

# Lagrange Relaxation: Introduction and Applications

Operations Research

Anthony Papavasiliou

## 1 Context

## 2 Applications

- Application in Stochastic Programming
- Unit Commitment

## 1 Context

## 2 Applications

- Application in Stochastic Programming
- Unit Commitment

# When to Use Lagrange Relaxation

Consider the following optimization problem:

$$\begin{aligned} p^* &= \max f_0(x) \\ f(x) &\leq 0 \\ h(x) &= 0 \end{aligned}$$

with  $x \in \mathcal{D} \subset \mathbb{R}^n$ ,  $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ ,  $h: \mathbb{R}^n \rightarrow \mathbb{R}^l$

Context for Lagrange relaxation:

- 1 *Complicating constraints*  $f(x) \leq 0$  and  $h(x) = 0$  make the problem difficult
- 2 Dual function is relatively easy to evaluate

$$g(u, v) = \sup_{x \in \mathcal{D}} (f_0(x) - u^T f(x) - v^T h(x)) \quad (1)$$

# Idea of Dual Decomposition

- Dual function  $g(u, v)$  is convex *regardless* of primal problem
- Computation of  $g(u, v)$ ,  $\pi \in \partial g(u, v)$  is relatively easy
- But...  $g(u, v)$  may be non-differentiable

Idea: minimize  $g(u, v)$  using algorithms that rely on linear approximation of  $g(u, v)$

- 1 Subgradient method
- 2 Cutting plane methods
- 3 Bundle methods

# Why Minimize the Dual Function?

We want to minimize the dual function for two reasons:

- Get as tight bounds to  $p^*$  as possible
- If we want to find primal optimal  $x^*$  by using the dual function, we have to minimize  $L(x, u, v)$  for  $(u, v)$  dual optimal

# Why Minimize the Dual Function: Bounding $p^*$

**Proposition:** If  $u \geq 0$  then  $g(u, v) \geq p^*$

Proof: If  $\tilde{x}$  is feasible and  $u \geq 0$  then

$$f_0(\tilde{x}) \leq L(\tilde{x}, u, v) \leq \sup_{x \in \mathcal{D}} L(x, u, v) = g(u, v).$$

Minimizing over all feasible  $\tilde{x}$  gives  $p^* \leq g(u, v)$

Conclusion: minimizing  $g(u, v)$  gives tightest possible bound to  $p^*$

# Dual Function Properties

**Proposition:**  $g(u, v)$  is convex lower-semicontinuous<sup>1</sup>. If  $(u, v)$  is such that (1) has optimal solution  $x_{u,v}$ , then  $\begin{bmatrix} -f(x_{u,v}) \\ -h(x_{u,v}) \end{bmatrix}$  is a subgradient of  $g$



---

<sup>1</sup>A function is lower-semicontinuous when its epigraph is a closed subset of  $\mathbb{R}^m \times \mathbb{R}^l \times \mathbb{R}$ .

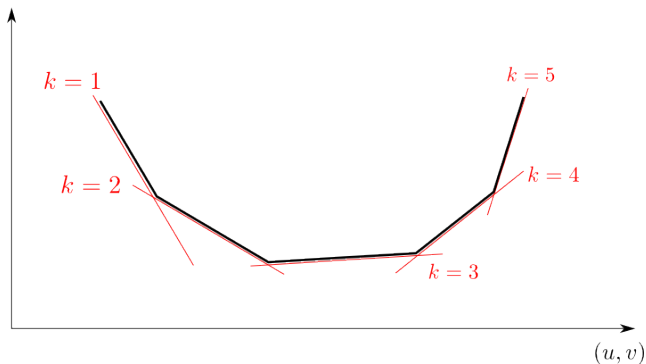


# Graphical Interpretation

Think of each  $x \in \mathcal{D}$  as an index  $k \in \mathcal{K}$ , then

$$g(u, v) = \max_{k \in \mathcal{K}} (f_{0k} - u^T f_k - v^T h_k)$$

$$g(u, v) = \max_k (f_{0k} - u^T f_k - v^T h_k)$$



## 1 Context

## 2 Applications

- Application in Stochastic Programming
- Unit Commitment

# Lagrange Relaxation of Stochastic Programs

Consider 2-stage stochastic program:

$$\begin{aligned} \min & f_1(x) + \mathbb{E}_\omega[f_2(y(\omega), \omega)] \\ \text{s.t.} & h_{1i}(x) \leq 0, i = 1, \dots, m_1, \\ & h_{2i}(x, y(\omega), \omega) \leq 0, i = 1, \dots, m_2 \end{aligned}$$

Introduce **non-anticipativity constraint**  $x(\omega) = x$  and reformulate problem as

$$\begin{aligned} \min & \mathbb{E}_\omega[f_1(x) + f_2(y(\omega), \omega)] \\ \text{s.t.} & h_{1i}(x) \leq 0, i = 1, \dots, m_1, \\ & h_{2i}(x(\omega), y(\omega), \omega) \leq 0, i = 1, \dots, m_2 \\ (\nu(\omega)) : & x(\omega) = x, \omega \in \Omega \end{aligned}$$

# Dual Function of Stochastic Program

Denote  $\nu = (\nu(\omega), \omega \in \Omega)$ , then

$$g(\nu) = g_1(\nu) + \mathbb{E}_\omega g_2(\nu(\omega), \omega)$$

where

$$\begin{aligned} g_1(\nu) = \inf_x \quad & f_1(x) + \left( \sum_{\omega \in \Omega} \nu(\omega) \right)^T x \\ \text{s.t.} \quad & h_{1,i}(x) \leq 0, i = 1, \dots, m_1, \end{aligned}$$

and

$$\begin{aligned} g_2(\nu(\omega), \omega) = \inf_{x(\omega), y(\omega)} \quad & f_2(y(\omega), \omega) - \nu(\omega)^T x(\omega) \\ \text{s.t.} \quad & h_{2,i}(x(\omega), y(\omega), \omega) \leq 0, i = 1, \dots, m_2 \end{aligned}$$

# Duality and Problem Reformulations

- Equivalent formulations of a problem can lead to very different duals
- Reformulating the primal problem can be useful when the dual is difficult to derive, or uninteresting

## Common reformulations

- Introduce new variables and equality constraints (we have seen this already)
- Rearrange constraints in subproblems

# Rearranging Constraints

Note alternative relaxation to the previous stochastic program:

$$\begin{aligned} \min \mathbb{E}_\omega [f_1(x(\omega)) + f_2(y(\omega), \omega)] \\ \text{s.t. } h_{1i}(x(\omega)) \leq 0, i = 1, \dots, m_1, \\ h_{2i}(x(\omega), y(\omega), \omega) \leq 0, i = 1, \dots, m_2 \\ x(\omega) = x \end{aligned}$$

This relaxation is probably useless (because subproblem involving  $x$  has no constraints)

# A Two-Stage Stochastic Integer Program [Sen, 2000]

Scenario	Constraints	Binary Solutions
$\omega = 1$	$2x_1 + y_1 \leq 2$ and $2x_1 - y_1 \geq 0$	$\mathcal{D}_1 = \{(0, 0), (1, 0)\}$
$\omega = 2$	$x_2 - y_2 \geq 0$	$\mathcal{D}_2 = \{(0, 0), (1, 0), (1, 1)\}$
$\omega = 3$	$x_3 + y_3 \leq 1$	$\mathcal{D}_3 = \{(0, 0), (0, 1), (1, 0)\}$

- 3 equally likely scenarios
- Define  $x$  as first-stage decision,  $x_\omega$  is first-stage decision for scenario  $\omega$
- **Non-anticipativity** constraint:  $x_1 = \frac{1}{3}(x_1 + x_2 + x_3)$

# Formulation of Problem and Dual Function

$$\max(1/3)y_1 + (1/3)y_2 + (1/3)y_3$$

$$\text{s.t. } 2x_1 + y_2 \leq 2$$

$$2x_1 - y_1 \geq 0$$

$$x_2 - y_2 \geq 0$$

$$x_3 + y_3 \leq 1$$

$$\frac{2}{3}x_1 - \frac{1}{3}x_2 - \frac{1}{3}x_3 = 0, (\lambda)$$

$$x_\omega, y_\omega \in \{0, 1\}, \omega \in \Omega = \{1, 2, 3\}$$

$$g(\lambda) = \max_{(x_\omega, y_\omega) \in \mathcal{D}_\omega} \left\{ \frac{2\lambda}{3}x_1 + \frac{1}{3}y_1 - \frac{\lambda}{3}x_2 + \frac{1}{3}y_2 - \frac{\lambda}{3}x_3 + \frac{1}{3}y_3 \right\}$$

$$= \max(0, 2\lambda/3) + \max(0, (-\lambda + 1)/3) + \max(1/3, -\lambda/3)$$



Dual function can be expressed equivalently

$$g(\lambda) = \begin{cases} \frac{1}{3} - \frac{2}{3}\lambda & \lambda \leq -1 \\ \frac{2}{3} - \frac{1}{3}\lambda & -1 \leq \lambda \leq 0 \\ \frac{2}{3} + \frac{1}{3}\lambda & 0 \leq \lambda \leq 1 \\ \frac{1}{3} + \frac{2}{3}\lambda & 1 \leq \lambda \end{cases}$$

- Primal optimal value  $p^* = 1/3$  with  $x_1^* = x_2^* = x_3^* = 1$ ,  $y_1^* = 0, y_2^* = 1, y_3^* = 0$
- Dual optimal value  $d^* = 2/3$  at  $\lambda^* = 0$

Conclusion: We have a duality gap of  $1/3$

# Alternative Relaxation

Add new explicit first-stage decision variable  $x$ , with the following *non-anticipativity constraints*:

$$x_1 = x, (\lambda_1)$$

$$x_2 = x, (\lambda_2)$$

$$x_3 = x, (\lambda_3)$$

Dual function:

$$\begin{aligned} g(\lambda) &= \max_{(x_1, y_1) \in \mathcal{D}_1} \left\{ \frac{1}{3} y_1 - \lambda_1 x_1 \right\} + \max_{(x_2, y_2) \in \mathcal{D}_2} \left\{ \frac{1}{3} y_2 - \lambda_2 x_2 \right\} + \\ &\quad \max_{(x_3, y_3) \in \mathcal{D}_3} \left\{ \frac{1}{3} y_3 - \lambda_3 x_3 \right\} + \max_{x \in \{0, 1\}} \left\{ (\lambda_1 + \lambda_2 + \lambda_3) x \right\} \\ &= \max(0, -\lambda_1) + \max\left(0, \frac{1}{3} - \lambda_2\right) + \max\left(\frac{1}{3}, -\lambda_3\right) + \\ &\quad \max(0, \lambda_1 + \lambda_2 + \lambda_3) \end{aligned}$$

# Closing the Duality Gap

Choosing  $\lambda_1 = 0$ ,  $\lambda_2 = \frac{1}{3}$  and  $\lambda_3 = -\frac{1}{3}$ , we have the following dual function value:

$$g(0, 1/3, -1/3) = 0 + 0 + \frac{1}{3} + 0 = \frac{1}{3}$$

# Conclusions of Example

- Different relaxations can result in different duality gaps
- Computational trade-off: introducing more Lagrange multipliers results in better bounds but larger search space

Consider the following notation

- Schedule  $I$  power plants over horizon  $T$
- Denote  $x_i^t$  as control of unit  $i$  in period  $t$ , denote  $x_i$  [resp.  $x^t$ ] as vector  $(x_i^t)_{t=1}^T$  [resp.  $(x_i^t)_{i \in I}$ ]
- Denote  $\mathcal{D}_i$  as feasible set of unit  $i$
- Denote  $C_i(x_i)$  as cost of producing  $x_i$  by unit  $i$
- Denote  $c^t(x^t) \leq 0$  as a *complicating constraint* that needs to be satisfied collectively by units, and suppose that it is additive:  $c^t(x^t) = \sum_{i \in I} c_i^t(x_i^t)$

Unit commitment problem:

$$\begin{aligned} \min \quad & \sum_{i \in I} C_i(x_i) \\ & x_i \in \mathcal{D}_i, i \in I \\ (u^t) : \quad & \sum_{i \in I} c_i^t(x_i^t) \leq 0, t = 1, \dots, T \end{aligned}$$

Relax *complicating constraints* to obtain the following Lagrangian:

$$L(x, u) = \sum_{i \in I} (C_i(x_i) + \sum_{t=1}^T u^t c_i^t(x_i^t))$$

What have we gained? We can solve one problem per plant:

$$\min_{x_i \in \mathcal{D}_i} (C_i(x_i) + \sum_{t=1}^T u^t c_i^t(x_i^t))$$

[1] S. Boyd, "Subgradient methods", EE364b lecture slides, <http://stanford.edu/class/ee364b/lectures/>

[2] C. Lemaréchal, "Lagrangian Relaxation", Computational combinatorial optimization. Springer Berlin Heidelberg, pp. 112-156, 2001.