Perturbational duality (continued) and Applications
*

Erik J. Balder

1 Polyhedral convexity and duality

A subset S of $\mathbb{R}^n$ is said to be **polyhedral** if it is the intersection of a finite number of closed halfspaces, i.e., if there exist $J \in \mathbb{N}$ and collections $\{y_1, \ldots, y_J\} \subset \mathbb{R}^n$, $\{\alpha_1, \ldots, \alpha_J\} \subset \mathbb{R}$ such that $S = \cap_{j=1}^J \{x \in \mathbb{R}^n : y_j^t x \leq \alpha_j\}$. A function $f : \mathbb{R}^n \to [-\infty, +\infty]$ is **polyhedral** if its epigraph epi $f \subset \mathbb{R}^{n+1}$ is a polyhedral set. Clearly, any polyhedral set is automatically convex and closed. Consequently, any polyhedral function is convex and lower semicontinuous (exercise). Polyhedral functions have a special form, which can also be used to obtain *refinements* of the previous duality results (for instance, refinements that include linear programming duality). We shall not go into the details of this, but the following example illustrates the main point, which is to be confirmed by Proposition 1.2 below: the conditions under which polyhedral functions are subdifferentiable are less stringent than those of arbitrary convex functions.

Example 1.1 a. Consider again the convex function $f : \mathbb{R} \to (-\infty, +\infty]$, defined by $f(x) := 1 - \sqrt{1 - x^2}$ if $|x| \leq 1$ and $f(x) := +\infty$ otherwise. Consider the following primal optimization problem (P): $\inf\{f(x) : 1 - x \leq 0\}$, which has the trivial optimal solution $\bar{x} = 1$. If we perturb the right hand side of the inequality in the usual way, we obtain (P_p): $\inf\{f(x) : 1 - x \leq p\}$, For $h(p) := \inf(P_p)$ this gives

$$h(p) = \begin{cases} +\infty & \text{if } p < 0 \\ f(1-p) = 1 - \sqrt{2p - p^2} & \text{if } 0 \leq p \leq 1 \\ f(0) = 0 & \text{if } p > 1 \end{cases}$$

Hence, the dual objective function is seen to be as follows. For $q < 0$ we get $-h^*(-q) = -\infty$. For $q \geq 0$ we can determine $h^*(q) = \max(-q, -q + \sup_{0 \leq t \leq 1}[tq - f(t)])$ by first showing that $\sup_{0 \leq t \leq 1}[tq - f(t)] = \sqrt{1 + q^2} - 1$ and then concluding that $-h^*(-q) = 1 - (q + \sqrt{1 + q^2})^{-1}$. So the dual problem (D) is to maximize $1 - (q + \sqrt{1 + q^2})^{-1}$ over all $q \geq 0$. This gives $\sup(D) = 1 = f(\bar{x}) = \min(P)$, but

*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.

it is clear that there exists no optimal dual solution. These two findings could also have been deduced directly from the calculation of h, which shows that (1) $h(p)$ is lower semicontinuous and finite at $p = 0$, but (2) $\partial h(0) = \emptyset$. In this problem the perturbation function h is clearly not polyhedral.

b. Next, consider the same primal problem, but now with $f(x) = |x|$ for $|x| \leq 1$ and $f(x) := +\infty$ otherwise. Then (P): $\inf\{f(x) : 1 - x \leq 0\}$ again has the trivial optimal solution $\bar{x} = 1$. For (P_p): $\inf\{f(x) : 1 - x \leq p\}$ and $h(p) := \inf(P_p)$ we now easily find

$$h(p) = \begin{cases} +\infty & \text{if } p < 0 \\ f(1-p) = 1 - p & \text{if } 0 \leq p \leq 1 \\ f(0) = 0 & \text{if } p > 1 \end{cases}$$

The dual objective function is as follows: $-h^*(-q) = -\infty$ if $q < 0$, $-h^*(-q) = q$ if $0 \leq q < 1$ and $-h^*(-q) = 1$ if $q \geq 1$. Clearly, the set of all optimal dual solutions is $[1, +\infty)$ (which is confirmed by $-\partial h(0) = [1, +\infty)$) and $\max(D) = 1 = \min(P)$. In this problem the perturbation function is polyhedral.

Proposition 1.2 *Let $f : \mathbb{R}^n \to [-\infty, +\infty]$, $f \not\equiv +\infty$, be polyhedral.*

(i) There exists a finite collection (possibly empty) of affine functions $a_i : \mathbb{R}^n \to \mathbb{R}$, $i \in I$ and a nonempty polyhedral set $P \subset \mathbb{R}^n$ such that

$$f(x) = \chi_P(x) \dot{+} \max_{i \in I} a_i(x) \text{ for every } x \in \mathbb{R}^n$$

and I may be empty, in which case the maximum is set to $-\infty$.[1]

(ii) Moreover, if $f \not\equiv -\infty$ then

$$\partial f(x_0) \neq \emptyset \text{ for every } x_0 \in \operatorname{dom} f.$$

Proof. By definition, epi f is the intersection of some number $J \geq 1$ of halfspaces $\{(x, r) \in \mathbb{R}^n \times \mathbb{R} : z_j^t x + \rho_j r \leq \alpha_j\}$. We now distinguish the "vertical halfspaces" from the non-vertical ones: let $I \subset \{1, \ldots, J\}$ be the set of all indices j with $\rho_j \neq 0$. If $j \in I$, then $\emptyset \neq \text{epi } f \subset \{(x, r) \in \mathbb{R}^n \times \mathbb{R} : z_j^t x + \rho_j r \leq \alpha_j\}$ implies $\rho_j < 0$ and we set $a_j(x) := (\alpha_j - z_j^t x)/\rho_j$. Also, we define $P := \cap_{j \notin I}\{x \in \mathbb{R}^n : z_j^t x \leq \alpha_j\} \subset \mathbb{R}^n$; then P is polyhedral. Now epi$f = \{(x, r) : x \in P\} \cap \cap_{j \in I}$epi a_j. If I is nonempty, then this states epi$f = \text{epi}(\chi_P + \max_{j \in I} a_j)$ and the desired identity follows; observe also that in this case $f > -\infty$. If I is empty, we find epi f $= P \times \mathbb{R}$ and the desired identity also follows. Finally, by the preceding lines the extra condition $f > -\infty$ implies that I is nonempty. In this case both $x \mapsto \max_{i \in I} a_i(x)$ and χ_P take values in $(-\infty, +\infty]$. The final step of the proof of part (ii) now follows by the Moreau-Rockafellar theorem in [OSC] (exercise). QED

Using certain representation results for polyhedral sets (which state that any such set can be decomposed as the sum $C + K$ of a convex hull C of finitely many points and a finitely generated cone K) it can be proved that if the function ϕ in

[1] Observe: in this situation the identity comes down to the following: $f(x) = -\infty$ if $x \in P$ and $f(x) = +\infty$ if $x \in \mathbb{R}^n \backslash P$.

section 3 of part I is polyhedral, then so is the perturbation function h [R] . Therefore, part (ii) of the above proposition implies that in this situation strong duality holds under less stringent conditions than those stated in Theorem 1.3 of Part I. A special, but important consequence of this development is the following version of Fenchel's duality theorem, which has a much weaker sufficient condition for strong stability than Theorem 3.2 in part I:

Theorem 1.3 *(Fenchel's duality theorem for polyhedral functions) 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* polyhedral *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* $0 \in \mathrm{dom}\ g - A(\mathrm{dom}\ f))$, *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*

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

2 Semidefinite programming duality

Until now, we have worked with optimization problems defined on the finite-dimensional vector space $\mathbb{R}^n$, whose inner product was given by $< x, y >:= x^t y$. In this section we shall consider another finite-dimensional vector space, namely the set $\mathbb{S}^n$ of all symmetric $n \times n$-matrices. This vector space has $\mathbb{R}$ as the field of scalars. It has the following inner product

$$\langle X, Y\rangle := \mathrm{tr}(XY) =:= \sum_i (XY)_{ii} = \sum_{i,j} X_{i,j}Y_{i,j}$$

and the corresponding inner product norm is $\|X\| := (\sum_{i,j} X_{i,j}^2)^{1/2}$. For fixed $i < j$ let $E^{i,j} \in \mathbb{S}^n$ be the $n \times n$-matrix consisting of zeros except for the (i,j)-th and (j,i)-th entries, which are set equal to $\frac{1}{2}\sqrt{2}$. Also, let $E^{i,i} \in \mathbb{S}^n$ be the $n \times n$-matrix consisting of zeros except for the (i,i)-th entry, which is set equal to 1. Then every

X in $\mathbb{S}^n$ can be decomposed as a linear combination of the $\frac{1}{2}n(n+1)$ matrices $E^{i,j}$, which therefore form a basis of $\mathbb{S}^n$:

$$X = \sum_{i,j,j \geq i} X_{i,j} E^{i,j} = \sum_{i,j,j \geq i} \langle X, E^{i,j} \rangle E^{i,j}$$

and this basis can easily be checked to be orthonormal (i.e., its elements are mutually orthogonal with respect to the above inner product and they all have unit length with respect to the corresponding inner product norm). We shall leave it to the reader to inspect that all the foregoing material in this course extends from $\mathbb{R}^n$ to an arbitrary finite-dimensional vector space, hence to $\mathbb{S}^n$ in particular. Two special convex cones in $\mathbb{S}^n$ are (1) $\mathbb{S}^n_+$, the set of all $n \times n$-matrices that are *positive semidefinite* (p.s.d.), and (2) $\mathbb{S}^n_{++}$, the set of all *positive definite* (p.d.) $n \times n$-matrices. Recall that $A \in \mathbb{S}^n$ is p.s.d. if $x^t A x \geq 0$ for all $x \in \mathbb{R}^n$ and A is p.d. if $x^t A x > 0$ for all $x \in \mathbb{R}^n \backslash \{0\}$.

Exercise 2.1 Prove the following: $\mathbb{S}^n_+$ and $\mathbb{S}^n_{++}$ are convex cones and, moreover, $\mathbb{S}^n_+$ is closed and $\mathbb{S}^n_{++}$ is the interior of $\mathbb{S}^n_+$.

Observe that $\mathbb{S}^n_+$ forms a convex subset of $\mathbb{S}^n$. However, it is a rather complicated convex subset; for instance, for $n = 2$ a matrix $X \in \mathbb{S}^2$ belongs to $\mathbb{S}^2_+$ if and only if

$$X_{11}, X_{22} \geq 0 \text{ and } X_{11} X_{22} \geq X_{12}^2. \tag{1}$$

Let $A_i \in \mathbb{S}^n$ and $b_i \in \mathbb{R}$ for $i = 1, \ldots, m$. The *primal semidefinite program* is

$$(P_{SDP}) \quad \inf_{X \in \mathbb{S}^n_+} \{\operatorname{tr}(CX) : \operatorname{tr}(A_i X) = b_i,\ i = 1,\ \ldots, m\}.$$

Of course, this can also be written as

$$(P_{SDP}) \quad \inf_{X \in \mathbb{S}^n_+} \{\mathcal{A}(X) = b\},$$

where $b := (b_1, \ldots, b_m)^t$ and where the linear mapping $\mathcal{A} : \mathbb{S}^n \to \mathbb{R}^m$ is given by

$$\mathcal{A}(X) := (\langle A_i, X \rangle)_{i=1}^m.$$

The form of (P_{SDP}) is similar to that of the standard linear program

$$\inf_{x \in \mathbb{R}^n_+} \{< c, x >:< a_i, x >= b_i,\ i = 1,\ \ldots, m\},$$

where the corresponding linear mapping is the matrix multiplication $x \mapsto Ax$, with A being the $n \times m$-matrix whose i-th row is the horizontal vector a_i^t. We shall now show that the natural dual problem of (P_{SDP}) is given by

$$(D_{SDP}) \quad \sup_{q \in \mathbb{R}^m} \{b^t q : C - \sum_{i=1}^m q_i A_i \in \mathbb{S}^n_+\}.$$

Conversely, by biduality (see the previous section), (P_{SDP}) can also be regarded as the dual of (D_{SDP}). Semidefinite programs form a broad class of convex optimization problems. They include linear programming, quadratic programming, but also fractional programming as the next example shows:

Example 2.1 Consider for an $m \times n$-matrix A and vectors $c, d \in \mathbb{R}^n$, $b \in \mathbb{R}^m$:

$$(P) \quad \inf_{x \in \mathbb{R}^n} \{ \frac{(c^t x)^2}{d^t x} : Ax + b \geq 0 \}.$$

We assume that $d^t x > 0$ for every x with $Ax + b \geq 0$. Then (P) can be rewritten as

$$\inf_{t \in \mathbb{R}, x \in \mathbb{R}^n} \{ t : Ax + b \geq 0, (c^t x)^2 \leq t d^t x \}$$

Note that the last constraint is of the same form as (1). That enables us to rewrite (P) in the same form as (D_{SDP}) (exercise).

Exercise 2.2 Execute the last line in Example 2.1.

We have the following result about duality for SDP:

Theorem 2.2 *(i) Suppose that (P_{SDP}) has a feasible solution that is p.d. Then* $\inf(P_{SDP}) = \max(D_{SDP})$, *provided that* $\sup(D_{SDP}) \in \mathbb{R}$.
(ii) Suppose that there exists $\tilde{q} \in \mathbb{R}^m$ such that $C - \tilde{q}_i A_i$ is p.d. Then $\min(P_{SDP}) = \sup(D_{SDP})$, *provided that* $\inf(P_{SDP}) \in \mathbb{R}$.

Proof.Step 1. In this step we prove that Fenchel's duality theorem applies. Let $K := \mathbb{S}^n_+$ and observe that the negative polar K^* equals the cone $-\mathbb{S}^n_+$ of all negative semidefinite symmetric matrices. Define $f : \mathbb{S}^n \to (-\infty, +\infty]$ and $g : \mathbb{R}^m \to \mathbb{R}$ by $f(X) := \langle C, X \rangle + \chi_K(X)$ and $g(y) := \chi_{\{b\}}(y)$. Then (P_{SPD}) can be rewritten as the following version of Fenchel's primal (P_F):

$$\inf_{X \in \mathbb{S}^n} [f(X) + g(\mathcal{A}(X))].$$

Observe that

$$f^*(Y) := \sup_{X \in \mathbb{S}^n} [\langle X, Y \rangle - f(X)] = \left\{ \begin{array}{ll} 0 & \text{if } Y - C \in K^* \\ +\infty & \text{otherwise} \end{array} \right\} = \chi_{C+K^*}(Y)$$

and that $g^*(-q) = \sup_{y \in \mathbb{R}^m} [(-q)^t y - \chi_{\{b\}}(y)] = -q^t b$. According to Fenchel the corresponding dual version of (P_{SDP}) is

$$\sup_{q \in \mathbb{R}^m} [-f^*(\mathcal{A}^*(q)) - g^*(-q)],$$

where $\mathcal{A}^* : \mathbb{R}^m \to \mathbb{S}^n$ is the so-called *adjoint* of the operator $\mathcal{A}$, which is defined via the following relationship:[2]

$$\langle X, \mathcal{A}^*(q) \rangle =< \mathcal{A}(X), q > \text{ for every } X \in \mathbb{S}^n,\ q \in \mathbb{R}^m.$$

[2]Notice that for a *matrix* operator $x \mapsto Ax$ from $\mathbb{R}^n$ to $\mathbb{R}^m$ the same kind of relationship leads to the adjoint being $q \mapsto A^t q$, because of $< Ax, q >= x^t A^t q$.

Combining the above, we conclude that the previous dual version of (P_{SDP}) can be rewritten as

$$\sup_{q\in\mathbb{R}^m} [-\chi_{C+K^*}(\mathcal{A}^*(q)) - (-q^t b)$$

and it is easy to see that this optimization problem coincides with the above problem (D_{SDP}).

Step 2. Based on step 1, we may apply Fenchel's duality theorem. To prove (i), we distinguish between the following cases:

Case 1: $\inf(P_{SDP}) = -\infty$. In this case weak duality implies $\sup(D_{SDP}) = -\infty$, which is impossible.

Case 2: $\inf(P_{SDP}) = +\infty$. This is impossible by the fact that (P_{SDP}) has a feasible solution that is p.d.

Case 3: $\inf(P_{SDP}) \in \mathbb{R}$. Because of the above substitutions, the sufficient condition for stability in Fenchel's duality theorem is as follows: $0 \in \text{int } (\{b\} - \mathcal{A}(K))$, i.e., $b \in \text{int } \mathcal{A}(K)$. This condition is satisfied, because of Exercise 2.3 below. In Exercise 2.4 the reader is invited to prove (ii). QED

Exercise 2.3 Show: if (P_{SDP}) has an feasible solution that is p.d., then $b \in \text{int } \mathcal{A}(K)$ holds in case 3 of part (i).

Exercise 2.4 Prove part (ii) of Theorem 2.2. *Hint:* Use the bidual approach of section 4 in part I. As an alternative, observe that (D_{SDP}) is *equivalent* to a problem of the following form: $\inf_{q\in\mathbb{R}^m}\{b^t q : c - \mathcal{A}^* q \in K\}$, where $\mathcal{A}^*, b, c$ and K are as in the proof of part (i). Using this as the primal problem, apply Fenchel's duality theorem (note here that $\mathcal{A}^{**} = \mathcal{A}$).

Exercise 2.5 Consider (P_{SDP}) for $n = 2$, $m = 1$, with $b = 0$, $C := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ and $A_1 := \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$.

a. Show: $\min(P_{SDP}) = 0$, but (D_{SDP}) has no feasible solutions. *Hint:* $M \in \mathbb{S}^2$ is s.p.d. if and only if $M_{11}, M_{22} \geq 0$ and $M_{11}M_{22} - M_{12}^2 \geq 0$.

b. Show explicitly that in this example the conditions of Theorem 2.2 do not hold.

3 Applications of duality: von Neumann's minimax theorem

As an important application of duality, we now present von Neumann's minimax theorem for zero-sum noncooperative games, which establishes equilibrium existence for mixed strategies. We describe a *zero-sum game* for two players I and II. Each player can choose from finitely many actions; more precisely, player I can choose between m actions $1, \ldots, m$ and player II between n actions $1, \ldots, n$. Let P be an $m \times n$-matrix; this is the *payoff matrix*: when I chooses action i and II chooses action j, then I has to pay II the amount of p_{ij} dollars (here p_{ij} stands for the element in the i-the row and j-th column of P – a negative number p_{ij} means, of course, that II is effectively paying a positive amount to I)).

The idea of the game is that player I wishes to choose i in such a way that the the amount p_{ij} (which she has to pay to II) is kept small, whereas player II, who is on the receiving end, wishes to keep p_{ij} large by choosing j prudently. A *pure equilibrium* for the game consists of a pair $(\bar{i}, \bar{j})$ such that

$$p_{\bar{i}\bar{j}} \leq p_{i\bar{j}} \text{ for } i = 1, \ldots, m \text{ and } p_{\bar{i}\bar{j}} \geq p_{\bar{i}j} \text{ for } j = 1, \ldots, n. \tag{2}$$

To understand this notion, observe that any deviation from $\bar{i}$ results in a loss for player I in that she has to pay more (or at least not less) to II; similarly, any deviation from $\bar{j}$ results in a loss for player II in that he will receive less from I. Unfortunately, such a pure equilibrium rarely exists, because (2) amounts to the matrix P having a **saddle point** $p_{\bar{i},\bar{j}}$. However, most matrices do not possess such saddle points. For instance, the 2×2 matrix P with $p_{11} = p_{22} = 1$ and $p_{12} = p_{21} = -1$ (this models the well-known game of "matching pennies") does not have a saddle point.

Von Neumann's resolution of this dilemma (which was already known to gamblers as early as 1713) is to allow both players to use *mixed strategies*. Formally, this means that players I and II each can choose a probability vector from the unit simplices $S_I := \{x \in \mathbb{R}^m_+ : \sum_{i=1}^n x_i = 1\}$ and $S_{II} := \{u \in \mathbb{R}^n_+ : \sum_{j=1}^m u_j = 1\}$ respectively. By basic expectation and probabilistic independence considerations the *expected payoff* $E(x, u)$ for player I is then defined as

$$E(x, u) := \sum_{i=1}^{m} \sum_{j=1}^{n} x_i u_j p_{ij} = x^t P u,$$

when I uses $x \in S_I$ and II uses $u \in S_{II}$. A *mixed equilibrium* for the game consists of a pair $(\bar{x}, \bar{u})$ such that

$$E(\bar{x}, \bar{u}) \leq E(x, \bar{u}) \text{ for all } x \in S_I \text{ and } E(\bar{x}, \bar{u}) \geq E(\bar{x}, u) \text{ for all } u \in S_{II}.$$

The interpretation of this notion is quite similar to the one given above for the pure equilibrium: players cannot become better off by deviating. Of course, the use of *expectations* in this equilibrium notion requires further justification, for which we refer to the literature on game theory. Observe that if a pure equilibrium should exist, this corresponds to having a special mixed equilibrium, namely one where the pair of probability vectors consists of two unit vectors. The success of the mixed extension of the pure equilibrium concept lies in the fact that for both the simple games considered here and their generalizations, a mixed equilibrium always exists:

Theorem 3.1 (von Neumann) *For the above game there exists at least one mixed equilibrium pair.*

PROOF. Observe first that $\max_j (x^t P)_j = \max_j x^t P^j = \sup_{u \in S_{II}} E(x, u)$ for every $x \in S_I$, because the maximum of the linear function $u \mapsto x^t P u$ over the unit simplex S_{II} is attained in the extreme point (= unit vector) e_j of S_{II} for which $x^t P e_j = x^t P^j$

has the largest value. Similarly, we get $\min_i (Pu)_i = \min_i P_i^t u = \inf_{x \in S_I} E(x,u)$ for every $u \in S_{II}$. Consider now the optimization problem

$$(P_I) \ \inf_{x \in S_I} \max_j x^t P^j.$$

By the Weierstrass theorem, there exists an optimal solution $\bar{x} \in S_I$ of (P_I). By the obvious identity $\max_j x^t P^j = \inf\{r \in \mathbb{R} : r \geq x^t P^j \text{ for } j = 1, \ldots, n\}$ we have $\inf\,(P_I) = \inf\,(P_I')$, where

$$(P_I') \ \inf_{x \geq 0, r \in \mathbb{R}} \{r : x^t P^j - r \leq 0, j = 1, \ldots, n, 1 - \sum_{i=1}^m x_i = 0\},$$

is an equivalent optimization problem. This has as its optimal solution the pair $(\bar{x}, \bar{r})$, with $\bar{r} := \max_j \bar{x}^t P^j$. The problem (P_I') is a special case of the convex programming problem in the Lagrangian duality theorem. We see this by the substitutions $S := \mathbb{R}_+^m \times \mathbb{R}$, $f(x,r) := r$, $g_j(x,r) := x^t P^j - r$, $A := (-1, \cdots, -1, 0)$ and $b := -1$. Observe that the regularity conditions of that theorem hold (including Slater's constraint qualification – notice that $(1/m, \cdots, 1/m)$ lies in $\operatorname{int} \mathbb{R}_+^m \cap L$ and that $g_j(0,-1) = -1 < 0$). Hence, there exist $\bar{u} \in \mathbb{R}_+^n$, $\bar{v} \in \mathbb{R}$ such that

$$\theta_1(\bar{u}, \bar{v}) = \bar{r} = \max_j (\bar{x}^t P)_j.$$

Here

$$\theta_1(u,v) := \inf_{x \geq 0, r \in \mathbb{R}} r + \sum_j u_j (x^t P^j - r) + v(1 - \sum_i x_i)$$

is the dual objective function. By rewriting this as

$$\theta_1(u,v) = \inf_{x \geq 0, r \in \mathbb{R}} v + r(1 - \sum_j u_j) + \sum_i x_i ((Pu)_i - v)$$

we obtain

$$\theta_1(u,v) = \begin{cases} v & \text{if } \sum_j u_j = 1 \text{ and } \min_i (Pu)_i \geq v, \\ -\infty & \text{otherwise.} \end{cases}$$

Since $\theta_1(\bar{u}, \bar{v}) > -\infty$, we have $\bar{u} \in S_{II}$ and $\min_i (P\bar{u})_i \geq \bar{v} = \theta_1(\bar{u}, \bar{v}) = \max_j (\bar{x}^t P)_j$. By what was said at the start of this proof, this amounts to $\inf_{x \in S_I} E(x, \bar{u}) \geq \sup_{u \in S_{II}} E(\bar{x}, u)$, and by $E(\bar{x}, \bar{u}) \geq \inf_{x \in S_I} E(x, \bar{u}) \geq \sup_{u \in S_{II}} E(\bar{x}, u) \geq E(\bar{x}, \bar{u})$ this shows that $(\bar{x}, \bar{u})$ is a mixed equilibrium pair. QED

Corollary 3.2 *If $(\bar{x}, \bar{u})$ is a mixed equilibrium pair, then*

$$\bar{x}_i((P\bar{u})_i - \bar{v}) = 0 \textit{ for } i = 1, \ldots, m \textit{ (equalizing property for player I)},$$

$$\bar{u}_j((P^t\bar{x})_j - \bar{v}) = 0 \textit{ for } j = 1, \ldots, n \textit{ (equalizing property for player II)},$$

where, as in the proof of Theorem 3.1,

$$\bar{v} := \inf_{x \in S_I} \sup_{u \in S_{II}} E(x,u) = \sup_{u \in S_{II}} \inf_{x \in S_I} E(x,u).$$

Proof. Observe that in the proof of Theorem 3.1 the following complementary slackness relation must hold, as a consequence of the duality theorem: $\bar{u}_j((\bar{x}^t P)_j - \bar{r}) = 0$ for every j. From that proof it is also clear that $\bar{r} = \bar{v}$. The other equalizing property follows by rewriting: $0 = \sum_j \bar{u}_j((\bar{x}^t P)_j - v) = \bar{x}^t P\bar{u} - v = \sum_i \bar{x}_i((P\bar{u})_i - v)$. The latter sum has only nonnegative terms. QED

The following examples demonstrate the importance of the equalizing property in determining equilibrium solutions for games; this is quite comparable to the importance of complementary slackness when solving ordinary NLP problems.

Let the *support* supp $\bar{u}$ of $\bar{u}$ be defined as the set of all j, $1 \le j \le n$, such that $\bar{u}_j > 0$; this is the set of all player II's actions j which have positive probability under the mixed strategy $\bar{u}$. Observe that the equalizing property for player I ensures that $(P^t\bar{x})_j = v$ for every $j \in$ supp $\bar{u}$. Similarly, the equalizing property for player II gives $(P\bar{u})_i = v$ for every $i \in$ supp $\bar{x}$.

Example 3.3 Consider the game with payoff matrix

$$P = \begin{pmatrix} 2 & 3 & 1 & 5 \\ 4 & 1 & 6 & 0 \end{pmatrix}$$

The four functions $(p_{2j} - p_{1j})x_2 + p_{1j}$, as well as their pointwise maximum, can be plotted easily. Thus, the minimizer of the pointwise maximum function is seen to be $\bar{x}_2 = 2/5$; this gives $\bar{x} = (3/5, 2/5)^t$ for player I's equilibrium strategy. Also, $v = 3$. Observe that $\bar{x}^t P = (14/5, 11/5, 3, 3)$, so $\bar{u}_1 = \bar{u}_2 = 0$ by the equalizing property for player I. Also, since supp $\bar{x} = \{1, 2\}$, player II's equalizing property leads to the equations $\bar{u}_3 + 5\bar{u}_4 = 3$ and $6\bar{u}_3 = 3$, which result in $\bar{u} = (0, 0, 1/2, 1/2)^t$.

The following surveillance example, with the same structure as Example 3.3 above, is from [KF]:

Example 3.4 [surveillance of a store] A store has two rooms A and B and a control room T with t.v. monitors. The store is being guarded by two surveillants. There is one (prospective) thief, who can strike in either A or B. When the thief strikes, then his/her chance of being caught in the act by a specific surveillant in T is 0.3 in room A and 0.5 in room B. The corresponding probabilities of being caught by a specific surveillant in A are 0.4 and 0.2, and these numbers change into 0.3 and 0.7 respectively when the specific surveillant is in room B. The thief ("player I") has two actions (A or B). Together, the two guards ("player II") can choose between the following six actions: TT, AA, BB, TA, TB and AB (for instance, TA means that one guard is in T and the other in room A, etc. – for simplicity we suppose that the surveillants must stick to their choice and cannot switch rooms during the play). A natural payoff matrix is

$$P = \begin{pmatrix} .51 & .64 & .19 & .58 & .37 & .46 \\ .75 & .36 & .91 & .60 & .85 & .76 \end{pmatrix}$$

To understand these entries, observe for instance that when the thief chooses room A (i.e., row 1) and the surveillants choose action TA (i.e., column 4), then, assuming the surveillants operate *independently* from each other, the probability of the thief *not* being caught is 0.7 * 0.6 = 0.42. Thus, the corresponding entry in P is 1 - 0.42 = 0.58, etc. By reasoning just as in Example 3.3, we find as equilibrium strategies $\bar{x} = (0.8, 0.2)^t$ for the thief and $\bar{y} = (0, 1/15, 0, 14/15, 0, 0)^t$ for the surveillants.

Example 3.5 Let us consider a game that is *symmetric*, i.e. a game whose payoff matrix is skew-symmetric: $P = -P^t$. It is easy to see that by $P = -P^t$ Theorem 3.1 implies here that $v = 0$. Also, it is easy to see that in a symmetric game a mixed equilibrium pair $(\bar{x}, \bar{u})$ can always be taken in such a way that $\bar{x} = \bar{u}$. More concretely, let us consider the skew-symmetric payoff matrix

$$P := \begin{pmatrix} 0 & 1 & -2 \\ -1 & 0 & 3 \\ 2 & -3 & 0 \end{pmatrix}$$

and let $(\bar{x}, \bar{u})$ be as in Theorem 3.1. We wish to compute this pair explicitly. Let supp $\bar{x}$ be the support of $\bar{x}$.

Case 1: supp $\bar{x} = \{1\}$. In this case $\bar{x}$ is the unit vector $(1, 0, 0)^t$. But then we would have $0 = v_I = \sup_{u \in S_{II}} E(\bar{x}, u) = \sup_j P_{1j} = 1$, which is impossible. Similar reasoning excludes the other two cases where supp $\bar{x}$ is a singleton.

Case 2: supp $\bar{x} = \{1, 2\}$. In this case $\bar{x}_3 = 0$ and $\bar{x}_1, \bar{x}_2 > 0$. By player II's equalizing property we find $\bar{u}_2 - 2\bar{u}_3 = 0$, $-\bar{u}_1 + 3\bar{u}_3 = 0$, which leads to $\bar{u} = (1/2, 1/3, 1/6)^t$. But then the equalizing property of player I gives three equations, with solution $\bar{x} = (0, 0, 0)^t$, which is impossible. Similar reasoning excludes the other two cases where supp $\bar{x}$ has precisely two elements.

Case 3: supp $\bar{x} = \{1, 2, 3\}$. This is the only remaining case. The equalizing property for player II implies that $(P\bar{u})_i = 0$ for all i. The resulting 3 equations can easily be solved, and we find $\bar{u} = (1/2, 1/3, 1/6)^t$, just as in the previous case. This time, however, application of player I's equalizing property does not start from $\bar{x}_3 = 0$, which resulted in the impossible $\bar{x} = (0, 0, 0)^t$ in case 2, but rather from $\bar{x}_i > 0$ for all i, and this gives $\bar{x} = (1/2, 1/3, 1/6)^t$. The pair $(\bar{x}, \bar{u})$ just found is the desired mixed equilibrium solution.

References

[OSC] E.J. Balder, On subdifferential calculus, syllabus of September 2010.

[LDPD] E.J. Balder, Lagrangian duality and perturbational duality I, syllabus of October 2010.

[BSS] M.S. Bazaraa, H.D. Sherali and C.M. Shetty, *Nonlinear Programming: Theory and Algorithms.* Wiley, New York, 1993.

[BL] J.M. Borwein and A.S. Lewis, *Convex Analysis and Nonlinear Optimization.* Springer-Verlag, 2000.

[ET] I. Ekeland and R. Temam, *Convex Analysis and Variational Problems.* Classics in Applied Mathematics No. 28, SIAM, Philadelphia, 1999.

[KF] A. Kaufman and R. Faure, *Invitation à la Recherche Opérationelle*, Dunod, Paris, 1963.

[L] J.-P. Laurent, *Approximation et Optimisation*, Herrmann, Paris, 1972.

[R] R.T. Rockafellar, *Convex Analysis*, Princeton Press, 1970.