APEX DUALITY FOR CONSTRAINED OPTIMIZATION

R. J. DUFFIN[1], D. F. KARNEY[2] AND E. Z. PRISMAN[3]

Abstract

In this paper, we develop a duality theory for the Apex dual in the case of primal constraints. As suggested by Duffin in [4], the objective function in this framework is a weighted average of the Legendre-Lagrangian function evaluated at key points. We show that whenever this new dual is feasible there is no duality gap for this dual, and moreover, no duality gap for both the Lagrangian and Wolfe duals too. We conclude with an outline of an algorithm to solve constrained minimization problems in the Apex framework.

1. Introduction

In this paper, we consider the problem of estimating the minimum value of a differentiable convex function subject to constraints by means of dual inequalities. As noted in the unconstrained case [4], [5], in order to numerically estimate the minimum value of a convex function by means of dual inequalities, the following three properties are essential:

a) A dual program must be found.

b) All functions should be given explicitly.

c) All equality constraints should be linear.

If any of these properties do not hold, then the estimation technique would be of little value numerically. In fact, in extending dual inequality estimation techniques to the constrained case, the above properties can be very restrictive. Linear programming duals and ordinary geometric programming duals satisfy properties

[1] Department of Mathematics, Carnegie-Mellon University, Pittsburgh, PA 15213
[2] School of Business, University of Kansas, Lawrence, KS 66045
[3] College of Business Administration, Department of Finance, Arizona State University, Tempe, AZ 85287

(a, b, c) but, in general, conjugate duals (property (b)), ordinary Lagrangian duals (property (b)) and Wolfe duals (property (c)) do not. In other words, either dual inequality estimation techniques must be restricted to a narrow class of constrained optimization problems or a new dual must be developed.

Given the above, in this paper we construct a new dual program for constrained convex optimization problems which satisfies properties (a, b, c). The proposed dual is based upon the Apex dual developed in [4] for the unconstrained case. As such, the new dual is also termed the Apex dual. As seen below, this nomenclature results in little, if any, ambiguity.

Since Apex duality is a relatively new idea, in the next section we provide a sketch of its formulation along with the pertinent duality results. In Section 3, we construct the Apex dual for the constrained case, while in Section 4 we derive its duality properties. We show, for instance, under a feasibility assumption, that the value of the Apex dual is the value of the original constrained minimization problem. Given that the Apex dual satisfies properties (a, b, c), this implies that constrained convex minimization problems can indeed be "solved" by means of dual inequality techniques. In addition, under the same feasibility assumption, we show that the value of the Apex dual equals the value of both the Lagrangian and Wolfe duals too. We conclude this paper with a strategy for solving constrained minimization problems on a computer, utilizing the Apex machinery.

2. The unconstrained case

In this section, we consider the problem of estimating the minimum value of a convex function f on R^n in the absence of constraints. Intuitively, the minimum value of f and its minimizer x^* (assume such an x^* exists) can be approximated in the following manner: first, select $n + 1$ distinct points in R^n (label these points $x_1, \ldots, x_{n+1}$) such that the convex hull of $x_1, \ldots, x_{n+1}$ contains x^* in its interior. Next, for each $i = 1, \ldots, n + 1$, form the half space H_i in R^{n+1} which contains the graph f and whose boundary supports (is "tangent" to) the graph of f at $(x_i, f(x_i))$. The $\bigcap_{i=1}^{n+1} H_i$ is then a convex cone which contains the graph of f. Moreover, in the case that this cone does not contain any "horizontal lines", the "y-coordinate" of the cone's apex is actually a lower bound on $f(x)$ and the "x-coordinate" is an approximation for x^*. In fact, as seen in [4], $f(x^*)$ is actually the supremum of all such y-coordinates, as long as one such cone exists.

In order to formalize the above, as in [4], we restrict our attention to the case that $f(x)$ is a differentiable convex function on R^2. The differentiability restriction forces each of the supporting hyperplanes H_i to be unique and completely determined by the partial derivatives of f at x_i. The restriction of x to R^2 is

strictly a matter of convenience for purposes of discussion, after which the extension to higher dimensions will be apparent.

Let $f(x)$ be a differentiable convex function with $x \in R^2$ and, for $i = 1, 2$, let $f_i'(x)$ be the partial derivative of f with respect to x_i. Then, the equation of the tangent plane to the graph of f at the point $(\bar{x}_1, \bar{x}_2)$ is given by the equation

$$x_3 = x_1 f_1'(\bar{x}_1) + x_2 f_2'(\bar{x}_2) + \phi(\bar{x}_1, \bar{x}_2)$$

where ϕ is the Legendre function of f; that is

$$\phi = f - x_1 f_1' - x_2 f_2'.$$

Now, let x, y, z, be three arbitrary but fixed points in R^2. Let Δ be the following matrix of the partial derivatives of f evaluated at x, y, z:

$$\Delta = \begin{pmatrix} 1 & 1 & 1 \\ f_1'(x) & f_1'(y) & f_1'(z) \\ f_2'(x) & f_2'(y) & f_2'(z) \end{pmatrix}.$$

In addition, let w_1, w_2, w_3, be the cofactors of the first row of the matrix Δ. In other words, let

$$w_1 = f_1'(y)f_2'(z) - f_1'(z)f_2'(y)$$
$$w_2 = f_1'(z)f_2'(x) - f_1'(x)f_2'(z)$$
$$w_3 = f_1'(x)f_2'(y) - f_1'(y)f_2'(x).$$

Finally, let C be the convex cone generated by the system of inequalities

$$x_3 \geqslant x_1 f_1'(x) + x_2 f_2'(x) + \phi(x)$$
$$x_3 \geqslant x_1 f_1'(y) + x_2 f_2'(y) + \phi(y) \qquad (2.1)$$
$$x_3 \geqslant x_1 f_1'(z) + x_2 f_2'(z) + \phi(z).$$

Note that by construction, each of the above inequalities is a supporting hyperplane of the graph of f.

By definition, the apex of the cone C above is the point (x_1, x_2, x_3) which satisfies each of the inequalities in (2.1) as an equality. Moreover, the "y-coordinate" of the apex, call it $\psi(x, y, z)$, is precisely

$$\frac{w_1 \phi(x) + w_2 \phi(y) + w_3 \phi(z)}{w_1 + w_2 + w_3}. \qquad (2.2)$$

To see this, first rewrite each of the inequalities in (2.1) as an equality. Then multiply the first, second and third equations by w_1, w_2 and w_3 respectively and add. Finally, observe that the resulting coefficients of x_1 and x_2 are zero, as the cofactors of a given row of a matrix are orthogonal to any other row.

In the case that each of w_1, w_2, w_3 is positive, $\psi(x, y, z)$ is actually a lower bound on the minimum value of f. This follows from the fact that $(x^*, f(x^*))$

satisfies the system of inequalities given by (2.1), and the fact that the above argument extends to the inequality case as long as each w_i is positive.

Given the above, it is now possible to define a dual problem to the problem which seeks the minimum value of f. Specifically, find

$$MA = \sup \psi(x, y, z) \tag{A}$$

where x, y, z are chosen so that each of w_1, w_2, w_3 is positive. As MA is generated by the "y-coordinate" of the apex of the various approximating cones of f, program (A) is called the *Apex dual*. As the following example illustrates, the Apex dual satisfies properties (a, b, c) of Section 1. In fact, since (a) and (c) are trivially satisfied, we need only verify (b).

EXAMPLE 1. Let $f(x_1, x_2) = x_1^2 + x_2^2$. Then $\phi(x_1, x_2) = -x_1^2 - x_2^2$.
Let x, y, z be three arbitrary but fixed points in R^2. Then

$$\Delta = \begin{pmatrix} 1 & 1 & 1 \\ -2x_1 & -2y_1 & -2z_1 \\ -2x_2 & -2y_2 & -2z_2 \end{pmatrix}$$

from which it follows that

$$w_1 = 4y_1 z_2 - 4y_2 z_1,$$
$$w_2 = -4x_1 z_2 + 4x_2 z_1,$$
$$w_3 = 4x_1 y_2 - 4x_2 y_1.$$

An explicit formula for $\psi(x, y, z)$ is now immediate, given the above formulae for w_1, w_2, w_3, $\phi(x_1, x_2)$, $\phi(y_1, y_2)$ and $\phi(z_1, z_2)$, and (2.2).

We say that a triple of points $[x, y, z]$ is a *feasible solution* for the Apex dual if the cofactors w_1, w_2, w_3 of the matrix Δ are positive. By construction, it follows that if $[x, y, z]$ is a feasible solution to the program (A), then $\psi(x, y, z) \leqslant f(x^*)$, which in turn implies that $MA \leqslant f(x^*)$. In fact, as seen in [4], the following stronger statement can be made.

THEOREM 2.1. (Duffin [4]). *If the minimum points of $f(x)$ form a bounded set, then the Apex dual (A) is feasible. Moreover, the supremum value of program (A) is equal to the minimum value of $f(x)$.*

Conversely, if the Apex dual is feasible, then the minimum value of $f(x)$ is attained and the set of minimum points is bounded. Moreover, the supremum value of program (A) is equal to the minimum value of $f(x)$.

As a consequence of the above, we note that if $u, x, y, z \in R^2$ are chosen so that u is in the domain of f, and if $[x, y, z]$ is feasible for the program (A), then $(f(u) + \psi(x, y, z))/2$ approximates $f(x^*)$ with an error of at most $(f(u) - \psi(x, y, z))/2$. Moreover, since x^* is in the interior of the triangle generated by

the points x, y, z (denoted by Δxyz), we can approximate x^* by x_A, the "x-coordinate" of the apex of cone C, and we can also estimate how far x_A is from x^*.

An iterative scheme for estimating x^* and $f(x^*)$ is now immediate. Observe that the point x_A divides Δxyz above into three subtriangles, Δxyx_A, Δxx_Az and Δx_Ayz, one of which contains x^*. In other words, except for the case discussed below, one of the three sets of points, $[x, y, x_A]$, $[x, x_A, z]$ and $[x_A, y, z]$ is a feasible solution for the program (A). Hence the above estimation process can be repeated for this new feasible solution. Terminate this process when either the diameter of the approximating triangle is sufficiently small or the functional approximation is good enough. Note, if a uniform decrease in the diameter of the approximating triangle is desired, then use the epicenter of Δxyz in place of x_A in the above.

The above scheme needs to be modified in the case that x^* is on the boundary of at least one of above subtriangles of Δxyz. In this case, replace x_A by any other point of the interior of Δxyz which is not on the line determined by x^* and x_A, and continue.

The algorithm outlined above has an important pedagogical implication. Basically, such an algorithm can be implemented on a personal computer by students in an introductory optimization course. Then, after successful implementation, students become better motivated and certain topics, such as rates of convergence, become more understandable.

3. The constrained case

In this section, we construct the Apex dual for the constrained minimization program (P) which seeks

$$MP = \inf f(x) \tag{P}$$

subject to

$$g_i(x) \leq 0; \quad i = 1, 2, \ldots, m$$

where $f(x)$ and each $g_i(x)$ is a differentiable convex function on R^n. (Note, if f and/or one or more of the g_i is not defined on all of R^n, then we shall assume that the respective domains of all the functions in the program (P) intersect.)

As a first step, we consider the ordinary Lagrangian dual (D) of the program (P) which seeks

$$MD = \sup_{\lambda \geq 0} \inf_x L(x, \lambda), \tag{D}$$

where $L(x, \lambda) = f(x) + \sum_{i=1}^m \lambda_i g_i(x)$.

Since, for each $\lambda \geqslant 0$, the function $L(x, \lambda)$ is a differentiable convex function of x, it follows that for each $\lambda \geqslant 0$, there exists an (unconstrained) Apex dual (A_λ) for the Lagrangian dual objective function $\inf_x L(x, \lambda)$. This implies that for each $\lambda \geqslant 0$, the optimal value MA_λ of the program (A_λ) is a lower bound for MD and hence for MP also. Thus the program (A) which seeks

$$MA = \sup_{\lambda \geqslant 0} MA_\lambda \qquad\qquad (A)$$

is a legitimate dual for the program (P). In fact Program (A) is called the *Apex dual* of the program (P). Before developing this dual further, however, an example is appropriate.

EXAMPLE 2. Find the minimum value of $x_1^2 + x_2^2$ subject to $x_1 + x_2 \geqslant 1$. Given the above,

$$L(x_1, x_2, \lambda) = x_1^2 + x_2^2 + \lambda(1 - x_1 - x_2).$$

For fixed λ, we have

$$L_1'(x_1, x_2, \lambda) = 2x_1 - \lambda$$

and

$$L_2'(x_1, x_2, \lambda) = 2x_2 - \lambda.$$

Thus

$$\phi_\lambda(x_1, x_2) = -x_1^2 - x_2^2 + \lambda.$$

Let x, y, z be three arbitrary but fixed points in R^2. Let λ be fixed. Then

$$\Delta_\lambda = \begin{pmatrix} 1 & 1 & 1 \\ 2x_1 - \lambda & 2y_1 - \lambda & 2z_1 - \lambda \\ 2x_2 - \lambda & 2y_2 - \lambda & 2z_2 - \lambda \end{pmatrix}.$$

Let $w_{1,\lambda}$, $w_{2,\lambda}$, and $w_{3,\lambda}$ be the cofactors of the first row of Δ_λ. Then

$$\psi_\lambda(x, y, z) = (w_{1,\lambda}\phi_\lambda(x) + w_{2,\lambda}\phi_\lambda(y) + w_{3,\lambda}\phi_\lambda(z))/(w_{1,\lambda} + w_{2,\lambda} + w_{3,\lambda})$$

and

$$MA_\lambda = \sup_{\lambda \geqslant 0} \psi_\lambda(x, y, z),$$

where x, y, z are chosen so that each of $w_{1,\lambda}$, $w_{2,\lambda}$ and $w_{3,\lambda}$ is positive.

Technically, as seen in the above example, the Apex dual as defined above does not satisfy property (b) in that MA_λ is not given explicitly. In order to satisfy property (b), the Apex dual (A) can be rewritten as follows: find

$$MA = \sup \psi_\lambda(x_1, \ldots, x_{n+1}),$$

where the supremum is take over $\lambda, x_1, \ldots, x_{n+1}$, where $\lambda \geqslant 0$ and $x_1, \ldots, x_{n+1}$ represent $n + 1$ distinct points of R^n which form a feasible solution for the program (A_λ).

Since the two formulations for the Apex dual (A) are equivalent, in the proof below we shall utilize the original formulation involving MA_λ while in the computational remarks below we shall defer to the reformulation involving ψ_λ.

We say that a vector $\lambda \geqslant 0$ is *feasible* for the Apex dual (A) if and only if the program (A_λ) is feasible, that is, if and only if there exists $n + 1$ distinct points in R^n for which the cofactors of the first row of the matrix Δ_λ are positive. Theorem 2.1 now implies

PROPOSITION 3.1. *A vector λ is feasible for the Apex dual if and only if the* $\inf_x L(x, \lambda)$ *is finite and attained on a bounded set.*

Let the set of feasible solutions to program (A) be denoted by dom(A), with a similar notation for the feasible solutions of program (D). Then by construction,

$$\text{dom}(A) \subseteq \text{dom}(D) \qquad (3.1)$$

and

$$MA \leqslant MD \leqslant MP. \qquad (3.2)$$

In the next section, we show that $MA = MP$ whenever the Apex dual is feasible. Given this, the above inequality now implies that the Apex dual is not feasible whenever there is a Lagrangian duality gap, that is, whenever $MD < MP$. As the next example illustrates, it can happen that the Apex dual is not feasible even when $MD = MP$.

EXAMPLE 3. Consider the problem

$$\inf \exp(x)$$

subject to $x \leqslant 0$.

By construction, $MP = MD = 0$ and $MA = -\infty$. To see the second claim, first let $\lambda = 0$. Then $\inf_x L(x, 0) = 0$. However this infimum is not attained. Thus Theorem 2.1 implies that the program (A_0) is not feasible. Now let $\lambda > 0$ be arbitrary but fixed. Then $\inf_x L(x, \lambda) = -\infty$. Again, Theorem 2.1 implies that the program (A_λ) is not feasible. Hence (A) is not feasible and $MA = -\infty$.

In order to characterize when the Apex dual is feasible, we need to study the behavior of f on unbounded sequences. Specifically, we say that the function f is *unbounded at infinity* if $f(x_k) \to +\infty$ whenever (x_k) is a sequence with $\|x_k\| \to +\infty$.

LEMMA 3.2. *Let $f(x)$ be a continuous convex function on R^n. Then f is unbounded at infinity if and only if the minimum set of f is nonempty and bounded.*

PROOF. Assume that f is unbounded at infinity and let (x_k) be a minimizing sequence of f. Then by assumption (x_k) is bounded from which it follows that (x_k) has a cluster point x^*. Without loss of generality we may assume that x^* is a limit point. Now $f(x^*) = \inf f$ since (x_k) is a minimizing sequence and f is continuous. Thus the minimum of f is attained. Moreover the minimum set of f is bounded since, by assumption, $f(x_k) \to +\infty$ whenever $\|x_k\| \to +\infty$.

Assume now that the minimum set of f is nonempty and bounded. In addition assume that f is not unbounded at infinity. Then there exists a sequence (x_k) in R^n with $\|x_k\| \to +\infty$ for which $f(x_k) \to M < +\infty$. Let y be a cluster point of the sequence $(x_k/\|x_k\|)$. Without loss of generality, we may assume that y is a limit point. Moreover, without loss of generality, we may also assume that the origin minimizes f. Then, for all $\alpha \geq 0$, αy is a minimizer of f. To see this, let α be an arbitrary but fixed positive number. Then, for all k sufficiently large, $\|x_k\| > \alpha$. Thus, by convexity, for these k

$$f(\alpha x_k/\|x_k\|) \leq (1 - \alpha/\|x_k\|)f(0) + (\alpha/\|x_k\|)f(x_k).$$

Now, by continuity, the left hand side of the above inequality converges to $f(\alpha y)$ as k tends to infinity and, by construction, the right hand side converges to $f(0)$. Thus

$$f(\alpha y) \leq f(0) \quad \text{for all } \alpha \geq 0.$$

But $f(0)$ is the minimum value of f. Thus

$$f(\alpha y) = f(0) \quad \text{for all } \alpha \geq 0$$

from which it follows that the minimum set of f is unbounded, a contradiction. In other words, f is unbounded at infinity whenever the minimum set of f is nonempty and bounded.

As an immediate consequence to the above lemma, we can now characterize when the Apex dual is feasible.

COROLLARY 3.3. *The vector λ is feasible for the Apex dual if and only if the function $L(x, \lambda)$ is unbounded at infinity.*

PROOF. By Proposition 3.1, λ is feasible for the Apex dual if and only if the minimizing set of $L(x, \lambda)$ is nonempty and bounded. By Lemma 3.2, the minimizing set of $L(x, \lambda)$ is nonempty and bounded if and only if $L(x, \lambda)$ is unbounded at infinity.

Note that the above lemma and its corollary imply that the Apex dual is feasible whenever the function f is unbounded at infinity. This follows from the fact that $L(x, 0) = f(x)$. Thus Theorem 4.4 below implies that there is no Apex duality gap whenever the objective function f is unbounded at infinity.

We now conclude this section with the following Lagrangian duality result for canonical programs, which is utilized in the proof of Theorem 4.4 and found in [1], [2]. At this time we note that program (P) is called *canonical* if there exists a real number α and a positive number δ such that

$$\{x|f(x) \leqslant \alpha\} \cap \left(\bigcap_{i=1}^{m} \{x|g_i(x) \leqslant \delta\} \right)$$

is nonempty and bounded.

THEOREM 3.4. (Duffin[1]). *Assume that the program* (P) *is canonical. If either of the programs* (P) *or* (D) *has a feasible solution then MP = MD with MP attained.*

For the benefit of the readers who are interested in abstract spaces, we now note that Theorem 3.4 holds in the case that the domain of each of the functions $f, g_1, \ldots, g_m$ is a reflexive Banach space(see [1]) and in the case that f and each g_i are continuous on a closed convex subset of the reflexive Banach space, perhaps taking the value $+\infty$ on the boundary of the set (see [2]). Finally, we note that in finite dimensions Theorem 3.4 holds as long as the functions f and all the g_i are lower semi-continuous [7].

4. Duality results

Given the above notation, assumptions and constructions, we are now ready to develop a duality theory for the Apex dual (A). As a first step, we prove the following type of accessibility lemma for Apex programs.

LEMMA 4.1. *If λ_A is feasible for the Apex dual and λ_D is feasible for the Lagrangian dual then λ is feasible for the Apex dual whenever $\lambda \in [\lambda_A, \lambda_D)$.*

PROOF. Let $0 < \varepsilon < 1$ and let $\lambda = (1 - \varepsilon)\lambda_D + \varepsilon\lambda_A$. The result follows from Corollary 3.3 once we show that $L(x, \lambda)$ is unbounded at infinity.

Let (x_k) be a sequence in R^n with $\|x_k\| \to +\infty$. Then for each k,

$$L(x_k, \lambda) = (1 - \varepsilon)L(x_k, \lambda_D) + \varepsilon L(x_k, \lambda_A)$$
$$\geqslant (1 - \varepsilon)\inf_x L(x, \lambda_D) + \varepsilon L(x_k, \lambda_A).$$

Now $\inf_x L(x, \lambda_D)$ is finite since λ_D is feasible for the Lagrangian dual. In addition, $L(x_k, \lambda_A) \to +\infty$ as $k \to +\infty$ by Corollary 3.3 since λ_A is feasible for the Apex dual. Thus the above inequality implies that $L(x_k, \lambda) \to +\infty$ as $k \to +\infty$. Hence $L(x, \lambda)$ is unbounded at infinity and as reasoned above, λ is feasible for the Apex dual.

The following corollary is now immediate. In its statement $ri(\ )$ denotes the relative interior of the set in $(\ )$.

COROLLARY 4.2. *If the Apex dual is feasible, then* $ri(\text{dom}(D)) \subseteq \text{dom}(A) \subseteq \text{dom}(D)$.

PROOF. Let λ_A be feasible for the Apex dual, that is $\lambda_A \in \text{dom}(A)$. Then $\lambda_A \in \text{dom}(D)$ by construction (3.1). Thus $\text{dom}(A) \subseteq \text{dom}(D)$.

Let $\lambda_D \in ri(\text{dom}(D))$. Then, by definition, there exists an $\varepsilon > 0$ with $[\lambda_A, \lambda_D + \varepsilon(\lambda_D - \lambda_A)] \in \text{dom}(D)$. Lemma 4.1 now implies that $\lambda_D \in \text{dom}(A)$. Hence, $ri(\text{dom}(D)) \subseteq \text{dom}(A)$.

The above corollary implies that whenever the Apex dual is feasible, the feasible regions for the Apex and Lagrangian duals differ at most on the boundary of the feasible region of the Lagrangian. As the next example illustrates, this difference can indeed occur.

EXAMPLE 4. Let $f(x) = (1 + x^2)^{1/2}$ and let $g(x) = -(x^2 - 1)^{1/2}$ for $x \geqslant 1$; $+\infty$ otherwise. Then a straightforward calculation shows that $\inf_x L(x, \lambda) = (2 - 2\lambda^2)^{1/2}$ for $0 \leqslant \lambda \leqslant 1$; $-\infty$ otherwise, with attainment of the infimum at $x = [(1 + \lambda^2)/(1 - \lambda^2)]^{1/2}$ for $0 \leqslant \lambda < 1$. Since this attainment is unique, Proposition 3.1 implies that λ is feasible for the Apex dual whenever $0 \leqslant \lambda < 1$. However, since $\inf_x L(x, 1)$, which is 0, is not attained, Proposition 3.1 implies that $\lambda = 1$ is not feasible for the Apex dual. In other words, $\lambda = 1$ is feasible for the Lagrangian dual but not the Apex dual.

In the example given above, it is easy to show that $MA = MD = MP$. As Theorem 4.4 below shows, this is not surprising. However before proving Theorem 4.4 we need the following technical result:

LEMMA 4.3. *Assume that the Apex dual is feasible. Let* (x_k) *be a sequence in the set* $\bigcap_{i=1}^m \{x \mid g_i(x) \leqslant \delta\}$ *where* $\delta > 0$ *with* $\|x_k\| \to +\infty$ *as* $k \to +\infty$. *Then* $f(x_k) \to +\infty$ *as* $k \to +\infty$.

PROOF. Let the sequence (x_k) be as above. Since the Apex dual is feasible, Corollary 3.3 imply that there exists a $\lambda \geqslant 0$ with

$$f(x_k) + \sum_{i=1}^{m} \lambda_i g_i(x_k) \to +\infty \qquad \text{as } k \to +\infty.$$

Now for each k,

$$f(x_k) + \sum_{i=1}^{m} \lambda_i g_i(x_k) \leqslant f(x_k) + \delta \sum_{i=1}^{m} \lambda_i$$

since each $\lambda_i \geqslant 0$. It now follows from the above that $f(x_k) \to +\infty$ as $k \to +\infty$, since $\delta \sum_{i=1}^{m} \lambda_i$ is a constant.

We now note that the above lemma implies that if the Apex dual is feasible, then the program (P) is canonical. This observation will be important in the proof of the next theorem.

THEOREM 4.4. *If the Apex dual is feasible, then $MA = MD = MP$ with MP attained.*

PROOF. As noted in the discussion following Lemma 4.3, the program (P) is canonical since the Apex dual is feasible. Moreover, since the Apex dual is feasible, Lemma 3.2 and Corollary 3.3 imply that (D) has a feasible solution. Thus Theorem 3.4 implies that $MP = MD$ with MP attained. The result will follow once we show that $MA = MD$.

Let λ_A be feasible for (A) and let (λ_m) be a maximizing sequence for (D). Let m be arbitrary but fixed. Since $\theta(\lambda) = \inf_x L(x, \lambda)$ is a concave function, there exists a $\lambda(m) \in [\lambda_A, \lambda_m)$ such that $\theta(\lambda(m)) \geqslant \theta(\lambda_m) - 1/m$. Now, Lemma 4.1 implies that $\lambda(m)$ is feasible for (A) which in turn implies, by Theorem 2.1, that $MA_{\lambda(m)} = \theta(\lambda(m))$. Since by construction $\theta(\lambda_m) \to MD$, the above inequality implies that $MA_{\lambda(m)} \to N \geqslant MD$. Since, for each m, $MA_{\lambda(m)} \leqslant MA$, it now follows that $MA \geqslant MD$. But by construction (3.2), $MA \leqslant MD$. Thus $MA = MD$ and the result follows as reasoned above.

If the Wolfe dual (W) to (P) is used as a dual, then we can prove results similar to Corollary 4.2 and Theorem 4.4 for the feasible regions of the Apex and Wolfe duals and their respective optimal values. This follows, for instance, from the fact that

$$\text{dom}(A) \subseteq \text{dom}(W) \subseteq \text{dom}(D).$$

To see the first inclusion, let λ be feasible for Apex dual. Then by Lemma 3.4 and Corollary 3.3 there exists an $x(\lambda)$ with $L(x(\lambda), \lambda) = \inf_x L(x, \lambda)$. Moreover

$$0 = \nabla f(x(\lambda)) + \sum_{i=1}^{m} \lambda_i \nabla g_i(x(\lambda))$$

where ∇f is the gradient of f. In other words, λ is feasible for the Wolfe dual.

Now the above inclusions in conjunction with Corollary 4.2 imply that if the Apex dual is feasible, then the feasible regions of the Wolfe and Lagrangian dual differ on at most the boundary of the feasible region of the Lagrangian dual. In addition, the above inclusions in conjunction with Theorem 4.4 imply that if the Apex dual is feasible, then $MA = MW = MD = MP$. This extends the duality results derived by Falk in [6] concerning the Wolfe dual (see for example [9]).

5. Computational remarks

We conclude with a few remarks outlining the implementation of the Apex framework on a personal computer.

As suggested in [5], first find a feasible solution to the Apex dual problem by means of a grid or Monte Carlo search. Include in this search the multiplier λ as a variable. Once a feasible solution and multiplier are found, fix the multiplier λ and generate the approximation to the Lagrangian $L(x, \lambda)$ by the iterative scheme suggested at the end of Section 2.

Finally, update the multiplier by either a penalty or first order multiplier update as discussed, for instance, in [8].

It follows from Theorem 4.4 that the above algorithm will converge to the optimal value of program (P) as long as the Apex dual is feasible. In addition, since the above algorithm can be implemented on a personal computer, the pedagogical benefits noted in Section 2 for the unconstrained case can now also be realized in the constrained case.

Acknowledgements

References

[1] R. J. Duffin, "Lagrange multiplier method for convex programs", *Proc. Nat. Acad. Sci. USA* 72 (1975), 1778–1781.
[2] R. J. Duffin, "Convex programs having some linear constraints", *Proc. Nat. Acad. Sci. USA* 74 (1977), 26–28.

[3] R. J. Duffin, "Transformations that aid numerical solution of nonlinear programs", *Opsearch* 18 (1981), 158–166.

[4] R. J. Duffin, "Numerical estimation of optima by use of dual inequalities", in *Semi-infinite programming and applications* (eds. A. V. Fiacco and K. O. Kortanek), (Springer-Verlag, Berlin, 1983), 118–127.

[5] R. J. Duffin, "Duality inequalities as a numerical aid", *Oper. Res. Lett.* 3 (1984), 65–68.

[6] J. E. Falk, "Lagrange multipliers and nonlinear programming", *J. Math. Anal. Appl.* 19 (1967), 141–159.

[7] D. F. Karney, "Asymptotic convex programming", Georgia Institute of Technology, Management Science Technical Report MS-81-9 (1981).

[8] D. G. Luenberger, *Linear and nonlinear programming*, 2nd ed., (Addison-Wesley, Reading, MA, 1984).

[9] O. L. Mangasarian, *Nonlinear programming* (McGraw-Hill, New York, 1969).