wy = h(\omega) - T(\omega)x\}$ sua relaxação contínua.
- $K_2 = \bigcap_{k=1}^K K_2(\xi^k)$
- Vamos considerar o Programa Inteiro Misto abaixo, relativo ao segundo estágio, para um cenário ξ .

$$(P1) \quad w(x, \xi) = \min \quad e^T(v^+ - v^-)$$
$$\text{s.t. } Wy + I(v^+ - v^-) = h(\omega) - T(\omega)x$$
$$y \in \mathbb{Z}_+^{n_2}$$
$$v^+, v^- \geq 0$$

- Seja P1R, a relaxação linear de P1, e $\bar{w}(x, \xi)$ o valor ótimo da relaxação da função objetivo de P1R.

- $x \in K_2(\xi) \iff w(x, \xi) = 0$.
- $x \in C_2(\xi) \iff \bar{w}(x, \xi) = 0$.
- Se a componente y da solução ótima de P1R, \bar{y} , satisfaz $\bar{y} \in \mathbb{Z}_+^{n_2}$ e $\bar{w}(x, \xi) = 0$, não precisamos fazer branching em P1. Neste caso, resolver P1R basta para concluir que $x \in K_2(\xi)$.
- Se $\bar{w}(x, \xi) > 0$ temos que $(K_2(\xi) \cap) C_2(\xi) = \emptyset$. Neste caso, podemos usar a solução do dual de P1R para obter um corte de viabilidade L-shaped para inserir no primeiro estágio. Não precisamos de branching para produzir um corte de viabilidade para o primeiro estágio.
- Se $\bar{w}(x, \xi) = 0$ e a solução \bar{y} de P1R não é inteira não podemos concluir sobre x pertencer ou não a $K_2(\xi)$, sem realizar branching.

Vamos assumir que $\bar{w}(x, \xi) = 0$, $\bar{y} \notin \mathbb{Z}^{n_2}$ e que aplicamos um algoritmo Branch-and-bound (BB) para resolver P1. Vamos supor que, ao longo da árvore BB, detectamos que $w(x, \xi) > 0$. Vamos considerar os nós folha da árvore BB.

- A cada um destes nós, é associado um domínio $Y^\rho \subset \mathbb{R}^{n_2}$.
- O conjunto Y^ρ foi definido adicionando à formulação de segundo estágio para ξ^k desigualdades válidas, separadas ao longo do BB, e também por restrições de branching.
- Vamos assumir que existam R nós folha na árvore BB:
 $Y^\rho : \rho = 1, \dots, R$.
- Observe que $x \in K_2(\xi) \iff x \in \bigcup_{\rho=1}^R K_2^\rho(\xi)$ onde $K_2^\rho(\xi) = K_2(\xi) \cap Y^\rho$.
- Podemos construir um corte de viabilidade para cada Y^ρ e impor, uma restrição no primeiro estágio que garanta que x satisfaça um pelo menos destes R cortes.

- Vamos assumir que $u_\rho^T x \leq d_\rho$ represente um corte de viabilidade para o nó ρ .
- Vamos usar variáveis binárias $\delta_\rho \in \{0, 1\} : \rho = 1, \dots, R$ para determinar quais das R restrições são satisfeitas pelo ponto x de primeiro estágio. Isto é, reformulamos o problema de primeiro estágio com os seguintes cortes e variáveis adicionais:

$$\sum_{\rho=1}^R \delta_\rho \geq 1$$

$$u_\rho^T x \leq d_\rho + M_\rho(1 - \delta_\rho) \quad \rho = 1, \dots, R,$$

$$\delta \in \{0, 1\} \quad \rho = 1, \dots, R$$

onde M_ρ é um parâmetro big-M apropriado, isto é, que garanta que $u_\rho^T x \leq d_\rho + M_\rho$ seja válida para todo $x \in K_1$.

- Nem sempre é trivial obter este valor M_ρ .

$$-y_1 + y_2 \leq \xi - x_1$$

$$y_1 + y_2 \leq 2 - x_2$$

$$y_1, y_2 \in \mathbb{Z}_+$$

- $\xi \in \{1, 2\}$ equiprováveis.
- Basta considerarmos ξ^1 , pois $K_2(1) \subset K_2(2)$.
- Primeira solução produzida pelo mestre de Benders: $\hat{x} = (2, 2)^T$ que foi cortada pela introdução do corte de Benders $x_1 + x_2 \leq 3$.
- Segunda solução ofertada pelo mestre de Benders:
 $\hat{x} = (1.4, 1.6)^T$.

$$(P1R) \quad \bar{w}(\hat{x}, 1) = \min v_1 + v_2$$

$$-y_1 + y_2 - v_1 \leq -0.4$$

$$y_1 + y_2 - v_2 \leq 0.4$$

$$y_1, y_2 \geq 0$$

(P1R)

$$\begin{aligned}\bar{w}(\hat{x}, 1) &= \min v_1 + v_2 \\ -y_1 + y_2 - v_1 &\leq -0.4 \quad \rightarrow s_1 \\ y_1 + y_2 - v_2 &\leq 0.4 \quad \rightarrow s_2 \\ y_1, y_2 &\geq 0\end{aligned}$$

Dicionário ótimo de (P1R)

$$\begin{aligned}\bar{w}(\hat{x}, 1) &= 0 + v_1 + v_2 \\ y_1 &= 0.4 + y_2 + s_1 - v_1 \\ s_2 &= 0 - 2y_2 - s_1 + v_1 + v_2\end{aligned}$$

- Resolver P1R não é conclusivo. Branching é necessário.
- O algoritmo BB faz branching em y_1 (fracionária), criando dois subproblemas ou nós, associados respectivamente às restrições: $y_1 \leq 0$ e $y_1 \geq 1$ ($R = 2$).

→ s_3 representa a folga associada às restrições de branching.

Para $Y^1 : y_1 \leq 0$.

$$w = 0.4 + y_2 + s_1 + s_3 + v_2$$

$$y_1 = 0 - s_3$$

$$s_2 = 0.4 - y_2 + s_3 + v_2$$

$$v_1 = 0.4 + y_2 + s_1 + s_3$$

Para $Y^2 : y_1 \geq 1$.

$$w = 0.6 + y_2 + s_2 + s_3 + v_1$$

$$y_1 = 1 + s_3$$

$$v_2 = 0.6 + y_2 + s_2 + s_3$$

$$s_1 = 0.6 - y_2 + s_3 + s_1$$

- Veja que, depois de abertos os subproblemas Y^1 e Y^2 , é possível concluir que $w(\hat{x}, 1) > 0.4 = \min\{0.4, 0.6\}$.
- Associando os multiplicadores duais $(-1, 0, -1)$ para os rhs $1 - x_1, 2 - x_2, 0$ de Y^1 , obtemos o corte $x_1 \leq 1$.
- Associando os multiplicadores duais $(0, -1, 1)$ para os rhs $1 - x_1, 2 - x_2, 1$ de Y^2 , obtemos o corte $x_2 \leq 1$.
- Ou seja, a solução de primeiro estágio x deve satisfazer:

$$x_1 \leq 1 \vee x_2 \leq 1$$

A variável x deve satisfazer:

$$x_1 \leq 1 \vee x_2 \leq 1$$

- O problema de primeiro estágio passa incorporar novas variáveis binárias δ_1, δ_2 , que assumirão valor 1 caso $x_1 \leq 1, x_2 \leq 1$, sejam satisfeitas, respectivamente (0, em caso contrário).
- $M_1 = M_2 = 1$.
- O problema de primeiro estágio incorpora novas restrições:

$$\delta_1 + \delta_2 \geq 1$$

$$x_1 \leq 2 - \delta_1$$

$$x_2 \leq 2 - \delta_2$$

$$\delta_1, \delta_2 \in \{0, 1\}$$