Nonlinear Equilibrium vs. Linear Programming.

Roman A. Polyak

Department of Mathematics
The Technion - Israel Institute of Technology
32000 Haifa, Israel
rpolyak@techunix.technion.ac.il

May, 2015

"Linear programming has been one of the most important postwar developments in economic theory."
Robert Dorfman, Paul Samuelson, Robert Solow

Paul Samuelson
Nobel Prize in Economics 1970;

Robert Solow
Nobel Prize in Economics 1987.

In 1975 L.V. Kantorovich and T.C. Koopmans
Shared the Nobel Prize in Economics
“for their contribution to the theory of optimal allocation of limited resources.”

T.Koopmans, G. Dantzig, L.Kantorovich (1970)

Kantorovich L.V. 1939, George Dantzig 1947:

$$(c, x^*) = \max\{(c, x) | Ax \leq b, x \geq 0\}$$
$$A : \mathbb{R}^n \to \mathbb{R}^m, b \in \mathbb{R}^m, c \in \mathbb{R}^n$$

John von Neumann 1947:

$$(b, \lambda^*) \quad = \min\{(b, \lambda) | A^T \lambda \geq c, \lambda \geq 0\}$$

1. $(c, x^*) = (b, \lambda^*)$
2. $(A^T \lambda^* - c, x^*) = 0, (b - Ax^*, \lambda^*) = 0$

Linear Equilibrium

LP assumptions

The LP approach uses two fundamental assumptions:
a) The price vector $c = (c_1, \ldots, c_n)^T$ for goods is fixed, given a priori and independent of the production output vector $x = (x_1, \ldots, x_n)^T$.
b) The resource vector $b = (b_1, \ldots, b_m)^T$ is also fixed, given a priori and the resource availability is independent of the resource price vector
$\lambda = (\lambda_1, \ldots, \lambda_m)^T$

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. Also, a small change of at least one component of the price vector c might lead to a drastic change in the primal solution. Similarly, a small variation of the resource vector b might lead to a dramatic change in the dual solution.

Example

$x_1 + x_2 \to \max$

$$\Omega = \begin{cases} (1-\varepsilon)x_1 + x_2 & \leq 1 \\ x_1 + (1-\varepsilon)x_2 & \leq 1 \\ x_1 \geq 0, x_2 \geq 0 & \end{cases}$$

$$x_1^* = \frac{1}{2-\varepsilon}, x_2^* = \frac{1}{2-\varepsilon}$$

$$(1+\varepsilon)x_1 + x_2 \to \max_{x \in \Omega} \Rightarrow x_1^* = 1, x_2^* = 0$$

$$x_1 + (1+\varepsilon)x_2 \to \max_{x \in \Omega} \Rightarrow x_1^* = 0, x_2^* = 1$$

Nonlinear Equilibrium

$$c \Rightarrow c(x) = (c_1(x), \ldots, c_n(x)) \in \mathbb{R}^n_+$$
$$b \Rightarrow b(\lambda) = (b_1(\lambda), \ldots, b_m(\lambda)) \in \mathbb{R}^m_+$$

$$x^* \in \mathbb{R}^n_+, \lambda^* \in \mathbb{R}^m_+ :$$

$$(c(x^*), x^*) = \max\{(c(x^*), X) | AX \leq b(\lambda^*), X \geq 0\}$$
$$(b(\lambda^*), \lambda^*) = \min\{(b(\lambda^*), \Lambda) | A^T\Lambda \geq c(x^*), \Lambda \geq 0\}$$

NE Assumptions

1. The matrix A has no zero rows or columns
2. The price vector-function $c : \mathbb{R}^n_+ \to \mathbb{R}^n_+$ for goods is continuous and strongly monotone decreasing, i.e. there is $\alpha > 0$ such that

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

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

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

The price $c(x)$ and the factor $b(\lambda)$ are said to be well defined if 2 and 3 hold.

Existence and Uniqueness of NE

Theorem:
If $c(x)$ and $b(\lambda)$ are well defined then NE $(x^*, \lambda^*) \in \mathbb{R}^n_+ \otimes \mathbb{R}^m_+$ exists and is unique.

1. $(c(x^*), x^*) = (b(\lambda^*), \lambda^*)$
2. $(A^T\lambda^* - c(x^*), x^*) = 0, (b(\lambda^*) - Ax^*, \lambda^*) = 0$

Nonlinear Equilibrium

Walras-Wald Equilibrium

$$b(\lambda) \equiv b, \Omega = \{x \in \mathbb{R}^n_+ : Ax \leq b\}$$
$$x^* \in \Omega : (c(x^*), x^*) = \max\{(c(x^*), x) : x \in \Omega\}$$

If such $x^* \in \Omega$ exists, then there is a solution for the dual problem

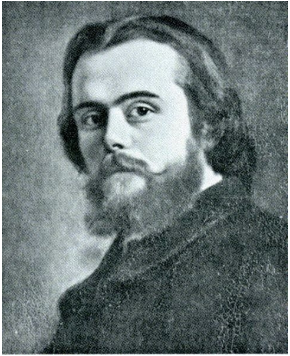
$$(b, \lambda^*) = \min\{(b, \lambda) : A^T\lambda \geq c(x^*), \lambda \in \mathbb{R}^m_+\}$$

and

1. $(c(x^*), x^*) = (b, \lambda^*)$
2. $(b - Ax^*, \lambda^*) = 0$
3. $(A^T\lambda^* - c(x^*), x^*) = 0$

L. Walras (1874), A. Wald (1935-1936), H. Kuhn (1956)

Marie-Esprit-Léon Walras (1834-1910)



Leon Walras has been hailed by Dr. William Jaffe (1898-1980), the leading historian of economic thought, as one of the "greatest of all economists".
Walras's most admired work "Elements of Pure Economics" was published in 1874.

Walras set forth the “Neoclassical Theory in a formal general equilibrium setting and is currently considered the Father of General Equilibrium Theory”.

In the aftermath of the “Elements . . . ”, Walras built up a correspondence with virtually every important economist of the time from America to Russia, but for the most part he was largely ignored or dismissed by both economic and mathematical mainstream.

Walras was appointed to Lausanne in 1870 and in 1893 was succeeded by his young admirer Vilfredo Pareto. The two men formed the core of the Lausanne School.

Walras’s “Elements . . . ” encompasses much of what is available to us in modern General Equilibrium Theory.

John F. Nash equilibrium in n-person concave game

Proc. of the 36th National Academy of Sciences Meeting, USA, 1950
Ω_i is a convex and compact set in $\mathbb{R}^{m_i}$ and $x_i \in \Omega_i$ is a strategy of the ith player, $i = 1, \ldots, n$.
Each player is associated with a payoff function

$$\varphi_i(x) = \varphi_i(x_1, \ldots, x_i, \ldots, x_n),\ i = 1, \ldots, n$$

which is defined on the Cartesian product

$$\Omega = \Omega_1 \otimes \cdots \otimes \Omega_i \otimes \cdots \otimes \Omega_n$$

The set

$$\Omega \subset \mathbb{R}^{m_1} \otimes \cdots \otimes \mathbb{R}^{m_n}$$

is convex and bounded.

Nash equilibrium in n-person concave game

The payoff function $\varphi_i(x_1,\dots,x_i,\dots,x_n)$ is continuous in x and concave in $x_i \in \Omega_i$.
The vector $x^* \in \Omega$ is called **Nash equilibrium in n-person concave game** if

$$\varphi_i(x_1^*,\dots,x_i^*,\dots,x_n^*) \geq \varphi_i(x_1^*,\dots,x_i,\dots,x_n^*),\ \forall x_i \in \Omega_i \\ i = 1,\dots,n$$

Consider $\Phi(x,y) = \sum_{i=1}^n \varphi_i(x_1,\dots,y_i,\dots,x_n)$.
$x^* \in \Omega$ is a normalized equilibrium if

$$\begin{aligned}\Phi(x^*,x^*) &= \max\{\Phi(x^*,y)|y \in \Omega\} \\ &= \max\left\{\sum_{i=1}^n \varphi_i(x_1^*,\dots,y_i,\dots,x_n^*)\middle| y=(y_1,\dots,y_n)\in\Omega\right\}\end{aligned}$$

Nobel Prize in Economics 1994

J. Rozen (1965); S. Zuchovitsky, R.Polyak, M.Primak (1968)

Any normalized equilibrium is a J. Nash equilbrium

Let

$$\omega(x) = \mathrm{Argmax}\{\Phi(x, y) | y \in \Omega\}$$

The existence of Nash equilibrium in an n-person concave game follows from the existence of the "fixed point" of ω, that is $x^* \in \omega(x^*)$.
The existence of the fixed point follows from the **Kakutani Theorem** (1941), because $\omega(x)$ is bounded, convex and upper semicontinuous.

Pseudo-gradient

$$\begin{aligned} g(x) &= \nabla_y \Phi(x, y)_{y=x} \\ = (\nabla_{y_1} \varphi_1(y_1, x_2, \ldots, x_n), \ldots, &\nabla_{y_i} \varphi_i(x_1, \ldots, y_i, \ldots, x_n), \\ &\ldots, \nabla_{y_n} \varphi_n(x_1, \ldots, x_{n-1}, y_n))_{y=x} \end{aligned}$$

Pseudo-Hessian

$$H(x) = J(g(x))$$

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

Under fixed $x = x^*$ we have a convex programming problem. Consider the pseudo-gradient at x^*

$$\nabla_y \Phi(x^*, y)_{y=x^*} = g(x^*)$$

Then

$$\max\{(g(x^*), y - x^*) | y \in \Omega\} \\ = (g(x^*), x^* - x^*) = 0$$

or x^* is a solution of the VI

$$(g(x^*), x - x^*) \le 0, \quad \forall x \in \Omega$$

Walras-Wald as J. Nash Equilibrium in n-person concave game

$$\varphi_i(x) = \varphi_i(x_1, \dots, x_i, \dots, x_n) = \int_0^{x_i} c_i(x_1, \dots, t, \dots, x_n)dt$$

We assume that for any $0 < t_1 < t_2 < \infty$ and any $x \in \Omega$ the following **price reduction condition** holds

$$c_i(x_1, \dots, t_1, \dots, x_n) \geq c_i(x_1, \dots, t_2, \dots, x_n)$$

Then $\varphi_i(x_1, \dots, x_i, \dots, x_n)$ is concave in x_i.

Everything but intellect is subordinate to the decreasing efficiency law.

J. Clark

The normalized equilibrium

$$x^* \in \operatorname{Argmax}\left\{ \sum_{i=1}^{n} \varphi_i(x_1^*, \dots, y_i, \dots, x_n^*) \middle| y = (y_1, \dots, y_i, \dots, y_n) \in \Omega \right\}$$

or

$$x^* \in \operatorname{Argmax}\{\Phi(x^*, y) | y \in \Omega\},$$

where

$$\begin{aligned} \Phi(x, y) &= \sum_{i=1}^{n} \varphi_i(x_1, \dots, y_i, \dots, x_n) \\ &= \sum_{i=1}^{n} \int_0^{y_i} c_i(x_1, \dots, t, \dots, x_n) dt \end{aligned}$$

The pseudo-gradient

$$\begin{aligned} g(x) = \nabla_y \Phi(x, y)_{y=x} \\ = (\nabla_{y_1} \varphi_1(y_1, x_2, \dots, x_n), \dots, \nabla_{y_i} \varphi_i(x_1, \dots, y_i, \dots, x_n), \\ \dots, \nabla_{y_n} \varphi_n(x_1, \dots, x_{n-1}, y_n))_{y=x} \\ = c(x) \end{aligned}$$

Walras-Wald equilibrium = J. Nash equilibrium

The optimality criteria

$$\max\{(c(x^*), x - x^*) | x \in \Omega\}$$
$$= (c(x^*), x^* - x^*) = 0$$

or

$$(c(x^*), x^*) = \max\{(c(x^*), x) | x \in \Omega\}$$

Finding the Walras-Wald equilibrium is equivalent to finding

the normalized J. Nash equilibrium in an n-person concave game.

S. Zuchovitsky, R.Polyak, M.Primak (1970, 1973)

A. Bakushinski, B. Polyak (1974), ...

Finding Nonlinear Equilibrium

$$(c(x^*), x^*) = \max\{(c(x^*), X) | AX \leq b(\lambda^*), X \geq 0\}$$
$$(b(\lambda^*), \lambda^*) = \min\{(b(\lambda^*), \Lambda) | A^T\Lambda \geq c(x^*), \Lambda \geq 0\}$$
$$y = (x, \lambda), Y = (X, \Lambda)$$

$$L(x,\lambda;X,\Lambda) = (c(x),X) - (AX - b(\lambda),\Lambda)$$

$$y^* = (x^*;\lambda^*) \in \mathbb{R}^n_+ \otimes \mathbb{R}^m_+$$

$$L(y^*;X,\Lambda) = (c(x^*),X) - (AX - b(\lambda^*),\Lambda)$$

$$y^* = (x^*;\lambda^*) : L(y^*;X,\lambda^*) \le L(y^*;x^*,\lambda^*) \le L(y^*;x^*,\Lambda),$$
$$\forall X \in \mathbb{R}^n_+,\ \forall \Lambda \in \mathbb{R}^m_+$$

$$
\begin{aligned}
\varphi_1(x,\lambda;X,\lambda) &= (c(x),X) - (AX - b(\lambda),\lambda) \\
&= (c(x) - A^T\lambda, X) + (b(\lambda),\lambda) \\
\varphi_2(x,\lambda;x,\Lambda) &= (Ax - b(\lambda),\Lambda) \\
\Phi(y;Y) &= (c(x) - A^T\lambda, X) + (\Lambda, Ax - b(\lambda)) + (\lambda, b(\lambda))
\end{aligned}
$$

$$y^* = (x^*; \lambda^*) :$$

$$\Phi(y^*; y^*) = \max\{\Phi(y^*; Y) | Y \in \mathbb{R}_+^{n+m}\}$$

$$\nabla_Y \Phi(y, Y)_{Y=y} = g(y)$$

$$\mathbf{g}(\mathbf{y}) = \mathbf{g}(\mathbf{x}; \lambda) = (\mathbf{c}(\mathbf{x}) - \mathbf{A}^{\mathbf{T}}\lambda; \mathbf{A}\mathbf{x} - \mathbf{b}(\lambda))$$

$$y^* \in \Omega = \mathbb{R}_+^n \otimes \mathbb{R}_+^m$$

$$(\mathbf{g}(\mathbf{y}^*), \mathbf{y} - \mathbf{y}^*) \leq \mathbf{0}, \forall \mathbf{y} \in \mathbb{R}_+^{\mathbf{n}} \otimes \mathbb{R}_+^{\mathbf{m}}$$

Pseudo-gradient strong monotonicity

$$
\begin{aligned}
&(g(y_1) - g(y_2), y_1 - y_2) \\
&= (c(x_1) - c(x_2), x_1 - x_2) - (A^T(\lambda_1 - \lambda_2), x_1 - x_2) \\
&\quad + (A(x_1 - x_2), \lambda_1 - \lambda_2) - (b(\lambda_1) - b(\lambda_2), \lambda_1 - \lambda_2) \\
&= (c(x_1) - c(x_2), x_1 - x_2) - (b(\lambda_1) - b(\lambda_2), \lambda_1 - \lambda_2) \\
&\leq -\nu\|x_1 - x_2\|^2 - \rho\|\lambda_1 - \lambda_2\|^2 \\
&\leq -\gamma\|y_1 - y_2\|^2
\end{aligned}
$$

where $\gamma = \min\{\nu, \rho\}$.
$\gamma = 0 \Rightarrow$ monotonicity

Lemma 1

Local strong monotonicity

$$(b(\lambda) - b(\lambda^*), \lambda - \lambda^*) \geq \beta \|\lambda - \lambda^*\|^2, \beta > 0, \forall \lambda \in \mathbb{R}_+^m$$
$$(c(x) - c(x^*), x - x^*) \leq -\alpha \|x - x^*\|^2, \alpha > 0, \forall x \in \mathbb{R}_+^n$$

If the operators b and c are locally strongly monotone, then the operator g is locally strongly monotone, i.e. for $\gamma = \min\{\nu, \rho\}$

$$(g(y) - g(y^*), y - y^*) \leq -\gamma \|y - y^*\|^2, \; \forall y \in \Omega$$

Local Lipschitz condition

$$\|b(\lambda) - b(\lambda^*)\| \leq L_b\|\lambda - \lambda^*\|, \; \forall \lambda \in \mathbb{R}^m_+$$
$$\|c(x) - c(x^*)\| \leq L_c\|x - x^*\|, \; \forall x \in \mathbb{R}^n_+$$

If b and c satisfy the local Lipschitz condition then the operator g satisfies the local Lipschitz condition, i.e. there is an $L > 0$ such that

$$\|g(y) - g(y^*)\| \leq L\|y - y^*\|, \; \forall y \in \Omega$$

Let $\varkappa = \gamma L^{-1}$ be the condition number of g.

Projection Properties

$$P_Q(u) = \operatorname{argmin}\{\|w-u\| | w \in Q\}$$

is the projection of u onto the set $Q \subset \mathbb{R}^n$. P_Q is non-expansive:

$$\|P_Q(u_1) - P_Q(u_2)\| \leq \|u_1 - u_2\|, \ \forall u_1, u_2 \in \mathbb{R}^n$$

A vector u^* is a solution to

$$(g(u^*), u - u^*) \leq 0, \forall u \in Q$$

if and only if

$$u^* = P_Q(u^* + tg(u^*)) \ , \forall t > 0$$

Projection on Ω

For a vector $u \in \mathbb{R}^q$, the projection on $\mathbb{R}^q_+$ is given by the formula

$$v = P_{\mathbb{R}^q_+}(u) = [u]_+ = ([u_1]_+, \ldots, [u_q]_+)^T$$

where

$$[u_i]_+ = \begin{cases} u_i, & u_i \geq 0 \\ 0, & u_i < 0 \end{cases}$$

Therefore the projection $P_\Omega(y)$ of $y = (x, \lambda) \in \mathbb{R}^n \otimes \mathbb{R}^m$ on $\Omega = \mathbb{R}^n_+ \otimes \mathbb{R}^m_+$ is defined by

$$P_\Omega(y) = [y]_+ = ([x]_+; [\lambda]_+)$$

Projected Pseudo-Gradient Method

$$g(y) = (c(x) - A^T\lambda; Ax - b(\lambda))$$

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

$$\begin{aligned} x_j^{s+1} &= [x_j^s + t(c(x^s) - A^T\lambda^s)_j]_+, \; j = 1, \dots, n \\ \lambda_i^{s+1} &= [\lambda_i^s + t(Ax^s - b(\lambda^s))_i]_+, \; i = 1, \dots, m \end{aligned}$$

A.Golstein (1964); E.Levitin,B.Polyak (1966).

PPM is equivalent to a projected Explicid Euler method for the linear ODE

$$\begin{aligned} \frac{dx}{dt} &= c(x) - A^T\lambda \\ \frac{d\lambda}{dt} &= Ax - b(\lambda) \end{aligned}$$

Theorem 2

If the operators b and c are well defined at the equilibrium and the local Lipschitz condition holds, then

1. for any $0 < t < 2\gamma L^{-2}$ the PPG method globally converges to NE $y^* = (x^*, \lambda^*)$ with Q-linear rate and the ratio $0 < q(t) = (1 - 2t\gamma + t^2 L^2)^{1/2} < 1$, i.e.

$$\|y^{s+1} - y^*\| \leq q(t)\|y^s - y^*\|$$

2. for $t = \gamma L^{-2} = \min\{q(t) | t > 0\}$ the following bound holds

$$\|y^{s+1} - y^*\| \leq (1 - \varkappa^2)^{1/2} \|y^s - y^*\|;$$

3. for the PPG complexity we have

$$Comp(PPG) = O(n^2 \varkappa^{-2} \ln \varepsilon^{-1})$$

where $\varepsilon > 0$ is the required accuracy.

Extra Pseudo-Gradient for finding NE

If $\gamma = 0$ we use the extra pseudo-gradient method for finding NE.
Two phase method.
Predictor:

$$\hat{y}_s = P_\Omega(y_s + tg(y_s)) = [y_s + tg(y_s)]_+$$

Corrector:

$$y_{s+1} = P_\Omega(y_s + tg(\hat{y}_s)) = [y_s + tg(\hat{y}_s)]_+$$

G. Korpelevich (1976)

A.Antipin (2002, 2003);Y.Censor, A.Gibali ,S Reich (2012); .A.Jusem, B.Svaiter (1997)...

Galya Korpelevich
With her son Misha currently Professor at the Technion

Predicted production:

$$\hat{x}_s = [x_s + t(c(x_s) - A^T \lambda_s)]_+$$

Predicted price:

$$\hat{\lambda}_s = [\lambda_s + t(Ax_s - b(\lambda_s))]_+$$

Corrected production:

$$x_{s+1} = [x_s + t(c(\hat{x}_s) - A^T \hat{\lambda}_s)]_+$$

Corrected price:

$$\lambda_{s+1} = [\lambda_s + t(A\hat{x}_s - b(\hat{\lambda}_s))]_+$$

Theorem 3

If c and b are monotone operators and the Lipschitz condition is satisfied, then for any $t \in (0, (\sqrt{2}L)^{-1})$ the EPG method generates a sequence $\{y_s\}_{s=1}^{\infty}$ that converges to NE, *i.e.* $\lim_{s\to\infty} y_s = y^*$.

Remark

$$\sup_{s\geq 1} \|y_{s+1} - y^*\| \left(\|y_s - y^*\|\right)^{-1} = q \leq 1$$

$$q < 1?$$

Theorem 4

If local strong monotonicity for b and c and the Lipschitz condition for g are satisfied then for $\nu(t) = 1 + 2\gamma t - 2(tL)^2$ and for the ratio $q(t) = 1 - 2\gamma t + 4(\gamma t)^2(\nu(t))^{-1}$ the following bound holds

1) $\|y_{s+1} - y^*\|^2 \le q(t)\|y_s - y^*\|^2$; $0 < q(t) < 1, \forall t \in (0, \sqrt{2}L)$
2) for $t = \frac{1}{2L}$ and $\varkappa = \gamma L^{-1}$ we have

$$q\left(\frac{1}{2L}\right) \le \frac{1+\varkappa}{1+2\varkappa};$$

3) for any $\varkappa \in [0, 0.5]$ the following bound holds

$$\|y_{s+1} - y^*\| \le \sqrt{1 - 0.5\varkappa}\|y_s - y^*\|;$$

4) $Comp(EPG) = O(n^2\varkappa^{-1} \ln \varepsilon^{-1})$.

Upper bound for L

If

a)

$$\|A\|_I = \max_{1 \le j \le n} \sum_{i=1}^{m} |a_{ij}| \le 1, \ \|A\|_{II} = \max_{1 \le i \le m} \sum_{i=1}^{n} |a_{ij}| \le 1$$

b)

$$|c_j(x_1) - c_j(x_2)| \le L_{c,j} \|x_1 - x_2\|$$

c)

$$|b_i(\lambda_1) - b_i(\lambda_2)| \le L_{b,i} \|\lambda_1 - \lambda_2\|$$

Then $L = O(\sqrt{n})$

Complexity of the EPG method

When $\gamma > 0$, the normalization $\|A\|_I \leq 1$, $\|A\|_{II} \leq 1$ forces:

$$\gamma \to \gamma n^{-1} = O(n^{-1})$$

$$Comp(EPG) = O(n^2 \varkappa^{-1} \ln \varepsilon^{-1})$$

$$\varkappa = \frac{\gamma}{L} = O(n^{-1.5})$$

$$Comp(EPG) = O(n^{3.5} \ln \varepsilon^{-1})$$

Example

$$c(x) = Cx + c$$

$$b(\lambda) = B\lambda + b$$

$$\Downarrow$$

$$\|A\|_1 \leq 1 \Rightarrow \|C\|_1 \leq 1$$

$$\|A\|_2 \leq 1 \Rightarrow \|B\|_2 \leq 1$$

$$\varkappa = \frac{\gamma}{\sqrt{n}}$$

$$Comp(EPG) = O(n^{2.5} ln \varepsilon^{-1})$$

Generalized Walras-Wald Law

$$\begin{aligned}
\lambda_i^* > 0 \quad &\Rightarrow \quad (Ax^*)_i = b_i(\lambda^*) \\
\lambda_i^* = 0 \quad &\Leftarrow \quad (Ax^*)_i < b_i(\lambda^*) \\
x_j^* > 0 \quad &\Rightarrow \quad (A^T\lambda^*)_j = c_j(x^*) \\
x_j^* = 0 \quad &\Leftarrow \quad (A^T\lambda^*)_j > c_j(x^*)
\end{aligned}$$

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

$$(A^T\lambda^* - c(x^*), x^*) = 0, \ (b(\lambda^*) - Ax^*, \lambda^*) = 0$$