Nonlinear Equilibrium vs. Linear Programming.

Roman A Polyak
George Mason University
Fairfax, VA 22030, USA

January 23, 2009

In memory of Professor Simon Israilevich Zuhovitsky.

In the early 60s the emerging field of Modern Optimization became the main focus of S. I.'s research.

The author and his late friend Matvey Primak were very lucky: it was the starting point of our long lasting collaboration with S. I, who opened for us the beautiful world of Mathematics and helped us to enter the world. For us it was a unique opportunity which would have been impossible without S.I. In fact S. I. saved us from spiritual death and changed our lives forever.

Abstract

The Generalized Walras -Wald Equilibrium (GE) was introduced by S.I.Zuchovitsky et al. in 1973(see [17]) as an alternative to the Linear Programming (LP) approach for optimal resource allocation.

There are two fundamental differences between the GE and LP approach for the best resource allocation.

First, the prices for goods (products) are not fixed as they are in LP; they are functions of the production output.

Second, the factors (resources) used in the production process are not fixed either; they are functions of the prices for resources.

In this paper we show that under natural economic assumptions on both price and factor functions the GE exists and unique. Finding the GE is equivalent to solving a variational inequality with a strongly monotone operator.

For solving the variational inequality, we introduce a projected pseudo–gradient method.

We prove that under the same assumptions on price and factor functions the projected pseudo–gradient method converges globally with Q-linear rate. This allows estimation of its computational complexity and finding parameters critical for the complexity bound.

The method can be viewed as a natural pricing mechanism for establishing an economic equilibrium.

Key Words: generalized equilibrium, variational inequality, projected sub–gradient method, Q–linear convergence rate.
MSC200 subject classification: Primary 90C30, 90C46, secondary 90C05.

1 Introduction

In the late 1930s L.V.Kantorovich in St Petersburg (then Leningrad) discovered that a number of real life problems arising in technological, engineering and economic applications lead to finding a maximum (minimum) of a linear function under linear inequality constraints.

$$(c, x^*) = \max \{(c, x) | Ax \leq b, x \geq 0\} \tag{1}$$

where $A : \mathbb{R}^n \to \mathbb{R}^m; c, x \in \mathbb{R}^n, b \in \mathbb{R}^m$. This was the beginning of the Linear Programming (LP) era.

During the Second World War and immediately after it become clear that LP has important military applications. In 1947 George Dantzig developed the Simplex method for solving LP, which was judged one of the ten best algorithms in the 20 century. Almost at the same time John von Neumann in a conversation with G.Dantzig introduced LP Duality, which became the foundation of LP theory.

With each LP problem (1) is associated the co-called dual problem:

$$(b, \lambda^*) = \min \left\{(b, \lambda) | A^T\lambda \geq c, \lambda \geq 0\right\} \tag{2}$$

The following Duality relations are fundamental for LP:

$$(c, x^*) = (b, \lambda^*) \tag{3}$$

and

$$(A^T\lambda^* - c, x^*) = 0 \qquad (b - Ax^*, \lambda^*) = 0, \tag{4}$$

Since then the interest to LP grew dramatically and over the last 60 years LP became one of the most advanced fields in modern optimization.

In 1949 S.I.Zuchovitsky in Kiev introduced a method for finding Chebyshev approximation for a system of linear equations $Ax = b$ i.e. finding

$$F(x^*) = \min\{F(x) | x \in \mathbb{R}^n\} \tag{5}$$

where $F(x) = \max\{|(Ax - b)_i| \quad | \quad i = 1, \ldots, m\}$ is a convex and piecewise linear function. The method was a realization of Fourier's idea: in order to find $F(x^*)$ one has to start at any point on a descent edge of the piecewise linear surface and following along the descent edges move from one vertex to another until the bottom of the surface will be reached.

S.I. was not aware of G.Dantzig's results; the "Iron Curtain" made it impossible. It would take another ten years before S.I. learned from E.Stiefel's paper [14] that his algorithm for solving the problem (5) is in fact the simplex method for solving the following equivalent to (5) LP problem

$$y^* = \min\{y| \quad |(Ax-b)_i| \le y, i = 1, \dots, m\} \tag{6}$$

From this point on Optimization became the main focus of S. I. research.

The efficiency of the original simplex method has been drastically improved over the last sixty years, and new very powerful interior point methods for solving LP problems have emerged in the last 20 years.

In 1975 L.V.Kantorovich and T.C.Koopmans shared the Nobel Prize in Economics "for their contributions to the theory of optimum allocation of limited resources".

The question however is: How adequately does LP reflect the economic reality when it comes to the best allocation of limited resources.

Two fundamental LP assumptions are:

a) The price vector $c = (c_1, \dots, c_n)$ for produced goods is fixed and independent of the production output vector $x = (x_1, \dots, x_n)$.

b) The resource vector $b = (b_1, \dots, b_m)$ is fixed and independent of the prices $\lambda = (\lambda_1, \dots, \lambda_n)$ for the resources.

   There are a few essential difficulties associated with these assumptions.

   1. Unfortunately such assumptions do not reflect the basic market law of supply and demand. Therefore the LP models might lead to solutions which are not always practical.
   2. Due to these assumptions a very small change of at least one component at the price vector c might lead to a drastic change of the primal solution.
   3. Small variations of the resource vector b can lead to a dramatic change of the dual solution.

The purpose of the paper is to develop an alternative to the LP approach for resources allocation which is based on GE [17].

The fixed vector c is replaced by a vector–function $c(x)$: the price for a good depends on the output $x = (x_1, \dots, x_n)$. The fixed factor vector b is replaced by a vector–function $b(\lambda)$: the factor availability depends on the prices $\lambda = (\lambda_1, \dots, \lambda_m)$ for the factors.

We introduce the notion of well defined vector–functions $c(x)$ and $b(\lambda)$ and show that for such functions the GE equilibrium exists and is unique.

Then we show that finding the GE is equivalent to solving a variational inequality with a strongly monotone operator on the Cartesian product of non-negative octants of the primal and dual spaces.

For solving the variational inequality we use the projected pseudo-gradient method and show that for well defined vector functions $c(x)$ and $b(\lambda)$ the method globally converges with

Q–linear rate. We estimate the computational complexity of the method and compare it with LP complexity.

The paper is organized as follows. In the following section we state the problem and describe the basic assumptions on the input data.

In the third section we provide some background related to the Nash equilibria of n-person concave games and the classical Walras-Wald equilibrium.

In section four we consider existence and uniqueness of the GE and show that finding the equilibrium is equivalent to solving a variational inequality for a strongly monotone operator on $\Omega = \mathbb{R}^n_+ \otimes \mathbb{R}^m_+$.

In section five we consider the projected pseudo–gradient method, establish its global convergence with Q–linear rate and estimate its computational complexity.

We conclude the paper with some remarks related to future research.

2 Statement of the problem and the basic assumptions.

We consider an economy which produces n goods (products) by consuming m factors (resources) in the production process. There are three sets of data required for problem formulation.

1) The technology matrix $A : \mathbb{R}^n \rightarrow \mathbb{R}^m$ which "transforms" production factors into goods,i.e. $a_{ij} \geq 0$ defines the amount of factor $1 \leq i \leq m$ is required to produce one item of good $1 \leq j \leq n$.

2) The price vector–function $c(x) = (c_1(x), \dots, c_n(x))$, where $c_j(x)$ is the price for one item of good j under the production output $x = (x_1, \dots, x_j \dots, x_n)$.

3) The factor vector–function $b(\lambda) = (b_1(\lambda), \dots, b_i(\lambda) \dots, b_m(\lambda))$, where $b_i(\lambda)$ is the availability of the factor i under the price vector $\lambda = (\lambda_1, \dots, \lambda_i, \dots, \lambda_n)$.

We will use the following assumptions.

A. The matrix A does not have zero rows or columns, which means that each factor is used for production at least one of the goods and each good requires of at least one of the production factors.

B. The price vector–function $c : \mathbb{R}^n_+ \rightarrow \mathbb{R}^n_+$ for goods is continuous and strongly monotone decreasing, i.e. there is $\alpha > 0$ such that

$$(c(x^2) - c(x^1), x^2 - x^1) \leq -\alpha \|x^1 - x^2\|^2, \quad \forall x^1, x^2 \in \mathbb{R}^n_+ \tag{7}$$

C. The factor vector–function $b : \mathbb{R}^m_+ \rightarrow \mathbb{R}^m_+$ is continuous and strongly monotone increasing, i.e. there is $\beta > 0$ such that

$$(b(\lambda^2) - b(\lambda^1), \lambda^2 - \lambda^1) \geq \beta \|\lambda^1 - \lambda^2\|^2, \quad \forall \lambda^1, \lambda^2 \in \mathbb{R}^m_+ \tag{8}$$

where $\|.\|$ is the Euclidean norm.

Assumption B implies that an increase of production x_j of good $1 \leq j \leq n$ when the rest is fixed leads to decrease of the price $c_j(x)$ per item of good j, moreover the margin of the price decrease has a negative upper bound.

Assumption C implies that the increase of the price λ_i of any factor $1 \leq i \leq m$ when the rest is fixed leads to the increase of the availability of the resource $b_i(\lambda)$ and the margin for the resource increase has a positive lower bound.

We will say that the price $c(x)$ and the factor $b(x)$ vector–functions are well defined if (7) and (8) hold.

It is worth mentioning that (7) is true if $c(x)$ is a gradient of a strongly concave function defined on $\mathbb{R}^n_+$, whereas (8) is true if $b(\lambda)$ is a gradient of a strongly convex function on $\mathbb{R}^m_+$.

The GE problem consists of finding $x^* \in \mathbb{R}^n_+$ and $\lambda^* \in \mathbb{R}^m_+$ such that

$$(c(x^*), x^*) = \max\{(c(x^*), X) | AX \leq b(\lambda^*), X \in \mathbb{R}^n_+\}) \tag{9}$$

$$(b(\lambda^*), \lambda^*) = \max\{(b(\lambda^*), \Lambda) | A^T\Lambda \geq c(x^*), \Lambda \in \mathbb{R}^m_+\}) \tag{10}$$

In the following section we provide some background, which helps to understand the relations between GE and both LP duality and the Classical Walras–Wald equilibrium, which is equivalent to the J. Nash equilibrium in an n–person concave game.

3 J. Nash equilibrium in an n-person concave game and the Walras -Wald equilibrium.

The notion of equilibrium in a concave n-person game was introduced by J. Nash in 1950(see [9]). In 1994 J. Nash received the Nobel Prize in Economics for his discovery.

For many years it was not clear at all whether the J. Nash equilibrium had anything to do with economic equilibrium introduced as early as in 1874 by Leon Walras in his most admired work "Elements of Pure Economics". Moreover,it was even not clear whether the Walras equations have a solution.

The first substantial contribution was due to Abraham Wald, who in the mid 30s proved the existence of the Walras equilibrium under some special assumptions on the price vector–function $c(x)$. These assumptions, unfortunately, were hard to justify from the economics standpoint (see [5]).

In the mid 50s Harold Kuhn modified the Walras-Wald model and proved under minimal assumptions on the input data the existence of the equilibrium. He used two basic tools: Kakutani's fixed point Theorem (1941) and LP Duality (3)-(4) (see [5] and references therein).

Our interest in J. Nash equilibria was inspired by Ben Rosen's paper [13] in mid 60s, where he discovered some relations between J.Nash equilibria and convex optimization. Soon after we developed (see S. I. Zuchovitsky et. al. [15]–[18]) a few numerical methods for finding a J. Nash equilibrium. Moreover, we proved that H. Kuhn's version of the Walras–Wald equilibrium is equivalent to the J. Nash equilibrium in a concave n–person game [16].

In this section we recall some basic facts and notations related to J. Nash equilibrium, which will be used later.

Let $\Omega_j \subset \mathbb{R}^{m_j}$ be a convex compact set, which defines the feasible strategies $x_j = (x_{j1}, \dots, x_{jm_j})$ for the players $1 \le j \le n$. The Cartesian product

$$\Omega = \Omega_1 \otimes \dots \otimes \Omega_j \otimes \dots \otimes \Omega_n$$

defines the feasible strategies set for a concave n–person game. The set $\Omega \in \mathbb{R}^m, m = \Sigma_{j=1}^n m_j$ is convex and bounded.

The payoff functions $\varphi_j : \Omega \to \mathbb{R}$ of the players $1 \le j \le n$, are continuous and $\varphi_j(x_1, \dots, x_j, \dots x_n)$ is concave in $x_j \in \Omega_j$.

The vector $x^* = (x_1^*, \dots x_j^*, \dots x_n^*) \in \Omega$ defines a J. Nash equilibrium if for every $1 \le j \le n$ the following inequality holds

$$\varphi_j(x_1^*, \dots, x_j, \dots, x_n^*) \le \varphi_i(x_1^*, \dots, x_j^*, \dots, x_n^*), \quad \forall x_j \in \Omega_j \tag{11}$$

which means that for every player $1 \le j \le n$ derivation from the equilibrium strategy can only reduce the corresponding payoff value if the other players accept their equilibrium strategies. Along with J. Nash equilibrium, let's consider the normalized equilibrium for a concave n–person game. Let $x = (x_1, \dots, x_j, \dots x_n) \in \Omega$. On $\Omega \otimes \Omega$ we consider the following normalized payoff function

$$\Phi(x, X) = \sum_{j=1}^{n} \varphi_j(x_1, \dots, X_j, \dots x_n)) \tag{12}$$

Vector $x^* \in \Omega$ is a normalized equilibrium in n-person concave game if

$$\Phi(x^*, x^*) = \max\{\Phi(x^*, X) | X \in \Omega\} \tag{13}$$

It is clear that each normalized equilibrium $x^* = (x_1^*, \dots, x_n^*) \in \Omega$ is a J. Nash equilibrium, but the opposite generally speaking is not true.

By finding $x^* \in \Omega$ from (13) we find the J. Nash equilibrium in a concave n-person game, because it follows from (13) that if any player violates his equilibrium strategy x_j^* and the rest accept their equilibrium strategies, then the player can only reduce his payoff, i.e. (11) holds.

Let's consider the vector–function

$$g(x) = \nabla_x \Phi(x, X)|_{X=x} = (\nabla_{x_1} \varphi_1(X_1, x_2, \dots, x_n), \dots, \nabla_{x_n} \varphi_n(x_1, \dots, x_{n-1}, X_n))|_{X=x} \tag{14}$$

$$= (\nabla_{x_1} \varphi_1(x), \dots, \nabla_{x_n} \varphi_n(x))$$

where

$$\nabla_{x_j}\varphi_j(x_1,\dots X_j,\dots x_n)|_{X_j=x_j} = \left(\frac{\partial\varphi_j(x)}{\partial x_{j1}},\dots,\frac{\partial\varphi_j(x)}{\partial x_{jm_j}}\right),\quad 1\le j\le n \tag{15}$$

The vector function $g(x)$ is called a pseudo-gradient of $\Phi(x,X)$ at $X=x$.

For any given $x\in\Omega$ the normalized payoff function $\Phi(x,X)$ is concave in $X\in\Omega$. Hence,

$$\omega(x)=\operatorname{Arg}\max\{\Phi(x,X)|X\in\Omega\} \tag{16}$$

is the set of solutions of a convex optimization problem. Therefore for any given $x\in\Omega$ the set $\omega(x)\subset\Omega$ is convex and bounded. Moreover, it follows from the continuity of $\Phi(x,X)$ that $\omega(x)$ is an upper semi–continuous set valued function on Ω (see [6]). Therefore the existence of a fixed point $x^*\in\Omega$ of the map $x\to\omega(x)$ i.e. $x^*\in\omega(x^*)$ is a direct consequence of the Shizuo Kakutani Theorem (see, for example, [6]).

If $x^*\in\Omega$ is a given normalized equilibrium, then finding

$$\Phi(x^*,x^*)=\max\{\Phi(x^*,X)|X\in\Omega\} \tag{17}$$

is a convex optimization problem.

The fact that $x^*\in\Omega$ is the solution of convex optimization problem (17) means that for the pseudo-gradient $g(x^*)=\nabla_x\Phi(x^*,X)|X=x^*$ the following inequality

$$(g(x^*),X-x^*)\le 0,\quad \forall X\in\Omega \tag{18}$$

holds.

On the other hand if $x^*\in\Omega$ is a solution of the variational inequality (18) then x^* is a J. Nash equilibrium in the n–person concave game. Therefore finding J. Nash equilibrium in a concave n–person game is equivalent to solving the variational inequality (18).

At the same time one can view (18) as an optimality criteria for the normalized equilibrium.

The monotonicity of the nonlinear operator $g:\Omega\to\mathbb{R}^m$, i.e.

$$(g(x^1)-g(x^2),x^1-x^2)\le 0,\forall(x^1,x^2)\in\Omega\times\Omega \tag{19}$$

is critical for solving the variational inequality (18).

The operator g is strongly monotone on Ω if there is $\gamma>0$ such that

$$(g(x^1)-g(x^2),x^1-x^2)\le -\gamma\|x^1-x^2\|^2\quad \forall(x^1,x^2)\in\Omega\times\Omega \tag{20}$$

We conclude the section by recalling that H. Kuhn's version of the classical Warlas–Wald equilibrium is equivalent to J. Nash equilibrium in n–person concave game (see [16]).

One obtains H. Kuhn's version from (9)- (10) by assuming that the factor vector–function $b(\lambda)$ is fixed, i.e. $b(\lambda)\equiv b$.

In other words a Walras-Wald equilibrium is a pair $(x^*,\lambda^*)\in\mathbb{R}^n_+\otimes\mathbb{R}^m_+$ that

$$(c(x^*), x^*) = \max\{(c(x^*), X)|AX \leq b, X \in \mathbb{R}^n_+\} \tag{21}$$

$$(b, \lambda^*) = \min\{(b, \Lambda)|A^T\Lambda \geq c(x^*)|\Lambda \in \mathbb{R}^m_+\} \tag{22}$$

Let's assume that the price for any good is a non increasing, continuous function of the corresponding output when the production of other goods are fixed, i.e.

$$t_2 > t_1 > 0 \Rightarrow c_j(x_1, \dots, t_2, \dots, x_n) \leq c_j(x_1, \dots, t_1, \dots, x_n), 1 \leq j \leq n.$$

Then the function

$$\varphi_j(x_1, \dots, x_j, \dots, x_n) = \int_0^{x_j} c_j(x_1, \dots, t, \dots, x_n)dt \tag{23}$$

is continuous in $X = (x_1, \dots, x_j \dots, x_n) \in \Omega = \{X : AX \leq b, X \in \mathbb{R}^n_+\}$, and concave in x_j.

Therefore one obtains Walras-Wald equilibrium by finding the normalized J. Nash equilibrium with payoff functions defined by (23), i.e. by finding $x^* \in \Omega$:

$$\Phi(x^*, x^*) = \max\left\{\sum_{j=1}^{n} \varphi_j(x_1^*, \dots, X_j, \dots, x_n^*)|X \in \Omega\right\} \tag{24}$$

Existence of the Walras–Wald equilibrium is a direct consequence of existence of a normalized J. Nash equilibrium for a concave n–person game with payoff functions (23) and feasible strategies set Ω.

In fact, using the optimality criteria (18) and keeping in mind that $g(x^*) = c(x^*)$ we can rewrite the problem (24) as follows

$$\max\{(c(x^*), X)|AX \leq b, X \in \mathbb{R}^m_+\} = (c(x^*), x^*)$$

Therefore using LP Duality one obtains $\lambda^* \in \mathbb{R}^m_+$:

$$(b, \lambda^*) = \min\{(b, \lambda)|A^T\lambda \geq c(x^*), \lambda \in \mathbb{R}^m_+\}$$

i.e. by finding J. Nash equilibrium from (24) one finds the Walras–Wald equilibrium (21)-(22).

Finding a Walras-Wald equilibrium from (24) is equivalent to solving a variational inequality (18). Solving (18), generally speaking, is more difficult then solving the Primal LP (1).

In the following section we consider the GE (9)-(10), which looks more difficult than (21)-(22), however for well defined $c(x)$ and $b(\lambda)$ finding GE turns out to be much easier then solving the LP.

4 Existence and Uniqueness of the GE.

The GE problem was revisited in 1977 by M. Primak and the author in [10]. It was shown that finding a GE is equivalent to solving a particular variational inequality. We will discuss it briefly in this section and use the equivalence later for establishing the existence and uniqueness of the GE as well as for developing the projected pseudo–gradient method for finding GE.

Let's fix $x \in \mathbb{R}^n_+$ and $\lambda \in \mathbb{R}^m_+$ and consider the following dual pair of LP

$$\max\{(c(x), X)|AX \leq b(\lambda), X \in \mathbb{R}^n_+\} \tag{25}$$

$$\min\{(b(\lambda), \Lambda)|A^T\Lambda \geq c(x), \Lambda \in \mathbb{R}^m_+\} \tag{26}$$

Solving the dual LP pair (25)-(26) is equivalent to finding on $\mathbb{R}^n_+ \otimes \mathbb{R}^m_+$ a saddle point for the corresponding Lagrangean

$$\mathbb{L}(x, \lambda; X, \Lambda) = (c(x), X) - (\Lambda, AX - b(\lambda)) \tag{27}$$

i.e. finding

$$\max_{X\in\mathbb{R}^m_+} \min_{\Lambda\in\mathbb{R}^m_+} \mathbb{L}(x, \lambda; X, \Lambda) = \min_{\Lambda\in\mathbb{R}^n_+} \max_{X\in\mathbb{R}^n_+} \mathbb{L}(x, \lambda; X, \Lambda), \tag{28}$$

under fixed $x \in \mathbb{R}^n_+, \lambda \in \mathbb{R}^m_+$.

The problem (28) is in turn equivalent to finding a J. Nash equilibrium of a concave two person game with the following payoff functions

$$\varphi_1(x, \lambda; X, \lambda) = (c(x), X) - (\lambda, AX - b(\lambda)) = (c(x) - A^T\lambda, X) + (\lambda, b(\lambda)) \tag{29}$$

and

$$\varphi_2(x, \lambda; x, \Lambda) = (Ax - b(\lambda), \Lambda) \tag{30}$$

where $X \in \mathbb{R}^n_+$ is the strategy of the first player and $\Lambda \in \mathbb{R}^m_+$ is the strategy of the second player.

Let $y = (x, \lambda), Y = (X, \Lambda) \in \mathbb{R}^n_+ \otimes \mathbb{R}^n_+ = \Omega$. The corresponding normalized payoff function is defined as follows

$$\Phi(y, Y) = (c(x) - A^T\lambda, X) + (Ax - b(\lambda), \Lambda) + (\lambda, b(\lambda)) \tag{31}$$

Therefore finding a saddle point is equivalent to finding a normalized J. Nash equilibrium of two person concave game, i.e. finding such $\bar{y} \in \Omega$ that

$$\Phi(\bar{y}, \bar{y}) = \max\{\Phi(\bar{y}, Y)|Y \in \Omega\} \tag{32}$$

Let us consider the corresponding pseudo–gradient

$$\nabla_Y\Phi(y, Y)|_{Y=y} = g(y) \equiv g(x, \lambda) = (c(x) - A^T\lambda, Ax - b(\lambda)) \tag{33}$$

So finding $\bar{y} \in \Omega$ from (32) is equivalent to solving the following variational inequality

$$(g(\bar{y}), Y - \bar{y}) \leq 0, \quad \forall Y \in \Omega \tag{34}$$

Let us fix $\bar{y} \in \Omega$. It follows from (34) that

$$\max\{(g(\bar{y}), Y - \bar{y}) | Y \in \Omega\} = (g(\bar{y}), \bar{y} - \bar{y}) = 0 \tag{35}$$

From (35) we obtain that $g(\bar{y}) \leq 0$, because by assuming that at least one component of the vector $g(\bar{y})$ is positive, we obtain

$$\max\{(g(\bar{y}), Y - \bar{y}) | Y \in \Omega\} = \infty \tag{36}$$

Therefore if $\bar{y} = (\bar{x}, \bar{\lambda}) \geq 0$ solves the variational inequality (34), then

$$c(\bar{x}) - A^T \bar{\lambda} \leq 0 \quad \text{and} \quad A\bar{x} - b(\bar{\lambda}) \leq 0 \tag{37}$$

Note that solving (35) is equivalent to finding

$$\max\left\{\sum_{j=1}^{n} (c(\bar{x}) - A^T \bar{\lambda}))_j X_j + \sum_{i=1}^{m} (A(\bar{x}) - b(\bar{\lambda}))_i \Lambda_i | X_j \geq 0, j = 1, \ldots, n, \Lambda_i \geq 0, i = 1, \ldots, m\right\}$$

Therefore for $1 \leq j \leq n$ we have

$$(c(\bar{x}) - A^T \bar{\lambda})_j < 0 \Rightarrow \bar{x}_j = 0, \bar{x}_j > 0 \Rightarrow (c(\bar{x}) - A^T \bar{\lambda}) = 0 \tag{38}$$

and for $1 \leq i \leq m$ we have

$$(A\bar{x} - b(\lambda))_i < 0 \Rightarrow \bar{\lambda} = 0, \bar{\lambda}_i > 0 \Rightarrow (A\bar{x} - b(\bar{\lambda}))_i = 0 \tag{39}$$

Hence $\bar{y} = (\bar{x}, \bar{\lambda}) \in \Omega$ is a primal–dual feasible solution, which satisfied the complementarily condition (38)-(39).

Therefore the vector $Y = \bar{y}$ is the solution for the following primal–dual LP

$$\max\{(c(\bar{x}), X) | AX \leq b(\bar{\lambda}), X \in \mathbb{R}^n_+\} = (c(\bar{x}), \bar{x}) \tag{40}$$

$$\min\{(b(\bar{\lambda}), \Lambda) | A^T \Lambda \geq c(\bar{x}), \Lambda \in \mathbb{R}^m_+\} = (b(\bar{\lambda}), \bar{\lambda}) \tag{41}$$

i.e. $\bar{y} = y^*$. On the other hand it is easy to see that GE y^* which is defined by (9)-(10), solves the variational inequality (34). Therefore finding GE y^* is equivalent to solving variational inequality (34).

Now let us show that $y^* \in \Omega$ exists.

The arguments used for proving the existence of a normalized J. Nash equilibrium in section 3 can't be used in case of (32), because the feasible set Ω is unbounded and Kakutani's Theorem can't be applied.

It turns out that if $c(x)$ and $b(\lambda)$ are well defined, then in spite of unboundedness Ω the GE exists. We start with the following technical Lemma.

Lemma 1. If vector–functions $c(x)$ and $b(\lambda)$ are well defined, then the pseudo–gradient $g : \Omega \to \mathbb{R}^{m+n}$ is a strongly monotone operator, i.e. for $\gamma = \min\{\alpha, \beta\} > 0$ the following inequality holds

$$\big(g(y^1) - g(y^2), y^1 - y^2\big) \leq -\gamma \|y^1 - y^2\|^2 \tag{42}$$

for any pair $(y^1; y^2) \in \Omega \otimes \Omega$.

Proof Let $y^1 = (x^1, \lambda^1), y^2 = (x^2, \lambda^2) \in \Omega$ then

$$(g(y^1)-g(y^2), y^1-y^2) = (c(x^1)-A^T\lambda^1-c(x^2)+A^T\lambda^2, x^1-x^2)+(Ax^1-b(\lambda^1)-Ax^2+b(\lambda^2), \lambda^1-\lambda^2)$$

$$= (c(x^1)-c(x^2), x^1-x^2)-(A^T(\lambda^1-\lambda^2), x^1-x^2)+(A(x^1-x^2), \lambda^1-\lambda^2)-(b(\lambda^1)-b(\lambda^2), \lambda^1-\lambda^2)$$

$$= (c(x^1) - c(x^2), x^1 - x^2) - (b(\lambda^1) - b(\lambda^2), \lambda^1 - \lambda^2).$$

Invoking (7) and (8) we obtain (42). We are ready to prove existence and uniqueness of the GE.

Theorem 1. If $c(x)$ and $b(\lambda)$ are well defined, then the GE $y^* = (x^*, \lambda^*)$ exists and is unique.

Proof. Let us consider $y_0 \in \Omega : \|y_0\| \leq 1$ and a large enough number $M > 0$. Instead of (32) we consider the following equilibrium problem.

$$\Phi(y^*_M, y^*_M) = \max\{\Phi(y^*_M, Y) | Y \in \Omega_M\} \tag{43}$$

where $\Omega_M = \{Y \in \Omega : \|Y\| \leq M\}$. The normalized function $\Phi(y, Y)$ defined by (31) is linear in $Y \in \Omega$ and Ω_M is a convex compact set. Therefore for a given $y \in \Omega_M$ the set

$$\omega(y) = \text{Argmax}\{\Phi(y, Y) | Y \in \Omega_M\} \tag{44}$$

is a solution set of a convex optimization problem. Therefore for any given $y \in \Omega_M$ the set $\omega(y) \subset \Omega_M$ is convex and compact. Moreover, the map $y \to \omega(y)$ is upper semi–continuous. In fact, let us consider a sequence $\{y^s\} \subset \Omega_M : y^s \to \bar{y}$ and any sequence of images $\{z^s \in \omega(y^s)\}$ converging to $\bar{z}$.

Due to the continuity of $\Phi(y, Y)$ in y and Y we have $\bar{z} \in \omega(\bar{y})$. Therefore $y \to \omega(y)$ maps convex compact Ω_M into itself and the map is upper semi–continuous, hence Kakutani's Theorem can be applied. Therefore there exists $y^*_M \in \Omega_M : y^*_M \in \omega(y^*_M)$.

Let's show that the constraint $\|Y\| \leq M$ is irrelevant in problem (43). Using the bound (42) for $y^1 = y_0$ and $y^2 = y^*_M$ one obtains

$$\gamma\|y_0 - y^*_M\|^2 \leq (g(y^*_M) - g(y_0), y_0 - y^*_M) = (g(y^*_M), y_0 - y^*_M) + (g(y_0), y^*_M - y_0) \tag{45}$$

Vector y^*_M is the solution of variational inequality (34) when Ω is replaced by Ω_M, hence

$$(g(y^*_M), y_0 - y^*_M) \leq 0 \tag{46}$$

It follows from (45)-(46) and the Cauchy–Schwarz inequality that

$$\gamma\|y_0 - y_M^*\|^2 \leq |(g(y_0), y_M^* - y_0)| \leq \|g(y_0)\|\|y_M^* - y_0\| \tag{47}$$

or

$$\|y_0 - y_M^*\| \leq \gamma^{-1}\|g(y_0)\| \tag{48}$$

Therefore,

$$\|y_M^*\| \leq \|y_0\| + \|y_M^* - y_0\| \leq 1 + \gamma^{-1}\|g(y_0)\|$$

Hence, for $M > 0$ large enough, the inequality $\|Y\| \leq M$ is irrelevant and can be removed from the constraint set in (43).

In other words $y_M^* = y^* = (x^*, \lambda^*)$ is the GE. It turns out that if the vector–functions $c(x)$ and $b(\lambda)$ are well defined, then GE not only exists, but it is unique as well.

In fact, assuming that $\bar{y} \in \Omega, y^* \in \Omega$ are two different solutions of the variational inequality (34) then one obtains $(g(\bar{y}), y^* - \bar{y}) \leq 0$ and $(g(y^*), \bar{y} - y^*) \leq 0$, therefore

$$(g(\bar{y}) - g(y^*), \bar{y} - y^*) \geq 0$$

On the other hand it follows from (42) for $y^1 = \bar{y}$ and $y^2 = y^*$ that

$$(g(\bar{y}) - g(y^*), \bar{y} - y^*) \leq -\gamma\|\bar{y} - y^*\|^2 \tag{49}$$

The contradiction proves uniqueness of the GE y^* .

5 Projected pseudo–gradient method for finding GE.

In this section we introduce the projected pseudo–gradient method for finding GE and show its global convergence with Q–linear rate. We estimate the ratio through the basic parameters of the input data. The global convergence with Q–linear rate allows us to estimate the computational complexity of the method.

For well defined $c(x)$ and $b(\lambda)$ the pseudo–gradient

$$g(y) = (c(x) - A^T\lambda, Ax - b(\lambda)) \tag{50}$$

is strongly monotone, i.e. (42) holds. Therefore finding GE can be reduced to solving the variational inequality (34) for a strongly monotone operator $g : \Omega \to \mathbb{R}^{n+m}$.

We start by recalling basic facts related to the projection operation (see [4],[8]).

Let Q be a closed convex set in $\mathbb{R}^n$. Then for each $x \in \mathbb{R}^n$ there is a unique point $y \in Q$ such that

$$\|x - y\| \leq \|x - z\|, \quad \forall z \in Q \tag{51}$$

The vector y is called the projection of x on Q, i.e.

$$y = P_Q x \tag{52}$$

and $P_Q : \mathbb{R}^n \to Q$ defined by (52) is called the projection operator. Let $x_0 \in Q$, then finding projection $y = P_Q x$ is equivalent to solving the following convex optimization problem:

$$d(y) = \min\{d(z) = \|z - x\|^2 | z \in Q_0\} \tag{53}$$

where $Q_0 = \{z : \|z - x\| \leq \|x_0 - x\|\}$.

The problem (53) has a compact convex feasible set Q_0 and strongly convex and continuous objective function $d(z)$. Therefore the projection $y = P_Q x$ exists and is unique for any given $x \in \mathbb{R}^n$.

Let's consider the optimality criteria for the projection y in (53). Keeping in mind that $\nabla d(z) = 2(z - x)$ and the fact that y is the point in Q_0 closest to the vector x we obtain

$$(y - x, z - y) \geq 0 \quad \text{or} \quad (y, z - y) \geq (x, z - y), \quad \forall z \in Q_0 \tag{54}$$

Now we would like to recall few properties of the projection operator P_Q which will be used later.

Lemma 2. Let Q be a closed convex set. Then the projection operator P_Q is non–expansive, that is

$$\|P_Q x_1 - P_Q x_2\| \leq \|x_1 - x_2\|, \forall x_1, x_2 \in \mathbb{R}^n \tag{55}$$

Proof. Let $y_1 = P_Q x_1$ and $y_2 = P_Q x_2$, then using the optimality criteria (54) for y_1 and y_2 we obtain

$$(y_1, z - y_1) \geq (x_1, z - y_1), \quad \forall z \in Q$$

and

$$(y_2, z - y_2) \geq (x_2, z - y_2), \quad \forall z \in Q.$$

From the first inequality one obtains

$$(y_1, y_2 - y_1) \geq (x_1, y_2 - y_1). \tag{56}$$

From the second we have

$$(y_2, y_1 - y_2) \geq (x_2, y_1 - y_2). \tag{57}$$

Combining (56)and (57) and using the Cauchy–Schwarz inequality we obtain

$$\|y_1 - y_2\|^2 \leq (x_1 - x_2, y_1 - y_2) \leq \|x_1 - x_2\|\|y_1 - y_2\|.$$

Hence

$$\|y_1 - y_2\| \leq \|x_1 - x_2\|.$$

Lemma 3. Let Q be a closed convex set in $\mathbb{R}^n$. Then x^* is a solution of the variational inequality

$$(g(x^*), x - x^*) \leq 0, \quad \forall x \in Q \tag{58}$$

if and only if for any $t > 0$ vector x^* is a fixed point of the map $P_Q(I + tg) : Q \to Q$, i.e.

$$x^* = P_Q(x^* + tg(x^*))$$

Proof. If x^* is the solution of the variational inequality then (58) holds. Multiplying the inequality by $t > 0$ and adding to both sides $(x^*, x - x^*)$ one obtains

$$(x^* + tg(x^*), x - x^*) \le (x^*, x - x^*) \quad \forall x \in Q.$$

It follows from (54) that $x^* = P_Q(x^* + tg(x^*))$. On the other hand if $x^* = P_Q(x^* + tg(x^*))$ then from (54) taking $y = x^*$ and $x = x^* + tg(x^*)$ we obtain $(x^*, x - x^*) \ge (x^* + tg(x^*), x - x^*)$, i.e. $(g(x^*), x - x^*) \le 0 \quad \forall x \in Q$.

We are ready to describe the projected pseudo–gradient method for finding GE.

Let $y^0 = (x^0, \lambda^0) \in \mathbb{R}^n_{++} \otimes \mathbb{R}^m_{++}$ be a starting point. Assume that approximation $y^s = (x^s, \lambda^s)$ has been found already. The next approximation one finds by the formula

$$y^{s+1} = P_\Omega(y^s + tg(y^s)) \tag{59}$$

Keeping in mind that $\Omega = \mathbb{R}^n_+ \otimes \mathbb{R}^m_+$ the method (59) translates into the following formulas for goods and prices updates

$$x_j^{s+1} = \left[x_j^s + t(c(x^s) - A^T\lambda^s)_j\right]_+, \quad j = 1, \dots, n \tag{60}$$

$$\lambda_i^{s+1} = [x_i^s + t(Ax^s - b(\lambda^s))_i]_+, \quad i = 1, \dots, m \tag{61}$$

where $[z]_+ = \begin{cases} z & z \ge 0, \\ 0 & z < 0. \end{cases}$

Before we specify the step length $t > 0$ let's consider the update formulas (60) and (61).

It follows from (60) that if the current price $c_j(x^s)$ exceeds the current expenses $(A^T\lambda^s)_j$ than the production of good $1 \le j \le n$ has to be increased. If the current price $c_i(x^s)$ is less than current expenses $(A^T\lambda^s)_j$ then the production of the good has to be reduced.

It follows from (61) that if the current consumption $(Ax^s)_i$ of factor $1 \le i \le m$ exceeds the current availability $b_i(\lambda^s)$, then one has to increase the price for the factor.

If availability $b_i(\lambda^s)$ of factor $1 \le i \le m$ exceeds consumption $(Ax^s)_i$, then the price should be reduced.

The projection operation keeps the new primal–dual approximation $y^{s+1} = (x^{s+1}, \lambda^{s+1})$ in Ω, i.e. y^{s+1} is non–negative.

We recall that for well defined $c(x)$ and $b(\lambda)$ the pseudo–gradient $g(y)$ is strongly monotone on $\Omega = \mathbb{R}^n_+ \times \mathbb{R}^m_+$, i.e. (42) holds.

Let us also assume that pseudo–gradient $g(y)$ satisfies the Lipschitz condition as well, i.e. there is $L > 0$ such that

$$\|g(y_1) - g(y_2)\| \le L\|y_1 - y_2\| \quad \forall y_1, y_2 \in \Omega \tag{62}$$

Theorem 2. If $c(x)$ and $b(\lambda)$ are well defined and (62) holds then for any given $0 < t < 2\gamma L^{-2}$ the projected pseudo–gradient method (60) globally converges to GE with Q–linear rate and the ratio $0 < q(t) = 1 - 2t\gamma + t^2L^2 < 1$. Also the following bound

$$\|y^{s+1} - y^*\| \le (1 - (\gamma L^{-1})^2)^{1/2}\|y^s - y^*\| \tag{63}$$

holds for $t = \gamma L^{-2} = \min_{t>0} q(t)$

Proof Let us estimate

$$\|y^{s+1} - y^*\|^2 = \|P_\Omega(y^s + tg(y^s)) - P_\Omega(y^* + tg(y^*))\|^2$$

Using Lemma 2 and Lemma 3 we obtain

$$\begin{aligned}\|y^{s+1} - y^*\|^2 &= \|P_\Omega(y^s + tg(y^s)) - P_\Omega(y^* + tg(y^*))\|^2 \\ &\leq \|y^s + tg(y^s) - (y^* + tg(y^*))\|\end{aligned}$$

Then

$$\begin{aligned}\|y^{s+1} - y^*\|^2 \leq \|y^s + tg(y^s) - (y^* + tg(y^*))\|^2 &= \\ (y^s - y^* + t(g(y^s) - g(y^*)), y^s - y^* + t(g(y^s) - g(y^*))) &= \\ \|y^s - y^*\|^2 - 2t(y^s - y^*, g(y^s) - g(y^*)) + t^2\|g(y^s) - g(y^*)\|^2 &\end{aligned}$$

Using the strong monotonicity (42) and Lipschitz condition (63) we obtain

$$\|y^{s+1} - y^*\|^2 \leq \|y^s - y^*\|^2 - 2t\gamma\|y^s - y^*\|^2 + t^2L^2\|y^s - y^*\|^2$$

$$= (1 - 2t\gamma + t^2L^2)\|y^s - y^*\|^2 = q(t)\|y^s - y^*\|^2,$$

Therefore for any $0 < t < 2\gamma L^{-2}$ we obtain $0 < q(t) < 1$.

Hence for any $0 < t < 2\gamma L^{-2}$ the projected pseudo–gradient method globally converges with Q–linear rate and the ratio $0 < q(t) < 1$. Also $q(\gamma L^{-2}) = \min_{t>0} q(t) = 1 - (\gamma L^{-1})^2$, i. e. the bound (63) holds.

The projected pseudo–gradient method (60) converges to GE (x^*, λ^*) :

$$x_j^* > 0 \Rightarrow c_j(x^*) = (A^T\lambda^*);$$

$$x_j^* = 0 \Leftarrow (A^T\lambda^*)_j > c_j(x^*)$$

$$\lambda_i^* > 0 \Rightarrow (Ax^*)_i = b_i(\lambda^*)$$

$$\lambda_i^* = 0 \Leftarrow (Ax^*)_i < b_i(\lambda^*).$$

Therefore if the market “is cleared” from goods for which the production cost exceeds the price and from factors, which exceed the needs for them, then for the remaining goods and factors we obtain the following equilibrium:

$$(A^T\lambda^*)_j = c_j(x^*), \quad (Ax^*)_i = b_i(\lambda^*)$$

and

$$(c(x^*), x^*) = (b(\lambda^*), \lambda^*),$$

which one can view as the Generalized Walras Law.

The pseudo–gradient method (60)-(61) can be viewed as a projected explicit Euler method for the following system of differential equations

$$\frac{dx}{dt} = c(x) - A^T\lambda$$

$$\frac{d\lambda}{dt} = Ax - b(\lambda)$$

If vector–functions $c(x)$ and $b(\lambda)$ are well defined and Lipschitz condition (62) holds then for any given $0 < t < 2\gamma L^{-2}$ the corresponding trajectory $\{y^s = (x^s, \lambda^s)\}_{s=0}^{\infty}$ converges to GE (x^*, λ^*) and the following bound holds

$$\|y^{s+1} - y^*\| \leq q\|y^s - y^*\|, \quad s \geq 1 \tag{64}$$

where $0 < q := (q(t))^{1/2} = (1 - 2t\gamma + t^2L^2)^{1/2} < 1$

Let small enough $0 < \epsilon << 1$ be the required accuracy. Due to (64) it takes $O((\ln \epsilon)(\ln q)^{-1})$ projected pseudo–gradient steps to get an approximation for GE (x^*, λ^*) with accuracy $\epsilon > 0$. The number of operations per step is bounded by $O(n^2)$ (assuming that $n > m$), therefore the total number of operations for finding the GE with given accuracy $\epsilon > 0$ is

$$\mathbb{N} = O(n^2(\ln \epsilon)(\ln q)^{-1})$$

This means that for well defined $c(x)$ and $b(\lambda)$ finding GE might require much less computational effort then solving an LP.

6 Concluding Remarks

Replacing the factor vector b by the factor vector–function $b(\lambda)$ leeds to "symmetrization" of the classical Walras–Wald equilibrium (see[5]). The symmetrization is not only justifiable from the economic standpoint, it also eliminates both the combinational nature of LP as well as the basic difficulties associated with finding classical Walras–Wald equilibrium by solving variational inequality (18) with $g(x) = c(x)$. In fact, instead of solving variational inequality (18) on a polyhedron, we end up solving the variational inequality (34) on $\mathbb{R}^n_+ \otimes \mathbb{R}^m_+$.

What is even more important, the "symmetrization" leads to the projected pseudo–gradient method (59), which is not only attractive numerically, but also represents a natural pricing mechanism for finding economic equilibrium.

All these important features of the projected pseudo–gradient method are due to our assumptions that vector–function $c(x)$ and $b(\lambda)$ are well defined and sufficiently smooth.

If this is not the case then other methods for solving the variational inequality (34) can be used (see [7]).

Lately the interest in gradient projected type methods has been revitalized and a number of important results have been obtained (see [1]–[3] and references therein). Some of the methods developed in [1]–[3] can be used for solving variational inequality (34).

Our main goal however is keeping the algorithms simple and efficient with full understanding of the pricing mechanism they represent.

We hope that the dual Nonlinear Rescaling methods in general (see [12] and reference therein) and the dual MBF method (see [11]) in particular can be used as prototypes for such algorithms. Some new results in this direction will be covered in an upcoming paper.

The research was supported by NSF Grant CCF-0836338.

References

[1] Auslender A., Teboulle M. “Interior projection–like methods for monotone variational inequality,” Math. Program Ser. A 104 (2005).

[2] Auslender A., Teboulle M.“Projected subgradient methods with non–Euclidean distances for non–differentiable convex minimization and variational inequalities” Mathematical Programming Ser B online Publication (2007).

[3] Beck A., Teboulle M., “ A linearly convergent dual based gradient projection Algorithm for Quadratically constrained convex optimization,” Math. of Operations Research, vol 31 n 2 (2006).

[4] Goldstein A. A.,“ Convex programming in Hilbert space,” Bull. Amer. Mat. Soc. 70 (1964).

[5] Kuhn H., “On a theorem of Wald,” Ann. of Math. Studies no 38 Princeton Univ. Press, Princeton NJ (1956).

[6] Lancaster K., “Mathematical Economics,” Dover Publication, NY (1987).

[7] Larsson T., Patriksson M., “ A class of gap functions for variational inequalities,” Mathematical Programming 64 n l (1994).

[8] Levitin E. S., Polyak B. T., “ Minimization methods with constraints” Zh. Vychisl. Matem. I Matem. Fiz. v6, n 5 (1966).

[9] Nash J. “ Non–cooperative games,” Ann. of Math. (2) 54 (1951).

[10] Polyak R., Primak M. “ Methods of controlling sequences for solving equilibrium problems,” Part2, Cibernetics 13 n 4, (1977).

[11] Polyak R.,“ Modified barrier functions(theory and methods),” Mathematical Programming 54, n 2 (1992).

[12] Polyak R. “ Nonlinear rescaling as interior quadratic prox method in convex optimization” Computational Optimization and Applications, 35 (2006).

[13] Rosen J. B. "Existence and uniqueness of equilibrium point for concave n–person game," Econometrica, vol 33, n 3 (1965).

[14] Stiefel E." Note on Jordan elimination, linear programming and Chebyshev approximation," Numerische Mathematic 2 (1960).

[15] Zuchovitsky S. I., Polyak R. A., Primak M. E. " Two methods for finding equilibrium points in N–person concave games," Soviet Math. Dokl., v 10, No, 2 (1969).

[16] Zuchovitsky S. I., Polyak R. A., Primak M. E." N–person concave game and one production model," Soviet Math. Dokl. v. 11, No. 2, (1970).

[17] Zuchovitsky S. I., Polyak R. A., Primak M. E." One general equilibrium problem," Soviet Math. Dokl. v. 14, No. 2, (1973).

[18] Zuchovitsky S. I., Polyak R. A., Primak M. E." Multiperson concave games," numerical methods Matekon 9 (1973).