LAGRANGE MULTIPLIERS IN FENCHEL DUALITY*

VIKRAM RAMANATHAN

Abstract. In this paper, we present the equivalence of Fenchel and Lagrangian duality. Then, for the problems of concave maximization and convex minimization, we show that the Lagrange multiplier is indeed equal to the constant η used in Fenchel duality.

1. Introduction.

The equivalence of Fenchel and Lagrangian duality is not explicitly stated in most of the popular convex optimization textbooks (including our class text, [2]). Magnanti pointed this out and goes on to show this equivalence in his 1974 paper [3]. Magnanti specifically equates the optimal value of the Lagrange dual problem with that of the Fenchel dual problem. For the interested reader, we recommend reading the original paper but the main contributions will be summarized later.

This paper specifically considers the equivalence of the Lagrange multiplier used in Lagrangian duality and the dual variable used in the Fenchel duality strategy presented in class. This equivalence was briefly questioned while solving a workshop problem and was left as an open problem. Below, we provide the necessary preliminaries as well as the main results from our investigation into this problem.

THEOREM 1.1 (Fenchel Duality [2]). *Assume that f and g are, respectively, convex and concave functionals on the convex sets C and D in a normed space X. Assume that $C \cap D$ contains points in the relative interior of C and D, i.e. $ri(C) \cap ri(D) \neq \emptyset$, and that either [f,C] or [g,D] has a nonempty interior. Suppose that $\mu = \inf_{x \in C \cap D} \{f(x) - g(x)\}$ is finite. Then,*

$$\mu = \inf_{x \in C \cap D} \{f(x) - g(x)\} = \max_{x^* \in C^* \cap D^*} \{g^*(x^*) - f^*(x^*)\}$$

where the maximum on the right is achieved by some $x_0^* \in C^* \cap D^*$.If the infimum on the left is achieved by some, $x_0 \in C \cap D$, then

$$\max_{x \in C} [< x, x_0^* > -f(x)] =< x, x_0^* > -f(x_0)$$

and,

$$\min_{x \in D} [< x, x_0^* > -g(x)] =< x, x_0^* > -g(x_0)$$

Now, we will separately state the Lagrange dual problem and Slater's constraint qualification. Consider the following optimization problem, termed "ordinary convex problem" by Rockafellar [4],

$$\begin{aligned} \text{minimize} \quad & f_0(x) \\ \text{subject to} \quad & f_i(x) \leq 0, \quad i = 1, ..., r \\ & h_i(x) = 0, \quad i = r+1, ..., m \end{aligned} \tag{1.1}$$

where $x \in C_0$, C_0 is a non-empty convex set in $\mathbb{R}^n$, f_i is a convex function on C_0 for $i = 0, 1, ...r$, and f_i is an affine function on C_0 for $i = r+1, ..., m$. The Lagrangian, $\mathcal{L}$, for this problem is defined as,

*Submitted to the MATH 116 staff 12/17/2020.

$$\mathcal{L}(x,\lambda,\nu) = f_0(x) + \sum_{i=1}^{r} \lambda_i f_i(x) + \sum_{i=r+1}^{m} \nu_i h_i(x) \tag{1.2}$$

where λ_i and ν_i are referred to as the Lagrange multiplier for the ith inequality constraint and the ith equality constraint respectively. Additionally, $\lambda = (\lambda_1, ...\lambda_r)$ and $\nu = (\nu_1, ...\nu_r)$. The corresponding Lagrange dual problem is stated as follows,

$$d^* = \max_{\lambda,\nu,\lambda\succeq 0} \inf_{x\in C_0} \mathcal{L}(x,\lambda,\nu) \tag{1.3}$$

Strong duality holds under Slater's condition.

THEOREM 1.2 (Strong Duality given Slater's condition [1]). *Slater's condition states that* $\exists x \in ri(C_0)$ *such that*

$$f_i(x) < 0, \quad i = 1, ..., r$$

$$h_i(x) = 0, \quad i = r+1, ..., m$$

i.e. x is strictly feasible. If the primal problem satisfies Slater's constraint qualification, strong duality holds, i.e. $p^* = d^*$.

Now that we have detailed the necessary background, we will present some of Magnanti's work. One important result is that he proves that given Slater's condition (strong duality), the Lagrangian dual can be written as a Fenchel dual. Subsequently, he shows that the main result from 1.1 holds.

Let $C_0 = C\text{x}D\text{x}\mathbb{R}^n$ and $x = (x_1, x_2, x_3)$. Then, the Fenchel problem can be restated as,

$$p_f^* = \inf_{x\in C_0} \{f_1(x_1) - f_2(x_2) | x_1 = x_3, x_2 = x_3\} \tag{1.4}$$

where p_f^* is the optimal value of the corresponding Fenchel primal problem. Intuitively, this means that x is chosen such that $x_1 \in C$ and $x_2 \in D$. The constraint, $x_1 = x_2 = x_3$, gives us, $x_1, x_2 \in C \cap D$.

Magnanti then notes that this is an ordinary convex problem, stated as,

$$\begin{aligned} \text{minimize} \quad & f_1(x_1) - f_2(x_2) \\ \text{subject to} \quad & h_1(x) = x_1 - x_3 = 0 \\ & h_2(x) = x_2 - x_3 = 0 \end{aligned} \tag{1.5}$$

whose Lagrangian dual problem is,

$$d_l^* = \max_{\lambda_j\in\mathbb{R}^n} \inf_{x\in C_0} \{f_1(x_1) - f_2(x_2) + \lambda_1(x_1 - x_3) + \lambda_2(x_2 - x_3)\} \tag{1.6}$$

With some algebraic manipulation, we get,

$$d_l^* = \max_{\lambda_j\in\mathbb{R}^n} \inf_{x\in C_0} \{f_1(x_1) - f_2(x_2) + \lambda_1 x_1 + \lambda_2 x_2 - (\lambda_1 + \lambda_2)x_3\} \tag{1.7}$$

Assuming that $p_f^* > -\infty$ and that the Slater condition holds, $p_f^* = d_l^*$. Magnani then states that the Lagrangian dual 1.8 can be written as the Fenchel dual and that the dual maximization occurs for some $x^* = -\lambda_1 = \lambda_2$. This last step is not exactly trivial and is not covered by Magnanti. So, we have written it out below.

Since $x_3 \in \mathbb{R}^n$, the infimum in the Lagrangian dual is $-\infty$ when $\lambda_1 + \lambda_2 \neq 0$. So, with $\lambda_1 + \lambda_2 = 0$, 1.8 reduces to,

$$\begin{aligned} d_l^* &= \max_{\lambda_j \in \mathbb{R}^n} \inf_{x \in C_0} \{f_1(x_1) - f_2(x_2) - \lambda_2 x_1 + \lambda_2 x_2\} \\ &= \max_{\lambda_j \in \mathbb{R}^n} \inf_{x_1 \in C} \{f_1(x_1) - \lambda_2 x_1\} + \inf_{x_2 \in D} \{-f_2(x_2) + \lambda_2 x_1\} \\ &= \max_{\lambda_j \in \mathbb{R}^n} -f_1^*(\lambda_2) + f_2^*(\lambda_2) \\ &= d_f^* \end{aligned} \tag{1.8}$$

where d_f^* denotes the optimal value of the Fenchel dual problem. Thus, $x^* = \lambda_2 = -\lambda_1$.

2. Main results. In class, we use Fenchel duality to either minimize a convex functional, f, or maximize a concave functional, g, defined on a convex set with an affine constraint. In this case, it is even simpler to see the equivalence of the Fenchel dual and the Lagrange multiplier. Below, we present this result for the problem of maximizing a concave functional. The result for the minimization of a convex functional can be very easily derived using similar steps.

2.1. Maximizing concave functional. We attempt to minimize the concave functional, g on D subject to an affine constraint.

2.1.1. Fenchel Duality Approach. Take $f = 0$. C is then taken to be the set where the affine constraint, $x_1 + \sum_{i=2}^n \alpha_i x_i = x_o$, holds. $y \in X^*$.

$$\begin{aligned} f^*(y) &= \sup_{x \in C}(y^T x - 0) \\ &= \sup_{x \in C}(y_1(x_1 + \alpha_2 x_2 + ... + \alpha_n x_n) - \sum_{i=2}^n (y_i - \alpha_i y_1) x_i) \end{aligned} \tag{2.1}$$

So $y_i = \alpha_i y_1$ for $i \in 2, ..., n$ for the supremum to exist. Thus, $y = (y_1, y_2, ..., y_n) = (\eta, \alpha_2 \eta, ..., \alpha_n \eta)$ and $C^* = \{(\eta, \alpha_2\eta, ..., \alpha_n\eta)\}$. Thus, C is the 1D subspace generated by $(1, \alpha_2, \alpha_3, ..., \alpha_n)$.

As a result,

$$f^*(y) = \eta x_0 \tag{2.2}$$

Using the Fenchel duality theorem, the original concave functional maximization problem reduces to the following dual optimization problem on η (where $y = (y_1, y_2, ..., y_n) = (\eta, \alpha_2\eta, ..., \alpha_n\eta)$).

$$\begin{aligned} \mu &= \max_{y \in C^* \cap D^*} g^*(y) - f^*(y) \\ &= g^*(y) - \eta x_0 \end{aligned} \tag{2.3}$$

2.1.2. Lagrange Duality Approach.

$$\begin{aligned} &\text{min.} \quad -g(x) \\ &\text{s.t} \quad x_1 + \sum_{i=2}^{n} \alpha_i x_i = x_o \end{aligned} \tag{2.4}$$

We assume the above satisfies Slater's condition, so strong duality holds. The Lagragian for this problem is given below.

$$\mathcal{L}(x, \nu) = -g(x) + \nu(x_1 + \sum_{i=2}^{n} \alpha_i x_i - x_o) \tag{2.5}$$

Thus, the Lagrangian dual problem is given by,

$$\begin{aligned} \mu &= \inf \mathcal{L}(x, \nu) \\ &= \inf_{x \in D} (-g(x) + \nu(x_1 + \sum_{i=2}^{n} \alpha_i x_i - x_o)) \\ &= -\nu x_0 + \inf_{x \in D} (-g(x) + \nu(x_1 + \sum_{i=2}^{n} \alpha_i x_i)) \\ &= -\nu x_0 + \inf_{x \in D} (-g(x) + \nu x_1 + \sum_{i=2}^{n} \nu \alpha_i x_i) \\ &= -\nu x_0 + g^*(y) \end{aligned} \tag{2.6}$$

where $y = (y_1, y_2, ..., y_n) = (\nu, \alpha_2 \nu, ..., \alpha_n \nu)$.

Hence, we have shown that the Lagrange multiplier, ν, is equal to the Fenchel duality η, i.e. $\nu = \eta$.

Acknowledgments. I would like to acknowledge Paul, Weifeng, and Kevin for all their help on class material. The course is very well-structured and has fascinating homework/workshop problems to work on. Without a doubt, this class was one of the best classes I have ever taken. I have learned a great deal of extremely interesting mathematical theory and potential applications. Thank you.

REFERENCES

[1] S. Boyd, S. P. Boyd, and L. Vandenberghe, *Convex optimization*, Cambridge university press, 2004.
[2] D. G. Luenberger, *Optimization by vector space methods*, John Wiley & Sons, 1997.
[3] T. L. Magnanti, *Fenchel and lagrange duality are equivalent*, (1974).
[4] R. T. Rockafellar, *Convex analysis*, no. 28, Princeton university press, 1970.