Working Paper

Noncooperative Convex Games: Computing Equilibrium By Partial Regularization

Sjur D. Flåm and Andrzej Ruszczyński

WP-94-42
May 1994

IIASA International Institute for Applied Systems Analysis □ A-2361 Laxenburg □ Austria
Telephone: +43 2236 71521 □ Telex: 079 137 iiasa a □ Telefax: +43 2236 71313

Noncooperative Convex Games: Computing Equilibrium By Partial Regularization

Sjur D. Flåm and Andrzej Ruszczyński

WP-94-42
May 1994

IIASA International Institute for Applied Systems Analysis ▫ A-2361 Laxenburg ▫ Austria
Telephone: +43 2236 71521 ▫ Telex: 079 137 iiasa a ▫ Telefax: +43 2236 71313

Abstract

A class of non-cooperative constrained games is analyzed for which the Ky Fan function is convex-concave. Nash equilibria of such games correspond to diagonal saddle points of the said function. This feature is exploited in designing computational algorithms for finding such equilibria.

Key words: Noncooperative constrained games, Nash equilibrium, subgradient projection, proximal point algorithm, partial regularization, saddle points, Ky Fan functions.

Noncooperative Convex Games: Computing Equilibrium By Partial Regularization

Sjur D. Flåm[1] *and Andrzej Ruszczyński*

1. Introduction

We shall consider noncooperative games on the following strategic (normal) form: each individual i belonging to a finite set I seeks, without any collaboration, to minimize his private loss (cost) $L_i(x) = L_i(x_i, x_{-i})$ with respect to his own strategy $x_i \in X_i \subset R^{n_i}$. In doing so the players are jointly restricted by the coupling constraint that $x \in C \subset R^n$, $n := \sum_{i \in I} n_i$. Here $x_{-i} := (x_j)_{j \in I \setminus i}$ is short notation for actions taken by player i's adversaries. Our problem is finding a *Nash equilibrium* x^*. By definition, such outcomes satisfy, for all $i \in I$, the optimality condition that x_i^* minimizes $L_i(x_i, x_{-i}^*)$ subject to $x_i \in X_i$ and $(x_i, x_{-i}^*) \in C$. A *normalized* equilibrium is a point $x^* \in X$ that minimizes $\sum_{i \in I} L_i(x_i, x_{-i}^*)$ subject to $x \in X := \prod_{i \in I} X_i \cap C$. Obviously, normalized equilibria are Nash equlibria, but the converse is not true in general, unless $C \supset \prod_{i \in I} X_i$.

Points of this particular sort are available under reasonable conditions. Indeed, supposing throughout that $\prod_{i \in I} X_i$ and C are convex with nonempty intersection X, we have - slightly generalizing (Rosen 1965, Thm.1) - the following result.

Proposition 1. (Existence of normalized equilibria) *Suppose X is compact. Also suppose that all functions $L_i(x_i, x_{-i})$ are jointly continuous on X and quasi-convex in x_i. Then there exists at least one normalized equilibrium.*

Proof. Consider the multi-valued mapping $F : X \to X$ given by

$$F(x) := \{x' \mid x' \text{ minimizes } \sum_i L_i(x_i', x_{-i}) \text{ over } X\}.$$

This mapping is upper semicontinuous with nonempty convex values. Whence, by Kakutanis fixed point theorem (see, e.g., (Aubin and Ekeland, 1984) , thm. 6.4.19), there exists an $x^* \in X$ such that $x^* \in F(x^*)$. Evidently, such a point x^* is a normalized equilibrium. □

Our concern here is with computation of such equilibria, not their existence. So henceforth, we take existence for granted and find it expedient that *each cost function $L_i(x_i, x_{-i})$*

[1]Support from IIASA and Ruhrgas is gratefully acknowledged

be convex in x_i and finite-valued near X. Equilibrium is then fully characterized by essential marginal costs, that is, by partial subdifferentials $M_i(x) := \partial_{x_i} L_i(x)$ and normal cones. Indeed, letting $N(\cdot)$ be the normal cone correspondence of X; P the orthogonal projection onto X, and $M(x) := (M_i(x))_{i \in I}$, we have by standard optimality conditions of convex programming (Rockafellar 1970), the following necessary and sufficient conditions.

Proposition 2. (Equilibria occur where essential marginal costs are zero) *The following three statements are necessary and sufficient for $x^* \in X$ to be a normalized equilibrium:*

(a) $\exists g \in M(x)$ such that $\langle g, x - x^* \rangle \geq 0$ for all $x \in X$;

(b) $0 \in M(x^*) + N(x^*)$;

(c) $x^* \in P[x^* - sM(x^*)]$ for all $s > 0$.

Conditions (a)-(c) beg the use of established techniques. In particular, (a) calls for algorithms applicable to solve variational inequalities (Harker and Pang 1990). Likewise, (b) directs attention to proximal point procedures (Rockafellar 1976), and especially, to splitting methods (Eckstein and Bertsekas 1992). Finally, (c) indicates that subgradient projections might offer a good avenue (Ermoliev and Uryasiev 1982), (Cavazzuti and Flåm 1992). In any event, to make good progress along any of these lines, it is desirable that the marginal cost correspondence $x \to M(x)$ be monotone in some appropriate sense. However, even with $x \to M(x)$ maximal monotone, each of the said approaches suffers from some difficulties. To wit, proximal point procedures, including those using splitting techniques - while yielding good convergence - are often difficult to implement. They typically require iterative solutions of similar perturbed games, each being almost as difficult to handle as the original one. Subgradient projection has opposite properties: implementation tends to come easy, but the method often produces exceedingly slow convergence.

These observations lead us to specialize on the data of the game, and to approach computation along different lines. Namely, first recall that $x^* \in X$ is a normalized equilibrium if and only if the *Ky Fan function*

$$L(x, y) := \sum_{i \in I} [L_i(x) - L_i(y_i, x_{-i})] \tag{1.1}$$

satisfies

$$\sup_{y \in X} L(x^*, y) \leq 0.$$

Second, when solving this last inequality system for x^*, it largely helps that $L(x, y)$ be convex in x. These last facts organize the inquiry below, and they make us focus on a special class of games, declared *convex.* These are the ones having a convex-concave Ky Fan function $L(x, y)$ (1.1). It turns out that such games admit Nash equilibria that furnish diagonal saddle points of the Ky Fan function. For motivation, some important examples are exposed in Section 2. Thereafter we proceed to our main object, that is, to compute equilibria for such games. Thus Section 3 brings out new algorithms, using partial regularizations, relaxed subgradient projections and averages of proposed solutions. The algorithms are specialized versions of general methods for saddle point seeking discussed in (Kallio and Ruszczyński, 1994).

2. Convex Games

As said, focus will be on games having convex-concave Ky Fan functions $L(x, y)$ (1.1). Such games may serve as standard models in their own right or as approximations to more complex data. The class at hand is more rich than might first be imagined.

Proposition 3. *Any zero-sum, two-person game with convex-concave cost $\Lambda(x_1, x_2)$ of player 1, is convex.*

Proof. Since $L(x, y) = \Lambda(x_1, y_2) - \Lambda(y_1, x_2)$, the conclusion is immediate. □

Proposition 4. **(a)** *Suppose the marginal cost correspondence $x \to M(x)$ is monotone and differentiable. If each cost function $L_i(x)$ is replaced by its second order approximation*

$$L_i(x^0) + \langle L_i'(x^0), x - x^0\rangle + \frac{1}{2}\langle x - x^0, L_i''(x^0)(x - x^0)\rangle,$$

then the resulting approximate game is convex.

(b) *Conversely, suppose each cost function $L_i(x)$ is exactly linear-quadratic in the sense that*

$$L_i(x) := \sum_{j \in I} [a_{ij}^T + x_i^T B_{ij}] x_j$$

for vectors a_{ij} and matrices B_{ij} of appropriate dimensions. Then, if the matrix

$$\begin{bmatrix} 2B_{11} & B_{12} & B_{13} & \cdots \\ B_{21} & 2B_{22} & B_{23} & \cdots \\ B_{31} & B_{32} & 2B_{33} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} \tag{2.1}$$

is positive semidefinite, the game is convex.

(c) *More generally, suppose individual loss equals*

$$L_i(x) := \Lambda_i(x_i) + x_i^T B_i x_{-i} + \Lambda_{-i}(x_{-i})$$

with $\Lambda_i(x_i)$ convex and satisfying

$$\Lambda_i(x_i') - \Lambda_i(x_i) \geq \langle g_i(x_i), x_i' - x_i\rangle + \frac{1}{2}\langle x_i' - x_i, B_{ii}(x_i)(x_i' - x_i)\rangle$$

for all x_i, x_i' and $g_i(x_i) \in \partial\Lambda_i(x_i)$. If, after replacing the i-th diagonal element of the matrix (2.1) by $2B_{ii}(x_i)$, we get a positive semidefinite matrix, then the game is convex.

Proof. In instance (b) define

$$\Lambda_i(x_i) := [a_{ii}^T + x_i^T B_{ii}] x_i;$$

$$x_i^T B_i x_{-i} := x_i^T \sum_{j \in I \setminus i} B_{ij} x_j$$

and

$$\Lambda_{-i}(x_{-i}) := \sum_{j \in I \setminus i} a_{ij}^T x_j$$

to have

$$L(x,y) = \sum_{i \in I} \left[\Lambda_i(x_i) + x_i^T B_i x_{-i} - \Lambda_i(y_i) - y_i^T B_i x_{-i} \right], \tag{2.2}$$

this form covering both cases (ii) and (iii). Case (i) also yields a form fitting (2.2). Evidently, the latter function $L(x,y)$ is concave in y, and upon differentiating twice with respect to x, partial convexity in that variable follows. □

We notice that any saddle point (x^*, y^*) of L furnishes a normalized equlibrium x^*. This simple observation makes us inquire whether a normalized equilibrium x^* can be duplicated to constitute a saddle point (x^*, x^*). As brought out in the next proposition, the answer is positive.

We start from the following technical result.

Lemma 1. *Assume that the game is convex. Then for every point $x \in X$*

$$\partial_x L(x,x) = -\partial_y L(x,x).$$

Proof. Define $h = x' - x$ with x' in a small neigborhood of x. By convexity of L with respect to x, for every $\alpha \in (0,1)$,

$$\alpha L(x+h, x+\alpha h) + (1-\alpha) L(x, x+\alpha h) \geq L(x+\alpha h, x+\alpha h) = 0.$$

Dividing by α and passing to the limit with $\alpha \downarrow 0$ we obtain

$$L(x+h, x) + \lim_{\alpha \downarrow 0} \left[\alpha^{-1} L(x, x+\alpha h) \right] \geq 0.$$

By concavity of L in the second argument, for every $g \in \partial_y L(x,x)$,

$$\lim_{\alpha \downarrow 0} \left[\alpha^{-1} L(x, x+\alpha h) \right] \leq \langle g, h \rangle.$$

Thus

$$L(x+h, x) \geq \langle -g, h \rangle.$$

Since $x+h$ can be arbitrary in a sufficiently small neighborhood of x (such that all function values in the analysis above are finite), $-g \in \partial_x L(x,x)$. Consequently,

$$\partial_x L(x,x) \supset -\partial_y L(x,x).$$

In a symmetric way, we can prove the converse inclusion. □

We can now state the main result.

Proposition 5. *If the game is convex, then the following statements are equivalent:*

(a) x^* *is a normalized equilibrium;*

(b) $\sup_{y \in X} L(x^*, y) = 0$;

(c) $\inf_{x \in X} L(x, x^*) = 0$;

(d) (x^*, x^*) *is a saddle point of* L *on* $X \times X$.

Proof.
(a)⇔(b). The equivalence follows directly from the definition of a normalized equilibrium.
(b)⇔(c). From (b) it follows that there is $g \in \partial_y L(x^*, x^*)$ such that $\langle g, x - x^* \rangle \leq 0$ for all $x \in X$. By Lemma 1, $-g \in \partial_x L(x^*, x^*)$, so

$$L(x, x^*) \geq \langle -g, x - x^* \rangle \geq 0 = L(x^*, x^*)$$

for every $x \in X$. The converse implication can be proved analogously.
((b)∧(c))⇔(d). The equivalence is obvious, because $L(x^*, x^*) = 0$. □

We can use this result to derive the following sufficient condition for the existence of normalized equilibria.

Proposition 6. *Assume that the game is convex and there exists a point* $\tilde{x} \in X$ *such that the function* $L(\tilde{x}, \cdot)$ *is sup-compact. Then there exists a normalized equilibrium.*

Proof. From theorem 6.2.7 of (Aubin and Ekeland, 1984) we conclude that there exists a point $x^* \in X$ satisfying condition (c) of Proposition 5. □

We end this section with two examples of convex games having normalized equilibria which furnish saddle points of the Ky Fan function.

Example 1: *Cournot oligopoly*

An important instance of noncooperative games is the classical oligopoly model of Cournot (1838). This model remains a workhorse within modern theories of industrial organization (Tirole 1988). Generalizing it to comprise k different goods, the model goes as follows: Firm $i \in I$ produces the commodity bundle $x_i \in R^k$, thus incurring convex production cost $c_i(x_i)$ and gaining market revenues $\langle p(\sum_{j \in I} x_j), x_i \rangle$. Here $p(\sum_{j \in I} x_j)$ is the price vector at which total demand equals the aggregate supply $\sum_{j \in I} x_j$. Suppose this inverse demand curve is affine and "slopes downwards" in the sense that $p(Q) = a - CQ$ where $a \in R^k$ and C is a $k \times k$ positive semidefinite matrix. Then

$$L_i(x) = c_i(x_i) - \langle a, x_i \rangle + \sum_{j \in I} \langle x_i, C x_j \rangle$$

and the resulting Cournot oligopoly is convex.

Example 2: *Multi-person matrix games*

Suppose that each player $i \in I$ has a finite number n_i of pure strategies. The cost of the ith player, when he uses strategy k_i and other players $j \neq i$ play strategies k_j, equals

$$l_i(k_i, k_{-i}) = \sum_{j \in I \setminus i} B_{ij}(k_i, k_j),$$

where $B_{ij}(k_i, k_j)$ denotes the (k_i, k_j)-th element of the pay-off matrix B_{ij}. Passing to mixed strategies $x_i \in X_i$, where X_i is the standard simplex in R^{n_i}, we get

$$L_i(x_i, x_{-i}) = \sum_{j \in I \setminus i} \langle x_i, B_{ij} x_j \rangle.$$

If the transfers for each pair of players sum to zero, i.e., if $B_{ij} = -B_{ji}^T$ for all $i \neq j$, the game is convex.

3. Partial Regularization Methods

Our purpose is to find a Nash equilibrium using only iterative, single-agent programming. In that endeavour, we shall adapt the general idea of a saddle point seeking procedure discussed in (Kallio and Ruszczyński, 1994). In our application the idea can be conceptually interpreted as follows.

Besides the individuals $i \in I$ actually considering a profile $y = (y_i)_{i \in I} \in X$, let there be a *coordinating agent* currently proposing the strategy profile $x \in X$. This agent predicts that individual $i \in I$ will select his next strategy

$$y_i^+ \in \arg\min L_i(\cdot, x_{-i})$$

thus generating a private cost reduction $L_i(x) - L_i(y_i^+, x_{-i})$. The agent is concerned with the overall cost reduction or regret

$$L(x, y^+) := \sum_i \left[L_i(x) - L_i(y_i^+, x_{-i}) \right]$$

of implementing x. To reduce such regret he would, if possible, change x in a "descent" direction

$$d_x \in -\partial_x L(x, y^+).$$

Similarly, individual $i \in I$ predicts that the coordinating agent will propose a profile x^+ satisfying $L(x^+, y) \leq 0$ or, a fortiori,

$$x^+ \in \arg\min L(\cdot, y).$$

Such beliefs induce a change of y_i, if possible, in a "descent" direction

$$d_{y_i} \in -\partial_{y_i} L_i(y_i, x_{-i}^+).$$

These loose ideas were intended to motivate and advertise the subsequent algorithms. They must be refined at several spots. First, some stability or inertia is needed in the predictions. For that purpose we shall introduce regularizing penalties of a quadratic nature. Second, the directions must be feasible. In this regard we shall rely on projections, also designed to enforce global, non-decomposable constraints. Third, when updating x and y along the proposed directions, appropriate step sizes are needed. At this juncture some techniques from subgradient projection methods will serve us well. Finally, equality of the coordinating profile and the individual strategies is maintained by introducing a special *compromise step* which generates an average proposal.

All the matters are accounted for and incorporated in the following algorithms.

ALGORITHM 1 (*Partial regularization in individual strategies*)

Initialization: Select an arbitrary starting point $x^0 \in X$ and set $\nu := 0$.

Predict individual strategies: Compute

$$y^{\nu+} := \arg\min \left\{ \sum_{i \in I} L_i(y_i, x_{-i}^{\nu}) + \frac{\rho}{2}\|y - x^{\nu}\|^2 \ : \ y \in X \right\}. \tag{3.1}$$

Test for termination: If $y^{\nu+} = x^{\nu}$, then stop: x^{ν} solves the problem.

Predict the coordinating strategy: Find $x^{\nu+} \in X$ such that

$$L(x^{\nu+}, x^{\nu}) \leq 0$$

and $\|x^{\nu+} - x^{\nu}\| \leq K$ for some constant K. In particular, $x^{\nu+} = x^{\nu}$ is one option.

Find directions of improvement: Select subgradients $g_x^{\nu} \in \partial_x L(x^{\nu}, y^{\nu+})$ and $g_{yi}^{\nu} \in \partial_{x_i} L_i(x_i^{\nu}, x_{-i}^{\nu+})$, $i \in I$, and define a direction $d^{\nu} := (d_x^{\nu}, d_y^{\nu})$ with

$$d_x^{\nu} := P_x^{\nu}\left[-g_x^{\nu}\right], \quad d_y^{\nu} := P_y^{\nu}[-g_y^{\nu}],$$

where P_x^{ν}, P_y^{ν} denote orthogonal projections onto closed convex cones T_x^{ν}, T_y^{ν} containing the tangent cone $T(x^{\nu})$ of X.

Calculate the step size: Let

$$\tau_{\nu} = \frac{\gamma_{\nu}\left[L(x^{\nu}, y^{\nu+}) - L(x^{\nu+}, x^{\nu})\right]}{\|d^{\nu}\|^2},$$

with $0 < \gamma_{\min} \leq \gamma_{\nu} \leq \gamma_{\max} < 2$.

Make a step: Update by the rules

$$x^{\nu++} := P\left[x^{\nu} + \tau_{\nu} d_x^{\nu}\right],$$

$$y^{\nu++} := P\left[x^{\nu} + \tau_{\nu} d_y^{\nu}\right],$$

where P is the orthogonal projection onto X.

Make a compromise: Let

$$x^{\nu+1} = \frac{1}{2}\left(x^{\nu++} + y^{\nu++}\right).$$

Increase ν by 1 and continue to predict strategies.

The second version of the method is symmetric to the first one, but this time with a more exact prediction of the coordinating strategy.

ALGORITHM 2 (*Partial regularization in the coordinating variable*)

The method proceeds as Algorithm 1, with the only difference in the prediction steps, which are replaced by the following.

Predict individual strategies: Find $y^{\nu+} \in X$ such that

$$\sum_{i \in I} L_i(y_i^{\nu+}, x_{-i}^{\nu}) \leq \sum_{i \in I} L_i(x_i^{\nu}, x_{-i}^{\nu})$$

and $\|y^{\nu+} - x^{\nu}\| \leq K$ for some constant K. In particular, $y^{\nu+} = y^{\nu}$ is an easy and acceptable choice.

Predict the coordinating strategy: Compute

$$x^{\nu+} \in \arg\min\left\{L(x, x^{\nu}) + \frac{\rho}{2}\|x - x^{\nu}\|^2 \; : \; x \in X\right\}. \tag{3.2}$$

Observe that in the absence of common constraints, i.e. with $X = \prod_{i \in I} X_i$, the prediction (3.1) of individual strategies decomposes into separate subproblems, one for each player. To execute (3.2) is generally more difficult than (3.1), given that $L(x, y)$ typically is less separable in x than in y.

4. Convergence

It simplifies the exposition to single out a key observation; namely, that our algorithmic mapping is a Fejér mapping (see (Eremin and Astafiev, 1976) and (Polyak, 1969)).

Lemma 2. *Assume that the game is convex and has a normalized equilibrium x^*. Define*

$$W_\nu := \|x^\nu - x^*\|^2,$$

where $\{x^\nu\}$ is the sequence generated by any of the two algorithms defined in the previous section. Then, for all ν

$$W_{\nu+1} \leq W_\nu - \frac{1}{2}\gamma_\nu(2 - \gamma_\nu)[L(x^\nu, y^{\nu+}) - L(x^{\nu+}, y^\nu)]^2/\|d^\nu\|^2.$$

Proof. Invoking the non-expansiveness of projection, we have

$$\begin{aligned}\|x^{\nu++} - x^*\|^2 &= \|P[x^\nu + \tau_\nu d_x^\nu] - P[x^*]\|^2 \\ &\leq \|x^\nu + \tau_\nu d_x^\nu - x^*\|^2 \\ &= \|x^\nu - x^*\|^2 + 2\tau_\nu \langle d_x^\nu, x^\nu - x^*\rangle + \tau_\nu^2 \|d_x^\nu\|^2.\end{aligned}$$

Use now the orthogonal decomposition $-g_x^\nu = d_x^\nu + n_x^\nu$, n_x^ν being in the negative polar cone of T_x^ν, and observe that $x^* - x^\nu \in T_x^\nu$, to obtain

$$\langle d_x^\nu, x^\nu - x^*\rangle \leq \langle g_x^\nu, x^* - x^\nu\rangle \leq L(x^*, y^{\nu+}) - L(x^\nu, y^{\nu+}).$$

Whence,

$$\|x^{\nu++} - x^*\|^2 \leq \|x^\nu - x^*\|^2 - 2\tau_\nu[L(x^\nu, y^{\nu+}) - L(x^*, y^{\nu+})] + \tau_\nu^2\|d_x^\nu\|^2.$$

Similarly,

$$\|y^{\nu++} - x^*\|^2 \leq \|x^\nu - x^*\|^2 + 2\tau_\nu[L(x^{\nu+}, x^\nu) - L(x^{\nu+}, x^*)] + \tau_\nu^2\|d_y^\nu\|^2.$$

By convexity of the squared norm,

$$\|x^{\nu+1} - x^*\|^2 \leq \frac{1}{2}\left(\|x^{\nu++} - x^*\|^2 + \|y^{\nu++} - x^*\|^2\right).$$

Combining the last three inequalities we have

$$W_{\nu+1} \leq W_\nu - \tau_\nu[L(x^\nu, y^{\nu+}) - L(x^*, y^{\nu+}) - L(x^{\nu+}, x^\nu) + L(x^{\nu+}, x^*)] + \frac{1}{2}\tau_\nu^2\|d^\nu\|^2.$$

Since, by Proposition 5, (x^*, x^*) is a saddle point of L, it follows that $L(x^*, y^{\nu+}) \leq L(x^{\nu+}, x^*)$. Therefore

$$W_{\nu+1} \leq W_\nu - \tau_\nu[L(x^\nu, y^{\nu+}) - L(x^{\nu+}, x^\nu)] + \frac{1}{2}\tau_\nu^2\|d^\nu\|^2.$$

Here apply the stepsize rule to arrive at the required result. □

The first convergence result can now be stated forthwith.

Theorem 1. *Assume that the game is convex and has a normalized equilibrium x^*. Then the sequence $\{x^\nu\}$ generated by Algorithm 1 is convergent to a normalized equilibrium.*

Proof. Since $L(x^{\nu+}, x^\nu) \leq L(x^\nu, x^\nu) = 0$ and $L(x^\nu, y^{\nu+}) \geq L(x^\nu, x^\nu) = 0$, from Lemma 1 we obtain,

$$W_{\nu+1} \leq W_\nu - \frac{1}{2}\gamma_\nu(2 - \gamma_\nu)[L(x^\nu, y^{\nu+})]^2/\|d^\nu\|^2.$$

Evidently $\{W_\nu\}$ is non-increasing, hence bounded. The sequence $\{d^\nu\}$ is bounded, so $L(x^\nu, y^{\nu+}) \to 0$. By the definition of $y^{\nu+}$, there exists a subgradient $g \in \partial_y L(x^\nu, y^{\nu+})$ such that

$$\langle g - \rho(y^{\nu+} - x^\nu), h\rangle \leq 0,$$

for every feasible direction h at $y^{\nu+}$. Thus, with $h = x^\nu - y^{\nu+}$, one has

$$L(x^\nu, x^\nu) - L(x^\nu, y^{\nu+}) \le \langle g, x^\nu - y^{\nu+} \rangle \le -\rho \|y^{\nu+} - x^\nu\|^2,$$

so

$$L(x^\nu, y^{\nu+}) \ge \rho \|y^{\nu+} - x^\nu\|^2.$$

Consequently,

$$\lim_{\nu \to \infty} \|y^{\nu+} - x^\nu\|^2 = 0.$$

Let $\hat{x}$ be an accumulation point of $\{x^\nu\}$ and y^+ be the associated accumulation point of $\{y^{\nu+}\}$. Then $y^+ = \hat{x}$, i.e.,

$$\hat{x} = \arg\min \left\{ \sum_{i \in I} L_i(y_i, \hat{x}_{-i}) + \frac{\rho}{2} \|y - \hat{x}\|^2 \ : \ y \in X \right\}.$$

This is necessary and sufficient for $\hat{x}$ to be a normalized equilibrium. Substituting it for x^* in the definition of W_ν, we conclude that the distance to $\hat{x}$ is non-increasing. Consequently, $\hat{x}$ is the only accumulation point of the sequence $\{x^\nu\}$. □

Theorem 2. *Assume that the game is convex and has a normalized equilibrium x^*. Then the sequence $\{x^\nu\}$ generated by Algorithm 2 is convergent to a normalized equilibrium.*

Proof. Proceeeding analogously to the proof of Theorem 2 we arrive at the relation:

$$\lim_{\nu \to \infty} \|x^{\nu+} - x^\nu\|^2 = 0.$$

Let $\hat{x}$ be an accumulation point of $\{x^\nu\}$ and x^+ be the associated accumulation point of $\{x^{\nu+}\}$. Then $x^+ = \hat{x}$, i.e.,

$$\hat{x} = \arg\min \left\{ L(x_i, \hat{x}_{-i}) + \frac{\rho}{2} \|x - \hat{x}\|^2 \ : \ x \in X \right\}.$$

By Proposition 5, this is necessary and sufficient for $\hat{x}$ to be a normalized equilibrium. Substituting it for x^* in the definition of W_ν, we conclude that $\hat{x}$ is the limit of the sequence $\{x^\nu\}$. □

The preceding developments can accomodate several minor modifications. First, the proximal parameter ρ may vary, provided it remains bounded away from 0 and ∞. Second, when designing the compromise, instead of equal weights $\frac{1}{2}$, we may apply stage-varying weights $\alpha_x^\nu \ge 0$, $\alpha_y^\nu \ge 0$, $\alpha_x^\nu + \alpha_y^\nu = 1$, provided the weight of the variable for which the calculation of direction was more elaborate (with minimization in the prediction step for the other variable) is bounded away from 0.

Finally, it should be stressed that instead of proximal operators in the prediction steps, we can use more general mappings with similar properties (see (Kallio and Ruszczyński, 1994)). We choose to present the idea with the use of quadratic regularizations just for simplicity, to avoid obscuring it with technical details.

References

[1] J.-P. Aubin and I. Ekeland, *Appplied Nonlinear Analysis*, John Wiley and Sons, New York (1984).

[2] E. Cavazzuti and S.D. Flåm, "Evolution to selected Nash equilibria," in F. Giannessi (ed.), *Nonsmooth Optimization Methods and Applications*, Gordon and Breach, London 30-41 (1992).

[3] A. Cournot, *Recherches sur les Principes Mathématiques de la Théorie des Richesses*, Paris(1838), English translation:*Researches into the mathematical principles of the theory of wealth* (ed. N. Bacon) Macmillan, New York(1897).

[4] J. Eckstein and D.P. Bertsekas, "On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators," *Mathematical Programming* 55 (3) 293-318 (1992).

[5] I.I. Eremin and N.N. Astafiev, *Introduction to the Theory of Linear and Convex Programming* (Nauka, Moscow, 1976).

[6] Yu.M. Ermoliev and S.P. Uryasiev, "On seeking a Nash equilibrium in n-person games," (in Russian) *Kibernetika* 3, 85-88 (1982).

[7] P.T. Harker and J.-S. Pang, "Finite-dimensional variational inequalities and nonlinear complementarity problems: A survey of theory, algorithms and applications," *Mathematical Programming* 48, 2, 161-220 (1990).

[8] M. Kallio and A. Ruszczyński, "Perturbation methods for saddle point computation," Working Paper WP-94-38, IIASA (1994).

[9] B.T. Polyak, "Minimization of nonsmooth functionals," *Zhurnal Vychislitelnoi Matematiki i Matematicheskoi Fiziki* 9 (1969) 509-521.

[10] R.T.Rockafellar, *Convex analysis*, Princeton University Press (1970).

[11] R.T. Rockafellar, "Monotone operators and the proximal point algorithm," *SIAM J. Control and Optimization* 14 (5) 877-898(1976).

[12] J.B. Rosen, "Existence and uniqueness of equilibrium points for concave n-person games," *Econometrica* 33 (3) 520-534 (1965).

[13] J. Tirole, *The Theory of Industrial Organization*, MIT Press, Cambridge (1988).