A NOTE ON NONDIFFERENTIABLE SYMMETRIC DUALITY

B. D. CRAVEN[1]

Abstract

Under suitable hypotheses on the function f, the two constrained minimization problems:

$\mathrm{MIN} f - f_y y$ subject to $x \geqslant 0$, $-f_y \geqslant 0$; $\mathrm{MAX} f - f_x x$ subject to $y \geqslant 0$, $f_x \geqslant 0$;

are well known each to be dual to the other. This *symmetric duality* result is now extended to a class of nonsmooth problems, assuming some convexity hypotheses. The first problem is generalized to:

$$\mathrm{MIN} f(x, y) - p^T y \quad \text{subject to } x \in T, \ -p^T \in S^* \cap \partial_y (-f)(x, y),$$

in which T and S are convex cones, S^* is the dual cone of S, and ∂_y denotes the subdifferential with respect to y. The usual method of proof uses second derivatives, which are no longer available. Therefore a different method is used, where a nonsmooth problem is approximated by a sequence of smooth problems. This duality result confirms a conjecture by Chandra, which had previously been proved only in special cases.

1. Introduction

Let $T \subset \mathbf{R}^n$ and $S \subset \mathbf{R}^m$ be closed convex cones; let $f: \mathbf{R}^n \times \mathbf{R}^m \to \mathbf{R}$ be a twice differentiable function, such that $f(\cdot, y)$ is convex on T for each $y \in S$, and $-f(x, \cdot)$ is convex on S for each $x \in T$. Writing f for $f(x, y)$ and taking vectors in the dual cones S^* and T^* as row vectors, consider the pair of problems

$$(\mathrm{PS}): \text{Minimize } f - f_y y \quad \text{subject to } x \in T, \ -f_y^T \in S^*;$$

$$(\mathrm{DS}): \text{Maximize } f - f_x x \quad \text{subject to } y \in S, \ f_x^T \in T^*.$$

Here f_x and f_y denote partial derivatives. Under appropriate "constraint qualification" hypotheses, that $\operatorname{int} S^* \neq \varnothing$ (or S a polyhedral cone) and that the Hessian matrix $f_{xx}(x^*, y^*)$ is nonsingular at a minimum (x^*, y^*) for (P), then

---

[1] Mathematics Department, University of Melbourne, Parkville, Victoria 3052

30

(see [3], Chapter 6) (DS) is a *strong dual* problem to (PS); thus weak duality holds (for feasible points, the object function for (PS) $\geq$ the objective function for (DS)), and (x^*, y^*) also maximizes (DS), the two objective functions being equal at this point (the "zero duality gap" property). The proof in [3] assumes an additional "closed-cone hypothesis", which appears superfluous for the linear constraint $x \in T$; and the proof of weak duality is deficient, but may easily be mended. The closed-cone hypothesis in [3] could be substituted for the hypothesis int $S^* \neq \varnothing$. If f is quadratic, then (PS) and (DS) have linear constraints, for which no Hessian hypothesis is relevant, or needed. But that well-known case does not contribute towards a nondifferentiable generalization.

Consider now a nondifferentiable version of problems (PS) and (DS), in which f is assumed locally Lipschitz, and satisfying the above convexity hypotheses (thus $f(\cdot, y)$ is convex on T and $-f(x, \cdot)$ is convex on S). Then f is differentiable almost everywhere, but second derivatives need not exist at all. The derivatives f_x^T and f_y^T are now replaced by the subdifferentials $\partial_x f(x, y)$ and $-\partial_y(-f)(x, y)$. The problems (PS) and (DS) now become:

(NPS): Minimize $f(x, y) - p^T y$ subject to $x \in T$, $-p^T \in S^* \cap \partial_y(-f)(x, y)$;

(NPD): Maximize $f(x, y) - q^T x$ subject to $y \in S$, $q^T \in T^* \cap \partial_x f(x, y)$.

This pair of problems is related to the pair proposed in [1], where weak duality was proved under convex/concave hypotheses, but strong duality could only be proved in a special case, where the subdifferentials could be represented so as to make the problem (NPD) differentiable, so that F. John conditions would hold at its minimum (x^*, y^*). *In the present paper, duality will be shown to hold generally, without any need for a differentiable representation.*

To prove weak duality, let (x, y) be feasible for (NPS), and let (u, v) be feasible for (NDS). Then

$$
\begin{aligned}
\big[f(x, y) - p^T y \big] &- \big[f(u, v) - q^T u \big] \\
&= \big[f(x, y) - f(x, v) + p^T(v - y) \big] + \big[f(x, v) - f(u, v) - q^T(x - u) \big] \\
&\quad + \big[-p^T v \big] + \big[q^T x \big] \\
&\geq 0 + 0 + 0 + 0 = 0.
\end{aligned}
$$

Let int denote interior, and let co denote convex hull and co$^-$ closed convex hull.

2. Strong duality for nonsmooth problems

Let $z = (x, y) \in \mathbf{R}^m \times \mathbf{R}^n$; assume (provisionally) that (NPS) reaches a strict local minimum at $(x, y, p) = (x^*, y^*, p^*)$; let $z^* = (x^*, y^*)$. For $\eta > 0$, let B_η denote the closed unit ball in $\mathbf{R}^{m+n}$ with centre 0 and radius η. Let $F(z) := f(x, y)$.

Following the method of [4], a smooth (at least C^2) function $F(\cdot : \eta)$ is defined, corresponding to a locally Lipschitz function F, by

$$F(z:\eta):= \int_{R^{m+n}} F(z-w)\psi(w)\, dw,$$

where ψ is a non-negative C^2 weighting function with support in B_η satisfying $\int_{B_\eta} \psi(t)\, dt = 1$, where dt denotes Lebesgue measure. Then [4]

$$F'(z^* : \eta) = \int F'(z^* - w)\psi(w)\, dw \in C(\eta)$$

$$:= \mathrm{co}^- \{ F'(z^* + s): \|s\| \leqslant \eta, \exists F'(z^* + s) \},$$

noting that, by Rademacher's theorem, F is Fréchet differentiable except on a set of zero measure, which does not affect the integral; and it follows that the Clarke generalized subdifferential (see [2]) $\partial F(z^*) = \bigcap_{\eta > 0} C(\eta)$, and is nonempty (see [4], and [2, Proposition 5]). Then (NPS) and (NDS) lead to the corrresponding smoothed problems:

$$(\mathrm{NP}:\eta): \text{Minimize } H(z, p:\eta):= F(z:\eta) - p^T y$$
$$\text{subject to } x \in T, \; p = F_y(z:\eta) \in -S^* + b'_\eta;$$
$$(\mathrm{ND}:\eta): \text{Maximize } K(z,q:\eta):= F(z:\eta) - q^T x$$
$$\text{subject to } y \in S, \; q^T = F_x(z:\eta) \in T^* + b''_\eta.$$

Here $z = (x, y)$, F_x and F_y denote partial derivatives with respect to x and y, and the vectors $b'_\eta = O(\eta)$ and $b''_\eta = O(\eta)$ (as $\eta \downarrow 0$) are defined using the above expression for $C(\eta)$, to ensure that the smoothed problems are feasible (compare [4, Lemma 1]). Note that $\psi(\cdot)$ depends parametrically on η.

Then "zero duality gap" will hold for the smoothed problems, by an existing symmetric duality theorem, noting that the convex properties assumed for f carry over to the smoothed version of this function. A limiting process is then required, to deduce "zero duality gap" for the given pair of nonsmooth problems. This is done in the following Theorem, by selecting suitable convergent subsequences, and choosing suitable functions $\psi(\cdot)$, depending on various values of η.

THEOREM 1. *For problems* (NPS) *and* (NPD), *assume that S and T are closed convex cones, $f: \mathbf{R}^m \times \mathbf{R}^n$ is locally Lipschitz, $f(\cdot, y)$ is convex on T for each $y \in S$, $-f(x, \cdot)$ is convex on S for each $x \in T$, int $S \neq \varnothing$ (or S is polyhedral). Let* (NPS) *attain a local minimum at $(x, y) = (x^*, y^*)$; and assume that, for sufficiently small $\eta > 0$,*

$$\varnothing \neq \partial_y f(x^*, y^*)$$
$$\cap \mathrm{int\, co}\{ f_y(x, y); \|(x, y) - (x^*, y^*)\| \leqslant \eta, (x, y) \in D \}, \qquad (Q)$$

where D denotes the set of points where f is Fréchet differentiable. Then (NDS) *is a strong dual problem to* (NPS).

PROOF. Weak duality is proved as in Section 1.

To prove "zero duality gap", assume provisionally that (NPS) reaches a strict local minimum at $(z, p) = (z^*, p^*) \equiv (x^*, y^*, p^*)$. Choose $\eta > 0$ sufficiently small that F is Lipschitz on $z^* + B_\eta$. Define $C(\eta)$ as above, then $F'(z^*; \eta) \in C(\eta)$; and similarly define

$$
\begin{aligned}
C_x(\eta) &:= \mathrm{co}^- \big\{ F_x(z^* + s) \colon \|s\| \leqslant \eta,\ \exists F'(z^* + s) \big\}, \\
C_y(\eta) &:= \mathrm{co}^- \big\{ F_y(z^* + s) \colon \|s\| \leqslant \eta,\ \exists F'(z^* + s) \big\}.
\end{aligned}
$$

Then [4, Lemma 1] shows that, for each $\varepsilon > 0$, there exists $\delta > 0$ for which the restriction $\{z \in z^* + B_\delta \colon x \in T,\ F_y(z \colon \eta) \in -S^* + b_\eta'\}$ of the feasible set of $(NP \colon \eta)$ to a δ-ball about z^* lies in an ε-neighborhood of the set $\{z \in z^* + B_\delta \colon x \in T,\ S^* \cap \partial_y(-F)(z) \neq \varnothing\}$, which is the corresponding restriction of the feasible set for the given problem (NPS). (The cited Lemma proves this for a constraint $p(z) \in K$, where K is a convex cone; here $p(x \colon \eta)$ is replaced by $F_x(z \colon \eta)$ or $F_y(z \colon \eta)$. The proof uses compactness.) A similar statement also holds, relating to $(ND \colon \eta)$ and (NDS). The Clarke generalized subdifferentials may here be equated to convex subdifferentials, in view of the convex hypotheses. Now [4, Lemma 2] shows that, given the above ε-neighborhood property, and the strict local minimum of (NPS), there exists a sequence $\{\eta_k\} \downarrow 0$, and corresponding weighting functions ψ_k (with support of ψ_k in the ball with radius η_k), for which the problem $(NP \colon \eta_k)$ attains a local minimum at some point z_k^*, such that $\{z_k^*\} \to z^*$ as $k \to \infty$, and $F_y(z_k^* \colon \eta_k)$ lies within distance ε of $S^* \cap \partial_y F(z^*)$ whenever k is sufficiently large.

From hypothesis (Q), it follows (see the Appendix) that the Hessian matrix $H_{yy}(\cdot \colon \eta)$ is nondegenerate (thus, has no zero eigenvalues) when η is sufficiently small, and so is nonsingular. The symmetric duality theorem for smooth problems, cited in the Introduction, then applies to the smoothed problems $(NP \colon \eta)$ and $(ND \colon \eta)$, showing that $(ND \colon \eta)$ is a strong dual problem to $(NP \colon \eta)$. Thus, to a minimum (z_k^*, p_k^*) of $(NP \colon \eta_k)$ there corresponds a maximum (z_k^*, q_k^*) of $(ND \colon \eta_k)$, with equal objective values.

By choosing successive subsequences of $\{z_k^*\}$, $\{F_y\{z_k^* \colon \eta_k\}\}$, $\{F_x(z_k^* \colon \eta_k)\}$, a subsequence $\{\eta_k \colon k \subset J\}$ (where $J \subset \mathbf{N}$) may be found, such that $\{z_k^*\} \to z^*$, $\{F_y(z_k^* \colon \eta_k)\} \to p^*$, $\{F_x(z_k^* \colon \eta_k)\} \to q^*$, with these limits satisfying $-p^* \in S^* \cap \partial_y(-F)(z^*)$ and $q^* \in T^* \cap \partial_x F(z^*)$. The local Lipschitz hypothesis ensures that these gradients lie in compact regions, ensuring that subsequence limits exist. Note that $-p^* \in S^*$ follows since S^* is closed and $-p_k = -F_x(z_k^* \colon \eta_k) \in S^*$, and

similarly $q^* \in T^*$. And $-p^* \in \partial_y(-F)(z^*)$ and $q^* \in \partial_x F(z^*)$ follow from

$$\partial_x F(z^*) = \bigcap_{\eta>0} C_x(\eta) \quad \text{and} \quad \partial_y(-F)(z^*) = \bigcap_{\eta>0} C_y(z^*).$$

Now assume that (NPS) reaches a local minimum at (z^*, p^*), but not necessarily a strict local minimum. To deal with this, add a term $\theta \|z - z^*\|^2$ to $F(z)$, where θ is a small positive parameter. The problem (NPS), so modified, has now a strict local minimum at z^*. (Convexity of $-f(x, \cdot)$ is upset, and must be restored later in the proof.) The previous proof now shows that "zero duality gap" holds between the modified (NPS) and the corresponding modified (NPD). Consider now a sequence of values of θ, say $\{\theta_j\}\downarrow 0$. By selecting a suitable subsequence, similarly to the subsequence argument used before, it may be assumed that the corresponding p_j and q_j tend to limits, $p^\#$ and $q^\#$ say. Then $(z^*, p^\#)$ and $(z^*, q^\#)$ are optimal for the given problems (NPS) and (NDS), and their objective functions are equal. (Convexity is only required for weak duality, already proved for these problems.)

3. Discussion

The essential subsequence arguments in the above proof use compactness of the unit sphere, hence must assume finite dimensional spaces. Note also that the proof of strong symmetric duality uses Fritz John, rather than Kuhn Tucker, conditions, hence does not need an additional constraint qualification. In (Q), int co can be written in place of int co$^-$, since these two are equal, for a nonempty set [5]. For (NPS) to be a dual to (NDP), ∂_y and f_y in (Q) are replaced by ∂_x and $\hat{f}_x$.

Appendix

Let $h: \mathbf{R}^k \to \mathbf{R}$ be locally Lipschitz, and satisfy $0 \in \partial h(a)$. Define $h(\cdot : \eta)$ from h in the same way as $f(\cdot : \eta)$ was defined above from f, with ψ depending parametrically on $\eta > 0$. Define

$$C^h(x, \eta) := \mathrm{co}^-\{h'(x+s): \|s\| \leqslant \eta, \exists h'(x+s)\};$$

$$Q(x, \eta) := \mathrm{co}\{h'(x:\eta): \mathrm{supp}\,\psi \subset B_\eta\}.$$

Since each point in $C^h(\eta)$ is a sequential limit of values of $\int h(x-s)\psi_j(s)\,ds$, the closure of $Q(x, \eta)$ equals $C^h(x, \eta)$. Since $0 \in \partial h(a)$, there exist points $e_i \in H(a, \eta)$ (say, at the vertices of a hypercube) such that $0 \in \mathrm{int}\,\mathrm{co}\{e_i: i = 1, 2, \ldots, 2^k\}$. So, by convexity, a neighbourhood of 0 is contained in $H(a; \eta)$.

Hence $h'(x^\# : \eta) = 0$ for some $x^\#$ for which $\|x^\# - a\| \leqslant \eta$. For $\|x - a\|$ sufficiently small,

$$h(x + y : \eta) = h(x^\#) + y^T A y/2 + o(\|y\|^2),$$

where the Hessian A is nondegenerate (thus, no zero eigenvalues) if and only if $0 \in \mathrm{int}\{y^T A : \|y\| < \delta\}$ for some $\delta > 0$, and that is implied by $0 \in \mathrm{int}\, C^h(a, \eta)$, noting that $y^T A$ differs from $h'(x^\# + y : \eta)$ by $o(\|y\|^2)$, and that $0 \in \mathrm{int}\, Q(a, \eta)$, which implies $0 \in \mathrm{int}\, Q(x, \eta)$ when $\|x - a\|$ is sufficiently small.

Now omit the assumption that $0 \in \partial h(a)$, and assume instead that

$$\varnothing \neq \partial h(a) \cap \mathrm{int}\, C^h(\eta)\big(= \partial h(a) \cap \mathrm{int}\, \mathrm{co}\{h'(x + s): \|s\| \leqslant \eta, \exists h'(x + s)\}\big),$$

for some $\eta > 0$. Thus, for some $c^T \in \partial h(a)$, defining g by $g(x) := f(x) - c^T x$, there follows $0 \in \partial g(a)$ and $0 \in \mathrm{int}\, C^g(a)$. Therefore the Hessian of f (or g) at A is nondegenerate, by the previous paragraph.

Acknowledgement

References

[1] S. Chandra, B. D. Craven and B. Mond, "Symmetric dual fractional programming," *Z. Oper. Res. Ser. Theory*, 29 (1985), 59–64.
[2] F. H. Clarke, "A new approach to Lagrange multipliers," *Math. Oper. Res.* 1 (1976), 165–174.
[3] B. D. Craven, *Mathematical programming and control theory*, (Chapman and Hall, London, 1978).
[4] B. D. Craven, "Nondifferentiable optimization by smooth approximation," Melbourne University, Mathematics Research Report 32 (1984), to appear in *Optimization* (1986).
[5] B. D. Craven and B. Mond, "Transposition theorems for cone-convex functions," *SIAM J. Appl. Math.* 24 (1973), 603–612.