Lagrange Relaxation:
Duality Gaps and Primal Solutions

Operations Research

Anthony Papavasiliou

Contents

Table of Contents

When to Use Lagrange Relaxation

Consider the following optimization problem:

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

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

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

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

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

Gaps and Feasible Solutions

Optimality gaps can guide termination

Dual function optimization does not solve the original problem

Table of Contents

Consider (without loss of generality) Lagrange relaxation without equality constraints:

$$\max f(x),\ x \in \mathcal{D},\ h_j(x) = 0, j = 1, \ldots, l \tag{2}$$

with Lagrangian function

$$L(x, v) = f_0(x) - \sum_{j=1}^{l} v_j h_j(x) = f_0(x) - v^T h(x) \tag{3}$$

and dual function

$$g(v) = \max_{x \in \mathcal{D}} L(x, v) \tag{4}$$

The dual problem is

$$\min g(v), \quad v \in \mathbb{R}^l$$

The Filling Property: Preliminary Definitions

Define the following sets:

$$\begin{aligned} \mathcal{D}(v) &= \{x \in \mathcal{D} : L(x, v) = g(v)\} \\ G(v) &= \{-h(x) : x \in \mathcal{D}(v)\} \end{aligned}$$

Interpretations:

- $\mathcal{D}(v)$: set of x that maximize the Lagrangian function at v
- $G(v)$: the image of $\mathcal{D}(v)$ through $-h(\cdot)$

The Filling Property

We know that $G(v) \subseteq \partial g(v)$, but when are they equal?

The **filling property** for (2) - (4) is said to hold at $v \in \mathbb{R}^l$ if $\partial g(v)$ is the convex hull of the set $G(v)$

When Does the Filling Property Hold?

The filling property holds at any $v \in \mathbb{R}^l$

- when $\mathcal{D}$ is a compact set on which f_0 and each h_j are continuous
- in particular, when $\mathcal{D}$ is a finite set (combinatorial optimization)
- in linear programming and in quadratic programming
- in problems where f_0 and h_j are l_p-norms, $1 \leq p \leq +\infty$

Perturbation Function

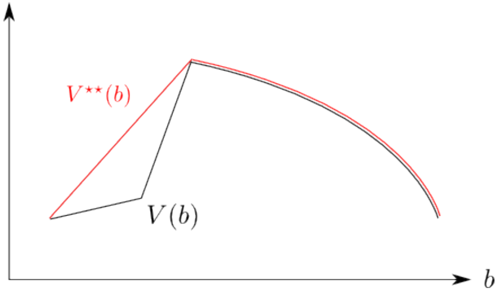
Define perturbation function as

$$V(b) = \max f(x)$$
$$x \in \mathcal{D}$$
$$h(x) = b$$

Concave Upper Semicontinous Hull

The concave upper semi-continuous hull of a function V is the smallest function $V^{\star\star}$ which is concave, upper semicontinuous and larger than V: $V^{\star\star}(b) \geq V(b)$

Graphical Illustration of USC Hull



Characterization of Dual Optimal Value

The dual optimal value is the value at 0 of the concave usc hull of the perturbation function: $\min g = V^{\star\star}(0)$

Convexifying $\mathcal{D}$ in Problems with Linear Data

Gap result 1: For an instance of (2) with linear data:

$$\max c^T x, \quad x \in \mathcal{D} \subset \mathbb{R}^n, \quad Ax = b$$

denote by $\bar{\mathcal{D}}$ the closed convex hull of $\mathcal{D}$. The dual minimal value $\min g$ is not smaller than the maximal value in the above equation, with $\mathcal{D}$ replaced by $\bar{\mathcal{D}}$.

Conclusion: the dual minimum is at least as large as the convex relaxation of the primal problem

Convexifying $\mathcal{D}$ in Problems with Linear Data: A Stronger Result

Gap result 2: Equality holds in *gap result 1* in any of the following cases:

1. $\bar{\mathcal{D}}$ is a bounded set in $\mathbb{R}^n$
2. for any $v \in \mathbb{R}^l$ close enough to 0, there exists $x \in \bar{\mathcal{D}}$ such that $Ax = b + v$
3. there exists $u^\star$ minimizing the dual function and the filling property holds at $u^\star$

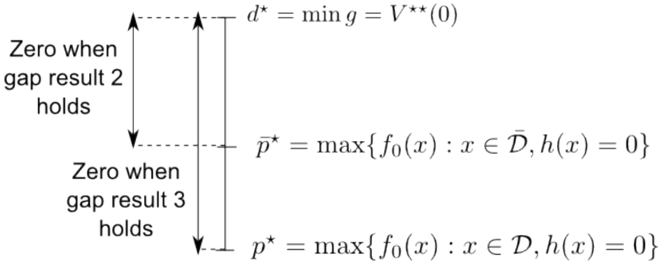
Conclusion: when *gap result 2* applies, Lagrange relaxation solves the "convex relaxation" of the primal problem

Zero Duality Gap

Gap result 3: Let (2) be a convex optimization problem: $\mathcal{D}$ is a closed convex set in $\mathbb{R}^n$, $f_0 : \mathcal{D} \to \mathbb{R}$ is a concave function, the constraint functions h_j are affine. Assume that the dual function (4) is not identically $+\infty$. Then there is no duality gap if one at least of the following properties holds:

1. $\mathcal{D}$ is a bounded set in $\mathbb{R}^n$, f_0 [resp. each inequality constraint] is upper [resp. lower] semicontinuous on $\mathcal{D}$
2. for any $v \in R^l$ close enough to 0, there is $x \in \mathcal{D}$ such that $h(x) = v$
3. there exists $u^\star$ minimizing the dual function, and the filling property holds at $u^\star$
4. $\mathcal{D}$ is a polyhedral set, f_0 and all constraints are affine functions (linear programming)

Duality Gap Relations



Counter-Example: A Convex Optimization Problem with a Non-Zero Duality Gap

Consider the following problem:

$$
\begin{aligned}
p^\star &= \min e^{-x} \\
&\text{s.t. } f(x,y) = x^2/y \leq 0 \\
&\mathcal{D} = \{(x,y) | y > 0\}
\end{aligned}
$$

- Show that x^2/y is a convex function for $y > 0$
- Conclude that this is a convex optimization problem
- Show that $p^\star = 1$

Relaxing $f(x, y)$, we have the following dual function:

$$g(u) = \min_{y>0}(e^{-x} + u\frac{x^2}{y})$$

with $\text{dom}(g) = \mathbb{R}_+$

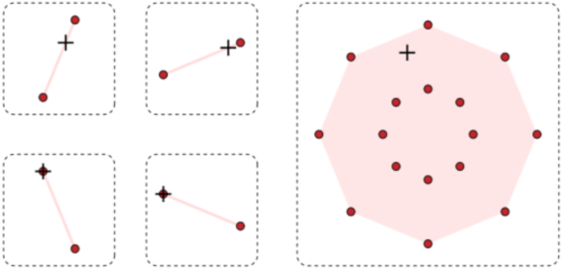
- Obviously, $g(u) \geq 0$ for all $u \geq 0$
- For any $\epsilon > 0$, $g(u) < \epsilon$ for all $u \geq 0$
- Therefore, $g(u) = 0$ for all $u \geq 0$
- Conclusion: $d^\star = \max g = 0 > p^\star = 1$
- Are the conditions of *gap result 1* satisfied? (Hint: no)
- Are the conditions of *gap result 2* satisfied? (Hint: no)
- Are the conditions of *gap result 3* satisfied? (Hint: no)

Shapley Folkman Lemma

Shapley-Folkman lemma: Let Y_i, $i = 1, \ldots, I$ be a collection of subsets of $\mathbb{R}^{m+1}$. Then for every $y \in conv(\sum_{i=1}^{I} Y_i)$ there exists a subset $I(y) \subset \{1, \ldots, I\}$ containing at most $m + 1$ indices such that

$$y \in \sum_{i \notin I(y)} Y_i + \sum_{i \in I(y)} conv(Y_i)$$

Graphical Illustration of Shapley-Folkman Lemma



- Consider four sets Y_i (left figure)
- The pink surface (right figure) indicates $conv(\sum_{i=1}^{I} Y_i)$
- Since $m + 1 = 2$, the point in the right is the sum of two points in Y_i and <u>two</u> points in $conv(Y_i)$

Relevance to Lagrange Relaxation

Consider an almost separable optimization problem with f_i, h_i linear[1]:

$$\max_{x_i \in \mathcal{D}_i} \sum_{i=1}^{n} f_i(x_i)$$

$$\text{s.t. } \sum_{i=1}^{n} h_i(x_i) = 0$$

Dual function:

$$g(\lambda) = \sum_{i=1}^{n} \max_{x_i \in \mathcal{D}_i} (f_i(x_i) + \lambda^T h_i(x_i))$$

[1] The linearity can be generalized, we use it to invoke gap result 2

Denote

$$\rho_i = \max_{x \in \bar{\mathcal{D}}_i} f_i(x) - \max_{x \in \mathcal{D}_i} f_i(x)$$

According to gap result 2,

$$\begin{aligned} d^\star &= \min g \\ &= \max_{x_i \in \bar{\mathcal{D}}_i} \{\sum_{i=1}^{n} f_i(x_i) : \sum_{i=1}^{n} h_i(x_i) = 0\} \end{aligned}$$

Apply the Shapley-Folkman theorem to the set $Y_i = \{(f_i(x), h_i(x))\}$ with $Y = \sum_{i=1}^{n} Y_i$ to get the following bound:

$$p^\star - d^\star \leq \frac{m+1}{n} E$$

where

$$E = \max_{i=i,\ldots,n} \rho_i$$

Graphical Illustration

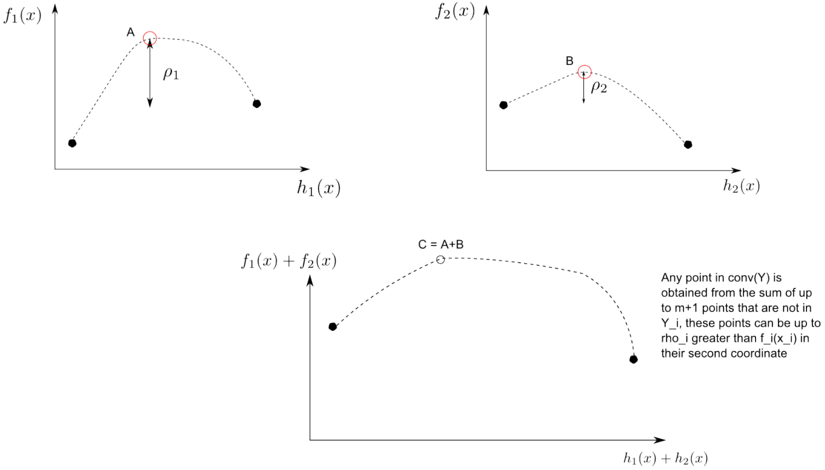


Table of Contents

Recovering Primal Optimal Solution

Optimality result 1: Let a dual algorithm produce $u^\star$ solving the dual problem. Suppose that

- the filling property holds
- appropriate convexity holds for X, c and L

Then (23) solves the primal problem

Primal Optimality in the Subgradient Algorithm

Optimality result 2: Let the subgradient method be applied with the following stepsizes:

$$a_k = \frac{\lambda_k}{\|\pi_k\|}, \quad \text{with } \lambda_k \downarrow 0 \text{ and } \sum_{k=1}^{\infty} \lambda_k = +\infty$$

Then g_k^{best} converges to $\inf g$

If the problem satisfies the assumptions of *optimality result 1*, then

$$\hat{x}_k = \frac{\sum_{j=1}^{k} a_j x_j}{\sum_{j=1}^{k} a_j}$$

converges to a primal optimal solution

References

[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.

[3] D. P. Bertsekas, N. R. Sandell, "Estimates of the Duality Gap for Large-Scale Separable Nonconvex Optimization Problems", IEEE Conference on Decision and Control, 1982.