Lagrangian duality and perturbational duality I *

Erik J. Balder

Our approach to the Karush-Kuhn-Tucker theorem in [OSC] was entirely based on subdifferential calculus (essentially, it was an outgrowth of the two subdifferential calculus rules contained in the Fenchel-Moreau and Dubovitskii-Milyutin theorems, i.e., Theorems 2.9 and 2.17 of [OSC]). On the other hand, Proposition B.4(v) in [OSC] gives an intimate connection between the subdifferential of a function and the Fenchel conjugate of that function. In the present set of lecture notes this connection forms the central analytical tool by which one can study the connections between an optimization problem and its so-called *dual* optimization problem (such connections are commonly known as *duality* relations). We shall first study duality for the convex optimization problem that figured in our Karush-Kuhn-Tucker results. In this simple form such duality is known as *Lagrangian duality.* Next, in section 2 this is followed by a far-reaching extension of duality to abstract optimization problems, which leads to *duality-stability* relationships. Then, in section 3 we specialize duality to optimization problems with cone-type constraints, which includes *Fenchel duality* for semidefinite programming problems.

1 Lagrangian duality

An interesting and useful interpretation of the KKT theorem can be obtained in terms of the so-called *duality* principle (or relationships) for convex optimization. Recall our standard convex minimization problem as we had it in [OSC]:

$$(P)\quad \inf_{x\in S}\{f(x) : g_1(x)\leq 0,\cdots,g_m(x)\leq 0, Ax-b=0\}$$

and recall that we allow the functions $f, g_1, \ldots, g_m$ on $\mathbb{R}^n$ to have values in $(-\infty,+\infty]$. Recall also that the feasible set of (P) is sometimes denoted by Z. We assume that the set of all feasible solutions is nonempty (in [OSC] that was not necessary, because of the role assigned to $\bar{x}$ in KKT-results). Let $\inf(P)$ denote the value of the above infimum in (P). We shall assume that $\inf(P)$ is not equal to $+\infty$; this excludes the trivial case where f is identically equal to $+\infty$ on Z. We can associate to (P) the following maximization problem:

$$(D)\quad \sup_{(u,v)\in\mathbb{R}^m_+\times\mathbb{R}^p}\theta(u,v),$$

*LNMB Ph.D. course "Convex Analysis for Optimization", October, 2010. These notes grew from earlier Ph.D. courses in Utrecht and Naples; the author is indebted to participants in those courses for helpful comments.

whose objective function is defined by

$$\theta(u,v) := \inf_{x\in S}[f(x) + \sum_{i=1}^{m} u_i g_i(x) + v^t(Ax-b)].$$

In this section we shall refer to (D) as the *(Lagrangian) dual* optimization problem,[1] associated to (P). Together with giving this name to (D), it is customary to refer to the original optimization problem (P) as the *primal* optimization problem. The function $\theta(u,v)$ is concave in (u,v), regardless of any convexity properties of (P). Also, θ cannot be equal to $+\infty$ anywhere, but it may take the value $-\infty$.

Exercise 1.1 a. Prove that the dual objective function $\theta(u,v)$ is concave in (u,v).
b. Prove that θ cannot be equal to $+\infty$ anywhere.

Example 1.1 The standard linear programming problem (P) is as follows:

$$(P)\quad \inf_{x\in S}\{c^t x : Ax - b = 0\},$$

where $S := \mathbb{R}^n_+$. Let us find the corresponding dual problem (D). Here (P) is a problem with no explicit inequality constraints; we can therefore do one of two things: (1) introduce the purely cosmetic inequality constraint $g_1(x) \le 0$ with $g_1 \equiv 0$ or (2) re-read the derivation of Corollary 3.5 in [OSC], this time based on using Theorem 2.10 instead of Theorem 3.1.[2] Both lead to essentially the same result (check this), but we prefer (2) because it dispenses with an otherwise purely cosmetic multiplier u_1. We have

$$\theta(v) := \inf_{x\ge 0}[c^t x + v^t(Ax-b)] = \begin{cases} -b^t v & \text{if } c + A^t v \ge 0 \\ -\infty & \text{otherwise} \end{cases}$$

Thus, writing $w := -v$, we can state (D) as follows:

$$(D)\quad \sup_{w\in\mathbb{R}^p}\{b^t w : A^t w \le c\}.$$

This is the familiar form of the dual problem in linear programming.

Example 1.2 Let A be a $p \times n$-matrix of rank $p < n$ and suppose $b \in \mathbb{R}^p\backslash\{0\}$. For $L := \{x \in \mathbb{R}^n : Ax = b\}$ as the feasible set (supposed to be nonempty) consider the following minimum norm optimization problem:

$$(P)\quad \inf_{x\in\mathbb{R}^n}\{|x|^2 : Ax - b = 0\}.$$

a. Before formulating the dual problem, we wish to solve (P) by means of the KKT theorem, because, unlike Example 1.1, this can be done explicitly. Just as

[1] The expression $f(x) + \sum_i u_i g_i(x) + v^t(Ax-b)$ is often called the *Lagrangian* of (P).

[2] Actually, in the absence of inequality constraints the "small" KKT theorem Theorem 2.10 could well be used to deal directly with problems without inquality constraints, provided that one borrows from Corollary 3.5 the identity stated in the first line of its proof; cf. Example 1.2.

in the previous example, problem (P) only has equality constraints and we treat it similarly to Example 1.1 by ignoring the inequality multiplier u. We set $f(x) := |x|^2$ and $S := \mathbb{R}^n$. Then S is open, and regularity comes down to $L = S \cap L \neq \emptyset$, i.e., to nonemptiness of L. So regularity holds and, since there are no inequality constraints, normality in the Lagrangian inclusion holds automatically. Hence, by the KKT theorem in [OSC] a vector $\bar{x} \in L$ is optimal if and only if there exists $\bar{v} \in \mathbb{R}^p$ such that

$$0 \in \partial f(\bar{x}) + A^t \bar{v},$$

where $\partial f(\bar{x}) = \{\nabla f(\bar{x})\} = 2\bar{x}$. Hence, the above Lagrangian inclusion gives $\bar{x} = -\frac{1}{2}A^t\bar{v}$. To determine the unknown multiplier $\bar{v}$, we apply A to both sides of the latter identity. By $\bar{x} \in L$ it gives $b = -\frac{1}{2}AA^t\bar{v}$, so $\bar{v} = -2(AA^t)^{-1}b$ (notice that AA^t is a regular $p \times p$ matrix by the original full rank condition for A). When this is substituted back into $\bar{x} = -\frac{1}{2}A^t\bar{v}$, it gives $\bar{x} = A^t(AA^t)^{-1}b$, an expression familiar to students of *regression theory* in statistics. We conclude that $\bar{x} = A^t(AA^t)^{-1}b$ is the (unique) optimal solution of (P).

b. Next, we determine the Lagrangian dual optimization problem associated to (P). Problem (P) has only equality constraints, so the dual objective function is given by

$$\theta(v) := \inf_{x \in S} [|x|^2 + v^t(Ax - b)].$$

This is an unconstrained infimum of a convex function, so setting the gradient of $x \mapsto |x|^2 + v^t(Ax - b)$ equal to zero (i.e., $2x + A^t v = 0$) gives the minimum value, and we find $\theta(v) = -v^t AA^t v/4 - b^t v$. For the Lagrangian dual problem this gives

$$(D) \quad \sup_{v \in \mathbb{R}^p} -v^t AA^t v/4 - b^t v.$$

Incidentally, we can solve this optimization problem immediately: setting the gradient of the concave function θ equal to zero gives $-AA^t v/2 - b = 0$; hence $\bar{v} := -2(AA^t)^{-1}b$ is the optimal solution of (D).

Exercise 1.2 The optimization problem studied in Example 1.2 can be equivalently written as

$$(P) \quad \inf_{x \in S} \{|x| : Ax - b = 0\},$$

with $S := \mathbb{R}^n$ and $f(x) := |x|$. Similar to what was done in that example, derive first the optimal solution for the optimization problem in this form. *Hint 1:* first demonstrate the following for $f(x) := |x|$:

$$\partial f(\bar{x}) = \begin{cases} \{\frac{\bar{x}}{|\bar{x}|}\} & \text{if } \bar{x} \neq 0 \\ \{x \in \mathbb{R}^n : |x| \leq 1\} & \text{if } \bar{x} = 0. \end{cases}$$

Next, show that the corresponding dual problem (D) is given by

$$\sup_{v \in \mathbb{R}^p} \{-b^t v : |A^t v| \leq 1\},$$

and thus it is completely different from (D) in the previous example, even though the primal (P)'s are equivalent. *Hint:* Prove first by the Cauchy-Schwarz inequality that

$$\inf_{x\in S}[|x| + w^t x] = \begin{cases} 0 & \text{if } |w| \leq 1 \\ -\infty & \text{otherwise} \end{cases}$$

Finally, prove that the optimal solution of the present dual problem (D) is given by

$$\bar{v} = \frac{-(AA^t)^{-1}b}{\sqrt{b^t(AA^t)^{-1}b}}.$$

Exercise 1.3 The optimization problem studied in Example 1.2 can also be equivalently written as

$$(P) \quad \inf_{x\in S} |x|,$$

with $S := L$ and $f(x) := |x|$. This problem has no explicit inequality or equality constraints. Derive its optimal solution by means of the "small" KKT Theorem 2.10; see also footnote 2. Propose also a suitable Lagrangian dual problem that is consistent with the general formula for θ by introducing the artificial constraint function $g_1 \equiv 0$.

The key to the Lagrangian duality interpretation of the KKT theorem of [OSC] is contained in the following exercise.

Exercise 1.4 Let $\bar{x}$ be feasible for (P) and let $(\bar{u}, \bar{v}) \in \mathbb{R}^m_+ \times \mathbb{R}^n$. In the context of [OSC], we indicate by (**OPT**) the statement that $\bar{x}$ is optimal solution of (P) (i.e., $\bar{x}$ is feasible for (P) and $f(\bar{x}) = \inf(P)$). Let (**CS**) abbreviate the complementary slackness relations and let (**MP**) the minimum principle in terms of the same $\bar{x}$ and multiplier vectors $\bar{u}$ and $\bar{v}$, i.e.,

$$\bar{x} \in \text{argmin}_{x\in S}[f(x) + \sum_{i=1}^m \bar{u}_i g_i(x) + \bar{v}^t(Ax - b)]$$

a. Prove the equivalence

$$\theta(\bar{u}, \bar{v}) = f(\bar{x}) \Leftrightarrow (\mathbf{CS}) + (\mathbf{MP}).$$

b. Use part a to prove that the KKT theorem can now be restated in the following form

$$(\mathbf{OPT}) \;\; \Leftrightarrow \;\; \exists_{\bar{u}\geq 0,\bar{v}}\theta(\bar{u}, \bar{v}) = f(\bar{x}),$$

provided that both the regularity condition and Slater's constraint qualification hold.

We can now state the main Lagrangian duality result. Like the KKT theorem in [OSC] it has an elementary part (i) and a deeper part (ii), the latter part being based on separating hyperplane theorems (or its proxies).

Theorem 1.3 (Lagrangian duality) (i) *For every $x \in Z$ and every $(u,v) \in \mathbb{R}^m_+ \times \mathbb{R}^p$*

$$\theta(u,v) \leq f(x) \text{ (weak duality).}$$

In particular, if

$$\theta(\bar{u},\bar{v}) = f(\bar{x})$$

for some $\bar{x} \in Z$ and some $(\bar{u},\bar{v}) \in \mathbb{R}^m_+ \times \mathbb{R}^p$, then

$$\bar{x} \text{ is optimal for } (P) \text{ and } (\bar{u},\bar{v}) \text{ is optimal for } (D).$$

and the complementary slackness relationship then also holds for $\bar{x}$.

(ii) *Conversely, if $\bar{x}$ is an optimal solution of (P) and if both the regularity condition and Slater's constraint qualification hold, then there exists $(\bar{u},\bar{v}) \in \mathbb{R}^m_+ \times \mathbb{R}^p$ such that*

$$\theta(\bar{u},\bar{v}) = f(\bar{x}) \text{ (strong duality).}$$

PROOF. (i) If $x \in S$ meets $g_1(x), \cdots, g_m(x) \leq 0$ and $Ax = b$, and if $(u,v) \in \mathbb{R}^m_+ \times \mathbb{R}^p$, then $\sum_i u_i g_i(x) \leq 0$ implies $\theta(u,v) \leq f(x) + \sum_i u_i g_i(x) + v^t(Ax-b) \leq f(x)$.

As for the additional part, for a $\bar{x} \in Z$ and a $(\bar{u},\bar{v}) \in \mathbb{R}^m_+ \times \mathbb{R}^p$ the equality $\theta(\bar{u},\bar{v}) = f(\bar{x})$ first implies $\theta(\bar{u},\bar{v}) = f(\bar{x}) \geq \theta(u,v)$ for any $(u,v) \in \mathbb{R}^m_+ \times \mathbb{R}^p$, and then also, in a second round, $f(\bar{x}) = \theta(\bar{u},\bar{v}) \leq f(x)$ for any $x \in Z$.

(ii) By part (ii) of Corollary 3.5 and Remarks 3.2 and 3.3 in [OSC] there exists $(\bar{u},\bar{v})$ in $\mathbb{R}^m_+ \times \mathbb{R}^p$ such that $\theta(\bar{u},\bar{v}) = f(\bar{x})$, as was already seen in Exercise 1.4. QED

Example 1.4 Example 1.2 gave the same (unique) optimal primal solution $\bar{x} = A^t(AA^t)^{-1}b$ and the (unique) dual optimal solution was found to be $\bar{v} := -2(AA^t)^{-1}b$. Since $f(x) = |x|^2$ and $\theta(v) = -v^tAA^tv/4 - b^tv$, we find empirically

$$f(\bar{x}) = b^t(AA^t)^{-1}b = -b^t(AA^t)^{-1}b + 2b^t(AA^t)^{-1}b = \theta(\bar{v}).$$

By Theorem 1.3(ii) this identity should indeed hold, because Slater's constraint qualification is valid (vacuously – this problem (P) has no inequality constraints) and regularity holds trivially.

2 Duality: towards an approach based on perturbations

After having used the "convex" KKT theorem of [OSC] to arrive at its duality reinterpretation in Theorem 1.3, we can now move toward far-reaching generalizations. First we shall only sketch the outlines of this approach in the context of our standard optimization problem

$$(P) \ \inf_{x \in S} \{f(x) : g_1(x) \leq 0, \cdots, g_m(x) \leq 0, Ax - b = 0\}.$$

We consider an infinite collection of *perturbed* optimization problems $(P_{y,z})$, $y \in \mathbb{R}^m$, $z \in \mathbb{R}^p$, given by

$$(P_{y,z}) \quad \inf_{x \in S} \{f(x) : g_1(x) \le y_1, \cdots, g_m(x) \le y_m, Ax - b = z\},$$

where we "perturb" the right hand side of the inequality and equality constraints. It is important to observe that the problem $(P_{0,0})$ coincides with (P). Here the following convention will be followed: *the infimum over the empty set is* $+\infty$ and, likewise, the supremum over the empty set is $-\infty$. The *perturbation function* associated to the above perturbations is $\nu : \mathbb{R}^m \times \mathbb{R}^p \to [-\infty, +\infty]$ is given by $\nu(y,z) := \inf(P_{y,z})$. Hence, $\nu(0,0) = \inf(P)$ gives the minimum value of the primal problem (P). By Exercise 2.1 below, the function ν is convex under the usual convexity conditions for (P). Let us now calculate the Fenchel conjugate function of ν (recall this notion from Appendix B in [OSC]). We have

$$\begin{aligned} \nu^*(-u,-v) &:= \sup_{y,z}[-u^t y - v^t z - \nu(y,z)] \\ &= \sup_{x \in S, y} \{-u^t y - v^t(Ax-b) - f(x) : g_1(x) \le y_1, \cdots, g_m(x) \le y_m\}. \end{aligned}$$

By the usual arguments this gives

$$\nu^*(-u,-v) = \begin{cases} \sup_{x \in S} -f(x) - \sum_{i=1}^m u_i g_i(x) - v^t(Ax-b) = -\theta(u,v) & \text{if } u \ge 0, \\ +\infty & \text{otherwise} \end{cases}$$

Hence, the dual optimization problem (D) can also be written as

$$(D) \quad \sup_{(u,v) \in \mathbb{R}^m \times \mathbb{R}^p} -\nu^*(-u,-v),$$

and it is then obvious that $\sup(D) = \nu^{**}(0,0)$. Also, the strong duality identity $\inf(P) = \theta(\bar{u},\bar{v})$ for $(\bar{u},\bar{v})$ in Theorem 1.3(ii) is *equivalent* to $\nu(0,0) = \theta(\bar{u},\bar{v}) = -\nu^*(-\bar{u},-\bar{v})$, which, in turn, is equivalent to

$$(-\bar{u},-\bar{v}) \in \partial\nu(0,0)$$

by Proposition B.4(v) of [OSC]. See figures 6.1 and 6.3 in [BSS] for a geometric interpretation (with only inequality constraints). This shows that the optimal dual solution is associated with marginal behavior with respect to perturbations of the right hand side. This is especially clear if ν is differentiable at $(0,0)$, for then for small increments δy and δz of the right hand side in (P) give, by the above, in first order approximation:

$$\inf(P_{\delta y,\delta z}) = \nu(\delta y,\delta z) \doteq \inf(P) + \bar{u}^t \delta y + \bar{v}^t \delta z. \tag{1}$$

This is useful to determine possible "bottlenecks" in a concrete problem: look for the highest values of $\bar{u}_i$ and of $|\bar{v}_j|$.

Exercise 2.1 Prove that, under the convexity conditions of Theorem 1.1, the perturbation function $\nu : \mathbb{R}^m_+ \times \mathbb{R}^p \to [-\infty, +\infty]$ is convex. Note here that the definition of convexity must be adapted: ν is said to be convex if $\nu(\lambda(y_1,z_1) + (1-\lambda)(y_2,z_2)) \le \lambda\nu(y_1,z_1) \dot{+} (1-\lambda)\nu(y_2,z_2)$. Here $\dot{+}$ is just the same as $+$, but with the additional convention $(-\infty)\dot{+}(+\infty) = (+\infty)\dot{+}(-\infty) = +\infty$.

3 Duality: an approach based on perturbations

In this section we present an abstract approach to duality; however, we still keep all dimensions finite-dimensional (if infinite dimensions are chosen, applications can also include problems in mechanics and approximation theory for instance – see [ET] or [L]). In the previous section we associated to our standard optimization problem

$$(P)\ \inf_{x \in S} \{f(x) : g_1(x) \le 0, \cdots, g_m(x) \le 0, Ax - b = 0\},$$

the collection of *perturbed* optimization problems $(P_{y,z})$, $y \in \mathbb{R}^m$, $z \in \mathbb{R}^p$, given by

$$(P_{y,z})\ \inf_{x \in S} \{f(x) : g_1(x) \le y_1, \cdots, g_m(x) \le y_m, Ax - b = z\}.$$

It was shown there that the associated *perturbation function* $\nu : \mathbb{R}^m \times \mathbb{R}^p \to [-\infty, +\infty]$, defined by $\nu(y, z) := \inf(P_{y,z})$, had the following Fenchel conjugate:

$$\nu^*(-u, -v) = \begin{cases} -\theta(u, v) & \text{if } u \ge 0, \\ +\infty & \text{otherwise} \end{cases}$$

This caused the usual Lagrangian dual optimization problem (D) to be rewritable as

$$(D)\quad \sup_{(u,v) \in \mathbb{R}^m \times \mathbb{R}^p} -\nu^*(-u, -v).$$

There we also noticed the following connections:

$$\inf(P) = \nu(0, 0) \text{ and } \sup(D) = \nu^{**}(0, 0).$$

It followed that *weak duality*, as defined in section 1, can be reproven by appealing to Proposition B.4(iv) in [OSC]. Moreover, the strong duality identity $f(\bar{x}) = \theta(\bar{u}, \bar{v})$ for $(\bar{u}, \bar{v})$ from section 1 was seen to be *equivalent* to $\nu(0, 0) = -\nu^*(-\bar{u}, -\bar{v})$, which, in turn, is equivalent to

$$(-\bar{u}, -\bar{v}) \in \partial\nu(0, 0)$$

by Proposition B.4(v) of [OSC].

We shall now generalize these findings as follows. For some $k \in \mathbb{N}$ let $\phi : \mathbb{R}^n \times \mathbb{R}^k \to (-\infty, +\infty]$ be a given convex function. Denote $\phi_0(x) := \phi(x, 0)$ and consider the *primal* optimization problem

$$(P)\quad \inf_{x \in \mathbb{R}^n} \phi_0(x),$$

associated to which we consider the family $\{(P_p)\}$ of *perturbed* optimization problems

$$(P_p)\quad \inf_{x \in \mathbb{R}^n} \phi(x, p).$$

Define $h(p) := \inf(P_p)$; then $h : \mathbb{R}^k \to [-\infty, +\infty]$, the so-called *perturbation function*, is a *convex* function (see Exercise 2.1). We suppose that $\inf(P) \notin \{-\infty, +\infty\}$, which is equivalent to $h(0) \in \mathbb{R}$.

Now let $h^* : \mathbb{R}^k \to (-\infty, +\infty]$ be the Fenchel conjugate of h. This is defined as in Definition B.1 of [OSC] (here $h(0) \in \mathbb{R}$ implies that h^* cannot take the value $-\infty$). Then the *dual* optimization problem is defined as

$$(D)\quad \sup_{q \in \mathbb{R}^k} -h^*(-q).$$

Theorem 3.1 *(general duality-stability theorem))* (i) *For every* $x \in \mathbb{R}^n$ *and every* $q \in \mathbb{R}^k$

$$-h^*(-q) \leq \phi_0(x) \ \textbf{(weak duality)}.$$

In particular,

$$\inf(P) = h(0) \geq h^{**}(0) = \sup(D).$$

(ii) *The following are equivalent:*
(iia) *the perturbation function* h *is lower semicontinuous at* 0 **(weak stability)**,
(iib) $\inf(P) = \sup(D)$ **(absence duality gap)**.

(iii) *If the perturbation function* h *is continuous at* 0 **((strong) stability)**, *then*

$$\inf(P) = \max(D) \ \textbf{(absence duality gap + existence optimal dual solution)}$$

and the set of all optimal dual solutions is $-\partial h(0) \neq \emptyset$.

Exercise 3.1 Prove the above Theorem 3.1. *Hint:* First prove that, by $h(0) \in \mathbb{R}$ and convexity of h, weak stability implies $h > -\infty$. Imitate parts of the proof of the Lagrangian duality theorem and use properties of Fenchel conjugates in Appendix B of [OSC]. Use also Lemma 2.16 of [OSC].

Exercise 3.2 Give a new proof of the Lagrangian duality theorem in section 1 (with inequality constraints only), by applying Theorem 3.1 to a suitably formulated convex function $\phi : \mathbb{R}^n \times \mathbb{R}^m \to (-\infty, +\infty]$. *Hint:* Argue that Slater's constraint qualification allows application of Lemma 2.16 of [OSC].

Exercise 3.3 Give an extension of the Lagrangian duality theorem in section 1 (with inequality constraints only), by considering the problem $\inf_{x \in \mathbb{R}^n} \{f(x) : g(x) \in K\}$. Here $g := (g_1, \ldots, g_m)$ and K is a closed convex cone in $\mathbb{R}^m$ (in the special case $K = \mathbb{R}^m_-$ we return to the problem mentioned in the first line of this exercise). *Hint:* Imitate parts of the proof of the Lagrangian duality theorem.

Theorem 3.2 *(Fenchel's duality theorem) Consider the following convex minimization problem*

$$(P_F) \quad \inf_{x \in \mathbb{R}^n} [f(x) + g(Ax)],$$

where $f : \mathbb{R}^n \to (-\infty, +\infty]$ *and* $g : \mathbb{R}^m \to (-\infty, +\infty]$ *are convex functions and* A *is an* $m \times n$*-matrix. Suppose that* $\inf(P_F) \in \mathbb{R}$.

Define the associated Fenchel dual problem *as follows:*

$$(D_F) \quad \sup_{q \in \mathbb{R}^m} [-f^*(A^t q) - g^*(-q)].$$

(i) *For every* $x \in \mathbb{R}^n$ *and every* $q \in \mathbb{R}^m$

$$-f^*(A^t q) - g^*(-q) \leq f(x) + g(Ax).$$

(ii) *If* $0 \in \text{int } (\text{dom } g - A(\text{dom } f))$, *then*

$$\inf(P_F) = \max(D_F).$$

Moreover, then $\bar{x} \in \mathbb{R}^n$ is optimal for (P_F) and $\bar{q} \in \mathbb{R}^k$ is optimal for (D_F) if and only if

$$A^t\bar{q} \in \partial f(\bar{x}) \text{ and } -\bar{q} \in \partial g(A\bar{x}). \tag{2}$$

Proof. We apply Theorem 3.1 to the following perturbed optimization problems

$$(P_p) \quad \inf_{x\in\mathbb{R}^n} [f(x) + g(Ax+p)].$$

That means we use $\phi(x,p) := f(x) + g(Ax+p)$. It gives

$$h^*(-q) = \sup_p\{-q^tp - \inf_x[f(x) + g(Ax+p)]\} = \sup_{p,x}[-q^tp - f(x) - g(Ax+p)].$$

Hence,

$$h^*(-q) = \sup_x\{-f(x)+\sup_p[-q^tp-g(Ax+p)]\} = \sup_x\{-f(x)+q^tAx+\sup_{p'}[-q^tp'-g(p')]\},$$

where we substitute $p' := Ax + p$ to work out the inner supremum. It thus follows that

$$h^*(-q) = g^*(-q) + \sup_x[x^tA^tq - f(x)] = g^*(-q) + f^*(A^tq).$$

Hence the Fenchel dual problem (D_F) is a special case of the dual problem (D) of Theorem 3.1, obtained by choosing ϕ as above. So we may apply that theorem, of which we only need parts (i) and (iii).

The thing that remains to be settled is the strong stability condition (i.e., the continuity of h at 0). Now note that the condition $0 \in \text{int}\,(\text{dom}\, g - A(\text{dom}\, f))$ implies that there exists $\epsilon > 0$ such that

$$\forall_{p,|p|<\epsilon}\exists_{x_p,y_p} p = y_p - Ax_p,\ f(x_p) < +\infty,\ g(y_p) < +\infty.$$

Then $h(p) := \inf(P_p) \le f(x_p) + g(Ax_p + p) = f(x_p) + g(y_p) < +\infty$ for all p with $|p| < \epsilon$. So $0 \in \text{int dom}\, h$, and this implies continuity of h in 0 by Lemma 2.16 in [OSC]. The final part is Exercise 3.5. QED

Remark 3.3 *Theorem 3.2 applies to the following particular cases:*
(a) If $g := \chi_{\{b\}}$, it applies to $\inf_{x,Ax=b} f(x)$.
(b) If $g := \chi_{b+K}$, it applies to $\inf_{x,Ax\in b+K} f(x)$. Here $K \subset \mathbb{R}^m$ is a convex cone and $Ax \in b + K$ called a conical constraint. *Special cones: $K = \{0\}$ (then we are back to (a)) or $K = \mathbb{R}^m_-$.*
(c) If $f(x) := \mu^{-1}c^tx - \sum_{i=1}^n \log(x_i) + \chi_{\mathbb{R}^n_{++}}(x)$ and $g := \chi_{\{b\}}$, then it applies to minimizing the so-called logarithmic barrier function. *Here $c \in \mathbb{R}^n$ and $\mu > 0$ is a scaling (penalty) parameter.*
(d) If $f(x) = \sum_{i=1}^n f_i(x_i)$ (additive separability) and $g = \chi_{\{b+K\}}$, then it applies to $\inf_{x,Ax\in b+K}\sum_i f_i(x_i)$. Here the f_i are convex functions on $\mathbb{R}$. Consider in particular the case $f_1 = \cdots = f_n$ and, even more in particular, the case with $f_i(x_i) := (x_i)^2$ for all i.

Exercise 3.4 Give a complete workout of Theorem 3.2 in each of the special situations (a), (b), (c) and (d), described in Remark 3.3.

Exercise 3.5 Prove the equivalence involving (2) in Theorem 3.2.

4 Bidual optimization problems

Observe that (D_F) is equivalent to

$$(D'_F) \quad \inf_{x \in \mathbb{R}^m} [g^*(x) + f^*((-A)^t x)]$$

(set $x := -q$) which is of the same form as (P_F) above, if we substitute $n := m$, $f := g^*$, $g := f^*$ and $A := -A^t$. So let us now take the dual of (D'_F). Then we obtain the *bidual* optimization problem:

$$(D''_F) \quad \sup_{q \in \mathbb{R}^n} [-g^{**}(-A^{tt}q) - f^{**}(-q)],$$

Of course, here $-A^{tt} = -A$, and if f and g are at least convex and lower semicontinuous functions, then $f^{**} = f$ and $g^{**} = g$ by the Fenchel-Moreau theorem. So then (D''_F) can be rewritten as

$$(D''_F) \quad \sup_{x \in \mathbb{R}^n} [-f(x) - g(Ax)],$$

which is equivalent to the original primal optimization problem (P_F). This finding can be generalized, as is done in the following exercise:

Exercise 4.1 Suppose that $\phi : \mathbb{R}^n \times \mathbb{R}^k \to (-\infty, +\infty]$ is convex and lower semicontinuous. Define $\psi := \phi^*$ and observe that $-h^*(-q) = -\psi(0, -q)$ holds for the dual objective function, which is analogous to $\phi_0(x) := \phi(x, 0)$, the objective function of the primal problem (P). Now develop a duality-stability scheme that uses the perturbed optimization problems (D_y) : $\sup_q -\psi(y, -q)$ to produce the general bidual optimization problem. Prove that it is equivalent to the original primal problem (P). Show also that the above bidual of (P_F) fits into that scheme.