Journal of the Operations Research Society of Japan
Vol. 60, No. 2, April 2017, pp. 44–49

DYNAMIC DUALIZATION IN A GENERAL SETTING

Hidefumi Kawasaki
Kyushu University

(Received February 15, 2016; Revised June 30, 2016)

Abstract Recently, Iwamoto, Kimura, and Ueno proposed dynamic dualization to present dual problems for unconstrained optimization problems whose objective function is a sum of squares. The aim of this paper is to show that dynamic dualization works well for unconstrained problems whose objective function is a sum of convex functions. Further we give another way to get dual problems, which is based on the infimal convolution. In both approaches we make clear the assumption for duality to hold.

Keywords: Dynamic programming, convex programming, duality theorem, dynamic dualization, infimal convolution

1. Introduction

Dual problems are usually defined for convex programming problems with constraints. Iwamoto, Kimura, and Ueno [1–3, 5] proposed dynamic dualization to present dual problems for unconstrained minimization problems whose objective function is a sum of squares. For example, for the following primal problem

$$\text{Minimize} \quad x^2 + (x+1)^2 + (x+2)^2, \quad x \in \mathbb{R}, \tag{1.1}$$

they introduce variables $u = x$, $v = x+1$, and $w = x+2$ to get a convex quadratic problem

$$\begin{aligned} \text{Minimize} \quad & u^2 + v^2 + w^2 \\ \text{subject to} \quad & x - u = 0,\ x+1-v = 0,\ x+2-w = 0. \end{aligned} \tag{1.2}$$

Its Lagrange dual is

$$\begin{aligned} \text{Maximize} \quad & -\frac{1}{4}(\lambda^2 + \mu^2 + \nu^2) + \mu + 2\nu \\ \text{subject to} \quad & \lambda + \mu + \nu = 0, \end{aligned} \tag{1.3}$$

where λ, μ, and ν are Lagrange multipliers corresponding to the constraints $x - u = 0$, $x + 1 - v = 0$, and $x + 2 - w = 0$, respectively. If we take $x - u = 0$ and $x + 1 - v = 0$ in (1.2), then we get another dual problem. If we take only $x - u = 0$ in (1.2), then we get another dual problem, see (4.4) in Section 4. Iwamoto, Kimura, and Ueno derived dual problems by computation, which is highly dependent on the sum of squares and perfect square. However dynamic dualization is effective in more general setting. In this paper, we deal with the following primal problem

$$(P) \qquad \text{Minimize} \quad \sum_{i=1}^{m} f_i(x), \quad x \in \mathbb{R}^n, \tag{1.4}$$

where $f_i : \mathbb{R}^n \to (-\infty, \infty]$ $(i = 1, \ldots, m)$ are convex functions not identically ∞. In Section 2, we show that dynamic dualization works well in a general setting. In Section 3, we show another route to reach the dual problem obtained by dynamic dualization, which is based on the infimal convolution. In Section 4, we discuss another type of dynamic dualization.

Throughout this paper, $\inf(P)$ and $\sup(D)$ denote the infimum of the primal problem (P) and the supremum of the dual problem (D), respectively. When the infimum is attained by some feasible solution of (P), we denote $\inf(P)$ by $\min(P)$. Similarly we denote $\sup(D)$ by $\max(D)$, where $\max(D)$ does not always implies $\max(D) \in \mathbb{R}$. When the value of the objective function is $-\infty$ for some feasible solution of (D), $\max(D) = -\infty$.

2. Dynamic Dualization

Dynamic dualization first introduces new variables $u_1, \ldots, u_m \in \mathbb{R}^n$ to (1.4), and transforms it into a minimization problem with m linear equality constraints:

$$\text{Minimize} \quad \sum_{i=1}^{m} f_i(u_i) \quad \text{subject to} \quad x - u_i = 0 \quad (i = 1, \ldots, m). \tag{2.1}$$

Its Lagrange function is given by

$$L(x, u_1, \ldots, u_m, y_1, \ldots, y_m) := \sum_{i=1}^{m} f_i(u_i) + \sum_{i=1}^{m} y_i^T (x - u_i), \tag{2.2}$$

where $y_i \in \mathbb{R}^n$, and its Lagrange dual problem is given by

$$(D_L) \qquad \text{Maximize} \quad \inf\{L(x, u_1, \ldots, u_m, y_1, \ldots, y_m) \mid x, u_1, \ldots, u_m \in \mathbb{R}^n\}. \tag{2.3}$$

The last step of dynamic dualization is to express the objective function by the dual variables $y_1, \ldots, y_m$.

The following theorem is given in Rockafellar [4, Corollary 28.2.2].

Theorem 2.1. (Lagrange duality theorem) Let (P_0) be a convex programming problem with m linear equality constraints

$$(P_0) \qquad \text{Minimize} \quad f(x) \quad \text{subject to} \quad Ax = b. \tag{2.4}$$

If (P_0) has a feasible solution and $\inf(P_0)$ is finite, then there exists $y \in \mathbb{R}^m$ such that

$$\inf\{L(x, y) \mid x \in \mathbb{R}^n\} = \inf(P_0), \tag{2.5}$$

where $L(x, y) := f(x) + y^T(Ax - b)$.

By applying Theorem 2.1 to (2.1), we obtain the following duality theorem.

Theorem 2.2. For the primal problem (1.4), its dual is given by

$$(D) \qquad \text{Maximize} \quad -\sum_{i=1}^{m} f_i^*(y_i) \quad \text{subject to} \quad \sum_{i=1}^{m} y_i = 0 \in \mathbb{R}^n, \tag{2.6}$$

where $f_i^*(y_i) := \sup\{y_i^T u_i - f_i(u_i) \mid u_i \in \mathbb{R}^n\}$, and if $\inf(P)$ is finite, then $\inf(P) = \sup(D)$.

Proof. It is evident that the primal problem (2.1) has a feasible solution. So Theorem 2.1 is applicable to (2.1). Since

$$\begin{aligned} &\inf\{L(x, u_1, \ldots, u_m, y_1, \ldots, y_m) \mid x, u_1, \ldots, u_m \in \mathbb{R}^n\} \\ = &\inf\left\{ \sum_{i=1}^{m}(f_i(u_i) - y_i^T u_i) + \sum_{i=1}^{m} y_i^T x \,\middle|\, x, u_1, \ldots, u_m \in \mathbb{R}^n \right\} \\ = &\begin{cases} -\sum_{i=1}^m f_i^*(y_i) & \text{if } \sum_{i=1}^m y_i = 0, \\ -\infty & \text{if } \sum_{i=1}^m y_i \neq 0, \end{cases} \end{aligned}$$

(D_L) becomes (D), and we get $\inf(P) = \sup(D)$. □

Example 2.1. Let $a_i \in \mathbb{R}$ and $f_i(x) = (x + a_i)^2$. Then since $f_i^*(y_i) = y_i^2/4 - a_i y_i$, the dual problem is

$$\begin{aligned} \text{Maximize} \quad & -\frac{1}{4}(y_1^2 + \cdots + y_m^2) + a_1 y_1 + \cdots + a_m y_m \\ \text{subject to} \quad & y_1 + \cdots + y_m = 0. \end{aligned}$$

3. Infimal Convolution

In this section, we present another route to reach the dual problem (2.6). For any convex function $f : \mathbb{R}^n \to (-\infty, \infty]$, its *conjugate* is defined by

$$f^*(y) := \sup\{y^T x - f(x) \mid x \in \mathbb{R}^n\}, \quad y \in \mathbb{R}^n. \tag{3.1}$$

The *epigraph* of f is defined by $\operatorname{epi} f := \{(x, r) \in \mathbb{R}^{n+1} \mid f(x) \leq r\}$. The *closure* $\operatorname{cl} f$ of f is defined so that $\operatorname{epi}(\operatorname{cl} f)$ is the closure of $\operatorname{epi} f$. The closure of f is the greatest lower semi-continuous function majorized by f.

The *infimal convolute* of $f_1, \ldots, f_m$ is defined by

$$(f_1 \Box f_2 \Box \cdots \Box f_m)(x) := \inf\{f_1(x_1) + \cdots + f_m(x_m) \mid x_1 + \cdots + x_m = x\}. \tag{3.2}$$

The operation $\Box$ is called *infimal convolution.*

Theorem 3.1. (Rockafellar [4, Theorem 16.4]) Let $f_i : \mathbb{R}^n \to (-\infty, \infty]$ $(i = 1, \ldots, m)$ be convex functions not identically ∞. Then for any $y \in \mathbb{R}^n$

$$(\operatorname{cl} f_1 + \cdots + \operatorname{cl} f_m)^*(y) = \operatorname{cl}(f_1^* \Box \cdots \Box f_m^*)(y) \in (-\infty, \infty]. \tag{3.3}$$

If the relative interior of $\operatorname{dom} f_i$ $(i = 1, \ldots, m)$ have a point in common, the closure operation can be omitted from (3.3), and

$$(f_1 + \cdots + f_m)^*(y) = \min\{f_1^*(y_1) + \cdots + f_m^*(y_m) \mid y_1 + \cdots + y_m = y\} \in (-\infty, \infty]. \tag{3.4}$$

Remark 3.1. In (3.4), $\min\{f_1^*(y_1) + \cdots + f_m^*(y_m) \mid y_1 + \cdots + y_m = y\} = \infty$ means that there exist $y_1, \ldots, y_m \in \mathbb{R}^n$ such that $y_1 + \cdots + y_m = y$ and

$$f_1^*(y_1) + \cdots + f_m^*(y_m) = \infty.$$

Applying Theorem 3.1 to the primal problem (1.4), we obtain the following duality theorem.

Theorem 3.2. Let $f_i : \mathbb{R}^n \to (-\infty, \infty]$ $(i = 1, \dots, m)$ be convex functions not identically ∞. Assume that the relative interior of dom f_i $(i = 1, \dots, m)$ have a point in common, then the dual of (1.4) is (2.6), and $\inf(P) = \max(D) \in [-\infty, \infty)$.

Proof. For $f := \sum_{i=1}^{m} f_i$, since $f^*(0) = \sup\{-f(x) \mid x \in \mathbb{R}^n\} = -\inf\{f(x) \mid x \in \mathbb{R}^n\}$, we get from (3.4)

$$\inf(P) = -f^*(0) = \max\{-f_1^*(y_1) - \cdots - f_m^*(y_m) \mid y_1 + \cdots + y_m = 0\} \in [-\infty, \infty). \quad (3.5)$$

□

Figure 1 illustrates a dual problem for $m = 2$. Since $y_2 = -y_1$, the dual problem is to maximize $-f_1^*(y_1) - f_1^*(-y_1)$. In both figures the slope of tangent lines are $\pm y_1$.

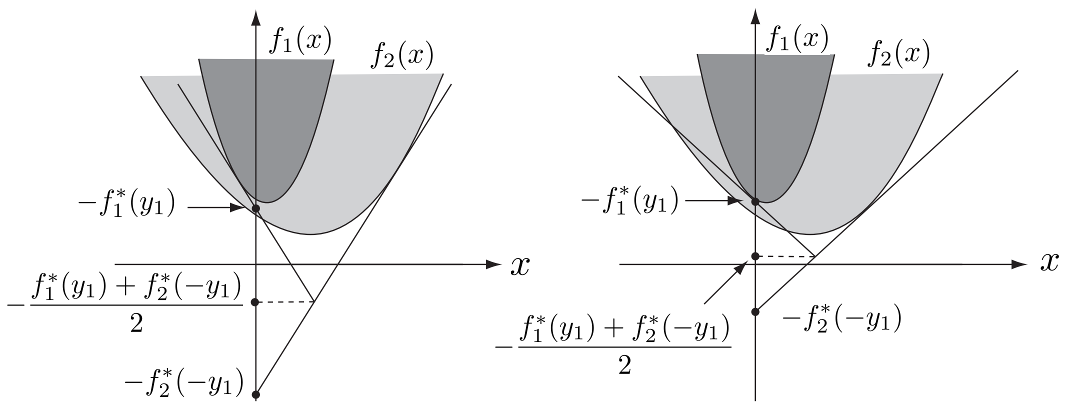

Figure 1: The dual problem for $m = 2$

4. Dynamic Dualization 2

As we mentioned in the introduction, Iwamoto, Kimura, and Ueno [1–3, 5] gave another type of dynamic dualization for the primal problem (1.4). The second dynamic dualization introduces variables u_i $(i \in I)$ and constraints $x - u_i$ $(i \in I)$, where I is a proper subset of $\{1, \dots, m\}$. Then the primal problem is equivalent to

$$\text{Minimize} \quad \sum_{i \in I} f_i(u_i) + \sum_{i \notin I} f_i(x) \quad \text{subject to} \quad x - u_i = 0 \quad (i \in I). \quad (4.1)$$

Since the Lagrange function is

$$L(x, \dots, u_i, \dots, y_i, \dots) = \sum_{i \in I} f_i(u_i) + \sum_{i \notin I} f_i(x) + \sum_{i \in I} y_i^T (x - u_i), \quad (4.2)$$

where $y_i \in \mathbb{R}^n$ $(i \in I)$, we have

$$\begin{aligned} & \inf\{L(x, \dots, u_i, \dots, y_i, \dots) \mid x, u_i \in \mathbb{R}^n \ (i \in I)\} \\ = \ & \inf \left\{ \sum_{i \in I} (f_i(u_i) - y_i^T u_i) + \sum_{i \notin I} f_i(x) + \sum_{i \in I} y_i^T x \ \middle| \ x, u_i \in \mathbb{R}^n \ (i \in I) \right\} \\ = \ & -\sum_{i \in I} f_i^*(y_i) - \left(\sum_{i \notin I} f_i \right)^* \left(-\sum_{i \in I} y_i \right). \end{aligned}$$

Applying Theorem 2.1 to (4.1), we get the following dual problem

$$(D_2) \qquad \text{Maximize} \quad -\sum_{i\in I} f_i^*(y_i) - \left(\sum_{i\notin I} f_i\right)^* \left(-\sum_{i\in I} y_i\right) \tag{4.3}$$
$$\text{subject to} \quad y_i \in \mathbb{R}^n \ (i \in I).$$

Theorem 4.1. If $\inf(P)$ is finite, then $\inf(P) = \sup(D_2)$.
Example 4.1. ([5]) Let $f_1(x) = x^2$, $f_2(x) = (x+1)^2$, $f_3(x) = (x+2)^2$, and $I = \{1\}$. Then

$$f_1^*(y_1) = \frac{1}{4}y_1^2, \quad (f_2 + f_3)^*(-y_1) = \frac{1}{8}y_1^2 + \frac{3}{2}y_1 - \frac{1}{2}.$$

Hence the dual problem (D_2) is written as

$$\text{Maximize} \quad -\frac{3}{8}y_1^2 - \frac{3}{2}y_1 + \frac{1}{2} \quad \text{subject to} \quad y_1 \in \mathbb{R}.$$

In [5] they put $\lambda = y_1/2$, and obtained

$$\text{Maximize} \quad -\frac{3}{2}\lambda^2 - 3\lambda + \frac{1}{2} \quad \text{subject to} \quad \lambda \in \mathbb{R}. \tag{4.4}$$

Remark 4.1. If we apply (3.4) to the second term of (4.3), then we have

$$\begin{aligned} & -\left(\sum_{i\notin I} f_i\right)^* \left(-\sum_{i\in I} y_i\right) \\ = & \max\left\{ -\sum_{i\notin I} f_i^*(z_i) \,\middle|\, \sum_{i\notin I} z_i = -\sum_{i\in I} y_i, \quad z_i \in \mathbb{R}^n \ (i \notin I) \right\} \\ = & \max\left\{ -\sum_{i\notin I} f_i^*(z_i) \,\middle|\, \sum_{i\in I} y_i + \sum_{i\notin I} z_i = 0, \quad z_i \in \mathbb{R}^n \ (i \notin I) \right\} \end{aligned}$$

Hence (D_2) reduces to (D). Therefore the second dynamic dualization can be regarded as a halfway point of the first dynamic dualization.

5. Concluding Remarks

We close this paper with making a comparison between dynamic dualization and infimal convolution. The former (Theorem 2.2) assumes only $\inf(P) \in \mathbb{R}$, and claims that $\inf(P) = \sup(D) \in \mathbb{R}$. The latter (Theorem 3.1) assumes that the relative interior of $\operatorname{dom} f_i$ ($i = 1, \ldots, m$) have a point in common, and claims that $\inf(P) = \max(D)$. The latter assumption is trivially satisfied if every f_i is real-valued. However $\inf(P) = \max(D)$ might be $-\infty$.
Example 5.1. Let $f_i(x) = a_i^T x + b_i$ ($i = 1, \ldots, m$). Since f_i is real-valued, Theorem 3.2 is applicable. Since

$$f_i^*(y_i) = \begin{cases} -b_i & y_i = a_i \\ \infty & y_i \neq a_i, \end{cases}$$

the dual problem is to maximize $\sum_{i=1}^m b_i$ subject to $\sum_{i=1}^m y_i = 0$ and $y_i = a_i$ ($i = 1, \ldots, m$), and we have

$$\inf(P) = \max(D) = \begin{cases} \sum_{i=1}^m b_i & \sum_{i=1}^m a_i = 0 \\ -\infty & \sum_{i=1}^m a_i \neq 0. \end{cases}$$

This example is outside of [1–3, 5].

Acknowledgements
The author would like to thank anonymous referees for careful reading and useful comments, which improved our paper. This research was supported by JSPS KAKENHI Grant Number 16K05278.

References

[1] S. Iwamoto: *Mathematics for Optimization II: Bellman Equation* (Chisen Shokan, Tokyo, 2013) (in Japanese).

[2] S. Iwamoto and Y. Kimura: Three dualizations—two-variable optimization. 20th Symposium of Informatics and Statistical Sciences, Kyushu University (2015) (in Japanese).

[3] Y. Kimura, T. Ueno, and S. Iwamoto: Two duals of one primal (preprint).

[4] R.T. Rockafellar: *Convex Analysis* (Princeton University Press, 1970).

[5] T. Ueno, Y. Kimura, and S. Iwamoto: Seven dual problems. Workshop "Mathematical Models of Decision Making Under Uncertainty and Ambiguity, and Related Topics", Research Institute for Mathematical Sciences, Kyoto University, (2015) (in Japanese).

Hidefumi Kawasaki
Faculty of Mathematics
Kyushu University
Moto-oka 744, Nishi-ku
Fukuoka 819-0395, Japan
E-mail: kawasaki@math.kyushu-u.ac.jp