# INTRODUCTION TO GENERALIZED NASH EQUILIBRIUM PROBLEMS

Vladimir Shikhman

Chemnitz University of Technology, Germany

Lecture 1

# GENERALIZED NASH EQUILIBRIUM PROBLEM

Players $\{1, \dots, N\}$

$\boxed{\text{PARAMETRIC OPTIMIZATION of Player } \nu\text{:}}$

$$\mathrm{P}_\nu(x^{-\nu}): \quad \min_{x^\nu} f^\nu(x^\nu, x^{-\nu}) \quad \text{s.t.} \quad x^\nu \in M^\nu(x^{-\nu}).$$

$\boxed{\text{FEASIBLE SET of Player } \nu\text{:}}$

$$M^\nu(x^{-\nu}) := \left\{ x^\nu \in \mathbb{R}^{n_\nu} \;\middle|\; \begin{array}{ll} g_j^\nu(x^\nu, x^{-\nu}) \geq 0, & j \in J^\nu, \\ G_j(x^\nu, x^{-\nu}) \geq 0, & j \in \mathcal{J} \end{array} \right\}$$

$f^\nu$ - **OBJECTIVE FUNCTION**
$g_j^\nu$ - **INDIVIDUAL CONSTRAINTS**
$G_j$ - **SHARED CONSTRAINTS**

# IMPORTANT FEATURES of GNEP

- Not only the objective function, but also the individual constraints depend on other players' variables. Hence, the feasible sets are not fixed as it is the case in the classical game theory, but may rather vary.

- The shared constraints explicitly relate the players' behaviour. Their appearance is usually due to the use of common resources, such as e.g. communication link, transportation facilities, or by facing common limitations, such as e.g. total pollution in a certain area, fishery quotas etc.

**GENERALIZED** NASH EQUILIBRIUM PROBLEM

# QUESTIONS

$$\boxed{\begin{array}{c} \left(\bar{x}^1, \ldots, \bar{x}^N\right) \text{ is GENERALIZED NASH EQUILIBRIUM} \\ \text{iff} \\ \bar{x}^\nu \text{ SOLVES } \mathrm{P}_\nu\left(\bar{x}^{-\nu}\right) \text{ for all } \nu = 1, \ldots, N \end{array}}$$

How is it possible:

- to describe the local structure of the set of generalized Nash equilibria, maybe, as a nonsmooth manifold with boundary,
- to compute generalized Nash equilibria by numerical schemes with a guaranteed convergence rate?

# OUTLINE

1. PRELIMINARIES

   Optimality conditions and singularities

2. STRUCTURE OF GNE-SET

   Nonsmooth analysis and genericity

3. MARKET EQUILIBRIA

   Convex duality and subgradient schemes

# NONLINEAR PROGRAMMING

$$\text{P}: \quad \min_{x\in\mathbb{R}^n} f(x) \quad \text{s.t.} \quad g_j(x) \geq 0, i \in J,$$

where all functions $f, g_j, j \in J$ are twice continuously differentiable.

- Fritz-John points
- Mangasarian-Fromovitz Constraint Qualification
- Linear Independence Constraint Qualification
- Karush-Kuhn-Tucker points
- Strict Complementarity
- Second-Order Sufficient Condition

$$\boxed{\text{NONDEGENERACY vs. SINGULARITIES}}$$

# FRITZ-JOHN POINTS

A feasible point $\bar{x}$ is called Fritz-John if there exist Lagrange multipliers $\bar{\delta}, \bar{\lambda}_j, j \in J$ – not all vanishing – such that it holds:

$$\begin{cases} \bar{\delta} \cdot Df(\bar{x}) & = & \displaystyle\sum_{j \in J} \bar{\lambda}_j \cdot Dg_j(\bar{x}) \\ \bar{\lambda}_j \cdot g_j(\bar{x}) & = & 0, \quad \bar{\lambda}_j \geq 0, \quad g_j(\bar{x}) \geq 0 \quad \text{for all } j \in J, \\ \bar{\delta} & \geq & 0. \end{cases}$$

**Theorem (First-order necessary optimality condition)**

*Let $\bar{x}$ solve (P). Then, it is a Fritz-John point.*

# COMPLEMENTARITY

Let a feasible point $\bar{x}$ be given. Then,

$$\bar{\lambda}_j \cdot g_j(\bar{x}) = 0, \quad \bar{\lambda}_j \geq 0, \quad g_j(\bar{x}) \geq 0 \quad \text{for all } j \in J$$

equivalently means

$$\bar{\lambda}_j \geq 0 \quad \text{for all } j \in J_0(\bar{x}), \quad \bar{\lambda}_j = 0 \quad \text{for all } j \notin J_0(\bar{x}),$$

where the index set of active inequality constraints is

$$J_0(\bar{x}) = \{ j \in J \mid g_j(\bar{x}) = 0 \}.$$

| |
|---|
| LAGRANGE MULTIPLIERS FOR INACTIVE INEQUALITY CONSTRAINTS VANISH |

# FRITZ-JOHN POINTS REVISITED

A feasible point $\bar{x}$ is Fritz-John if there exist Lagrange multipliers $\bar{\delta}, \bar{\lambda}_j, j \in J_0(\bar{x})$ – not all vanishing – such that it holds:

$$\left\{ \begin{array}{rcl} \bar{\delta} \cdot Df(\bar{x}) & = & \sum\limits_{j \in J_0(\bar{x})} \bar{\lambda}_j \cdot Dg_j(\bar{x}) \\ \bar{\lambda}_j & \geq & 0 \quad \text{for all } j \in J_0(\bar{x}), \\ \bar{\delta} & \geq & 0. \end{array} \right.$$

IF $\bar{\delta} = 0$, OBJECTIVE FUNCTION
IS NOT PRESENT IN OPTIMILATY CONDITIONS

# CONSTRAINT QUALIFICATIONS

- Mangasarian-Fromovitz Constraint Qualification (MFCQ) is said to hold at a feasible point $\bar{x}$ if there exist a vector $\xi \in \mathbb{R}^n$, such that it holds:

$$Dg_j(\bar{x}) \cdot \xi > 0 \quad \text{for all } j \in J_0(\bar{x}).$$

- Linear Independence Constraint Qualification (LICQ) is said to hold at a feasible point $\bar{x}$ if the following vectors are linearly independent:

$$Dg_j(\bar{x}), j \in J_0(\bar{x}).$$

**Lemma**

- *LICQ implies MFCQ.*
- *Under MFCQ, $\bar{\delta} \neq 0$ at a Fritz-John point.*

# KARUSH-KUHN-TUCKER POINTS

DIVIDE FRITZ-JOHN CONDITIONS BY $\bar{\delta} \neq 0$

A feasible point $\bar{x}$ is called Karush-Kuhn-Tucker if there exist Lagrange multipliers $\bar{\lambda}_j$, $j \in J_0(\bar{x})$, such that it holds:

$$\begin{cases} Df(\bar{x}) & = & \sum\limits_{j \in J_0(\bar{x})} \bar{\lambda}_j \cdot Dg_j(\bar{x}) \\ \bar{\lambda}_j & \geq & 0 \quad \text{for all } j \in J_0(\bar{x}). \end{cases}$$

**Theorem (First-order necessary optimality condition under LICQ)**

*Let $\bar{x}$ solve (P) and fulfil LICQ. Then, it is a Karush-Kuhn-Tucker point, moreover, the corresponding Lagrange multipliers $\bar{\lambda}_j$, $j \in J_0(\bar{x})$ are uniquely determined.*

# NECESSARY OPTIMALITY CONDITION

Let a KKT-point $\bar{x}$ satisfying LICQ be given.

- The Lagrange function is

$$L(x) = f(x) - \sum_{j \in J_0(\bar{x})} \bar{\lambda}_j \cdot g_j(x),$$

  where $\bar{\lambda}_j$, $j \in J_0(\bar{x})$ are the corresonding Lagrange multipliers.
- The tangent space is

$$T_{\bar{x}}M = \left\{ \xi \in \mathbb{R}^n \mid Dg_j(\bar{x}) \cdot \xi = 0, j \in J_0(\bar{x}) \right\}.$$

**Theorem (Second-order necessary optimality condition)**

*Let $\bar{x}$ solve (P) and fulfil LICQ. Then, the Hessian $D^2L(\bar{x})$ of the Lagrange function restricted to the tangent space $T_{\bar{x}}M$ is positive semi-definite.*

# SUFFICIENT OPTIMALITY CONDITION

Let a KKT-point $\bar{x}$ satisfying LICQ be given.

- Strict Complementarity (SC) is said to hold if the uniquely determined Lagrange multipliers of $\bar{x}$ are positive, i. e.

$$\bar{\lambda}_j > 0 \quad \text{for all} \quad j \in J_0(\bar{x}),$$

- Second-Order Sufficient Condition (SOSC) is said to hold if the Hessian $D^2 L(\bar{x})$ of the Lagrange function restricted to the tangent space $T_{\bar{x}} M$ is positive definite, i. e.

$$V^T \cdot D^2 L(\bar{x}) \cdot V \succ 0,$$

  where the columns of $V$ form a basis of $T_{\bar{x}} M$.

**Theorem (Second-order sufficient optimality condition)**

*Let a KKT-point $\bar{x}$ fulfil LICQ, SC, and SOSC. Then, $\bar{x}$ solves (P).*

# NONDEGENERACY

A minimizer $\bar{x}$ is called nondegenerate if it fulfills

- LICQ,
- SC,
- SOSC.

**Theorem (Nondegeneracy is generic)**

*Let $\mathcal{H} \subset C^2(\mathbb{R}^n, \mathbb{R}^1) \times C^2(\mathbb{R}^n, \mathbb{R}^{|J|})$ denote the subset of defining functions $(f, g)$ for which each minimizer is nondegenerate. Then, $\mathcal{H}$ is $C_s^2$-open and dense.*

**Remark (Strong or Whitney-topology)**

*$C_s^2$-topology is generated by allowing perturbations of the functions and their derivatives up to second order which are controlled by means of continuous positive functions.*

# FRITZ-JOHN SYSTEM

Consider the Fritz-John system:

$$\begin{cases} \bar{\delta} \cdot Df(\bar{x}) & = & \displaystyle\sum_{j \in J} \bar{\lambda}_j \cdot Dg_j(\bar{x}) \\ \bar{\lambda}_j \cdot g_j(\bar{x}) & = & 0, \quad \bar{\lambda}_j \geq 0, \quad g_j(\bar{x}) \geq 0 \quad \text{for all } j \in J, \\ \bar{\delta} & \geq & 0, \\ 1 & = & \bar{\delta} + \displaystyle\sum_{j \in J} \bar{\lambda}_j. \end{cases}$$

Last equation guarantees that not all Lagrange multipliers vanish.

# KOJIMA-TRICK

We define the positive/negative parts of $a \in \mathbb{R}$:

$$a^+ = \max\{a, 0\}, \quad a^- = \min\{a, 0\}.$$

Then, we can equivalently represent:

$$\bar{\lambda}_j \cdot g_j(\bar{x}) = 0, \quad \bar{\lambda}_j \geq 0, \quad g_j(\bar{x}) \geq 0$$

by setting $\bar{\lambda}_j = (\bar{\gamma}_j)^+$ and ensuring that

$$g_j(\bar{x}) + (\bar{\gamma}_j)^- = 0.$$

$$\boxed{\text{INTRODUCTION OF NEW NONSMOOTH VARIABLES } \bar{\gamma}_j}$$

# NONSMOOTH REFORMULATION

Rewrite it by using nonsmooth variables:

$$\begin{cases} \left(\bar{\delta}\right)^{+} \cdot Df(\bar{x}) & = & \sum_{j \in J} (\bar{\gamma}_j)^{+} \cdot Dg_j(\bar{x}) \\ g_j(\bar{x}) + (\bar{\gamma}_j)^{-} & = & 0 \quad \text{for all } j \in J, \\ 1 & = & \left(\bar{\delta}\right)^{+} + \sum_{j \in J} \left(\bar{\lambda}_j\right)^{+}. \end{cases}$$

$$\boxed{\#\text{Variables} = n + |J| + 1, \quad \#\text{Equations} = n + |J| + 1}$$

The violation of LICQ, SC, SOSC produces additional equations, but there are no available degrees of freedom. This ensures that nondegenericity of minimizers is a generic property.

# EXERCISES

- Show that the minimizers of the following nonlinear optimization problems are degenerate.
- Perform arbitrarily small perturbations of the defining functions in order to achieve nondegeneracy of the corresponding minimizer.

**Example (1)**

$$\min_{x \in \mathbb{R}} x^4$$

**Example (2)**

$$\min_{x \in \mathbb{R}} x^2 \quad \text{s.t.} \quad x \geq 0$$

# PARAMETRIC OPTIMIZATION

$$\mathrm{P}(t): \quad \min_{x \in \mathbb{R}^n} f(x,t) \quad \text{s.t.} \quad g_j(x,t) \geq 0, i \in J,$$

where all functions $f, g_j, j \in J$ are thrice continuously differentiable.

**Theorem (Type 1)**

*Let $\bar{x}$ be a nondegenerate local minimizer of $P(\bar{t})$ with Lagrange multipliers $\bar{\lambda}_j$, $j \in J_0(\bar{x}, \bar{t})$. Then, there exist twice continuously differentiable mappings $x(t)$ and $\lambda_j(t)$, $j \in J_0(\bar{x}, \bar{t})$, such that for all $t$ sufficiently close to $\bar{t}$ it holds:*

> *$x(t)$ is the unique local minimizer of $P(t)$ in a neighborhood of $\bar{x}$ with Lagrange multipliers $\lambda_j(t)$, $j \in J_0(\bar{x}, \bar{t})$ with*
>
> $$x(\bar{t}) = \bar{x}, \quad \lambda_j(\bar{t}) = \bar{\lambda}_j, j \in J_0(\bar{x}, \bar{t}).$$

# TRACKING OF MINIMIZERS

- Write the corresponding Karush-Kuhn-Tucker system:

$$\mathcal{T}(x,\lambda)=0 \Longleftrightarrow \left\{ \begin{array}{rcl} D_x f(x,t) & = & \displaystyle\sum_{j\in J_0(\bar{x},\bar{t})} \lambda_j \cdot D_x g_j(x,t), \\ & & \\ g_j(x,t) & = & 0 \quad \text{for all } j \in J_0(\bar{x},\bar{t}). \end{array} \right.$$

- Nondegeneracy of $\bar{x}$ for P $(\bar{t})$ with Lagrange multipliers $\bar{\lambda}$ implies that the matrix $D_{x,\lambda}\mathcal{T}(\bar{x},\bar{\lambda})$ is nonsingular.
- Apply the implicit function theorem to obtain the Karush-Kuhn-Tucker point $x(t)$ with Lagrange multipliers $\lambda(t)$.
- Use continuity reasons to argue that the Karush-Kuhn-Tucker point $x(t)$ fulfils LICQ, SC, SOSC.
- Due to second-order optimality condition, $x(t)$ solves $\mathrm{P}(t)$.

# SINGULARITIES

DEGENARACIES CANNOT BE AVOIDED
IN PARAMETRIC OPTIMIZATION

It may happen that at a minimizer of $\mathrm{P}(t)$ some of the conditions LICQ, SC, SOSC fail. More importantly, these degeneracies can be stable, i.e. it will be impossible to perturb defining functions in order to avoid this phenomenon. In fact, any violation of LICQ, SC, SOSC produces an additional equation to be satisfied, but now the Fritz-John system has $dim(t) > 0$ degrees of freedom available. Hence, singularities of different kinds naturally emerge by violating LICQ, SC, SOSC in various respect.

# EXERCISES

- Is $\bar{x} = 0$ for $\bar{t} = 0$ a Karush-Kuhn-Tucker or Fritz-John point?
- Check MFCQ, LICQ, SC, SOSC at $\bar{x} = 0$ for $\bar{t} = 0$.
- Derive the formula for the minimizer $x(t)$ in dependence of $t$.

**Example (Type 2)**

$$\min_{x \in \mathbb{R}^n} (x_1 - t)^2 + \sum_{j=2}^{n} x_j^2 \quad \text{s.t.} \quad x_1 \geq 0.$$

**Example (Type 4)**

$$\min_{x \in \mathbb{R}^n} -x_1 \quad \text{s.t.} \quad t - \sum_{j=1}^{n} x_j^2 \geq 0.$$

# EXERCISES

- Is $\bar{x} = 0$ for $\bar{t} = 0$ a Karush-Kuhn-Tucker or Fritz-John point?
- Check MFCQ/LICQ, SC, SOSC at $\bar{x} = 0$ for $\bar{t} = 0$.
- Derive the formula for the minimizer $x(t)$ in dependence of $t$.

**Example (Type 5.1)**

$$\min_{x \in \mathbb{R}^n} \sum_{j=1}^{n} x_j \quad \text{s.t.} \quad x_j \geq 0, j = 1, \ldots, n, \quad t - \sum_{j=1}^{n} x_j \geq 0.$$

**Example (Type 5.2)**

$$\min_{x \in \mathbb{R}^n} \sum_{j=1}^{n} j \cdot x_j \quad \text{s.t.} \quad x_j \geq 0, j = 1, \ldots, n, \quad -t + \sum_{j=1}^{n} x_j \geq 0.$$

# ONE-PARAMETRIC OPTIMIZATION

- Type 2: exactly one Lagrange multiplier vanishes, Karush-Kuhn-Tucker point ✓, SC ↯,
- Type 4: exactly one active gradient is linearly dependent from the others, Karush-Kuhn-Tucker point ↯, MFCQ ↯,
- Type 5.1: number of active inequalities exceeds dimension by exactly one, Karush-Kuhn-Tucker point ✓, MFCQ ↯,
- Type 5.2: number of active inequalities exceeds dimension by exactly one, Karush-Kuhn-Tucker point ✓, MFCQ ✓, LICQ ↯.

**Theorem (Five Types)**

*Up to a differentiable change of coordinates, the only possible singularities, which may generically occur at the minimizers in one-parametric optimization, are of Types 2, 4, 5.1, and 5.2.*

# LITERATURE


- Jongen, H. ; Meer, K. ; Triesch, E. (2004). Optimization Theory. Kluwer Academic Publishers, Dordrecht.

- Jongen, H.Th. ; Jonker, P. ; Twilt, F. (1986). Critical sets in parametric optimization. Mathematical Programming, Vol. 34, p. 333-353.

- Jongen, H. Th. ; Shikhman, V. (2012). Bilevel optimization: on the structure of the feasible set. Mathematical Programming, Vol. 136, p. 65-89.

# INTRODUCTION TO GENERALIZED NASH EQUILIBRIUM PROBLEMS

Vladimir Shikhman

Chemnitz University of Technology, Germany

Lecture 2

# GENERALIZED NASH EQUILIBRIUM PROBLEM

Players $\{1, \ldots, N\}$

$\boxed{\text{PARAMETRIC OPTIMIZATION of Player } \nu\text{:}}$

$$\mathrm{P}_\nu(x^{-\nu}): \qquad \min_{x^\nu} f^\nu(x^\nu, x^{-\nu}) \quad \text{s.t.} \quad x^\nu \in M^\nu(x^{-\nu}).$$

$\boxed{\text{FEASIBLE SET of Player } \nu\text{:}}$

$$M^\nu(x^{-\nu}) := \left\{ x^\nu \in \mathbb{R}^{n_\nu} \;\middle|\; \begin{array}{ll} g_j^\nu(x^\nu, x^{-\nu}) \geq 0, & j \in J^\nu, \\ G_j(x^\nu, x^{-\nu}) \geq 0, & j \in \mathcal{J} \end{array} \right\}$$

$f^\nu$ - **OBJECTIVE FUNCTION**
$g_j^\nu$ - **INDIVIDUAL CONSTRAINTS**
$G_j$ - **SHARED CONSTRAINTS**

# STRUCTURE OF GNE-SET

$$\left(\bar{x}^1, \ldots, \bar{x}^N\right) \text{ is GENERALIZED NASH EQUILIBRIUM}$$
$$\text{iff}$$
$$\bar{x}^\nu \text{ SOLVES } \mathrm{P}_\nu\left(\bar{x}^{-\nu}\right) \text{ for all } \nu = 1, \ldots, N$$

GOAL: Describe the set of GNEs locally around a given GNE $\bar{x}$.

- Consider Fritz-John system with Lagrange multipliers.
- Rewrite it by using nonsmooth variables.
- Study the solution set of this equation system.

# FRITZ-JOHN POINTS

Since $\bar{x}^\nu$ solves $\mathrm{P}_\nu(\bar{x}^{-\nu})$, there exist Lagrange multipliers

$$\bar{\delta}^\nu, \bar{\lambda}_j^\nu, j \in J^\nu, \bar{\Lambda}_j^\nu, j \in \mathcal{J},$$

– not all vanishing – such that it holds:

$$\left\{ \begin{array}{rcl} \bar{\delta}^\nu \cdot D_{x^\nu} f^\nu(\bar{x}) & = & \sum_{j \in J^\nu} \bar{\lambda}_j^\nu \cdot D_{x^\nu} g_j^\nu(\bar{x}) + \sum_{j \in \mathcal{J}} \bar{\Lambda}_j^\nu D_{x^\nu} G_j(\bar{x}) \\ \bar{\lambda}_j^\nu \cdot g_j^\nu(\bar{x}) & = & 0, \quad \bar{\lambda}_j^\nu \geq 0, \quad g_j^\nu(\bar{x}) \geq 0 \quad \text{for all } j \in J^\nu, \\ \bar{\Lambda}_j^\nu \cdot G_j(\bar{x}) & = & 0, \quad \bar{\Lambda}_j^\nu \geq 0, \quad G_j(\bar{x}) \geq 0 \quad \text{for all } j \in \mathcal{J}, \\ \bar{\delta}^\nu & \geq & 0. \end{array} \right.$$

# FRITZ-JOHN SYSTEM

By concatenating we get for all $\nu = 1, \ldots, N$:

$$\left\{ \begin{array}{rcl} \bar{\delta}^{\nu} \cdot D_{x^{\nu}} f^{\nu}(\bar{x}) & = & \sum_{j \in J^{\nu}} \bar{\lambda}_j^{\nu} \cdot D_{x^{\nu}} g_j^{\nu}(\bar{x}) + \sum_{j \in \mathcal{J}} \bar{\Lambda}_j^{\nu} D_{x^{\nu}} G_j(\bar{x}) \\ \bar{\lambda}_j^{\nu} \cdot g_j^{\nu}(\bar{x}) & = & 0, \quad \bar{\lambda}_j^{\nu} \geq 0, \quad g_j^{\nu}(\bar{x}) \geq 0 \quad \text{for all } j \in J^{\nu}, \\ \bar{\Lambda}_j^{\nu} \cdot G_j(\bar{x}) & = & 0, \quad \bar{\Lambda}_j^{\nu} \geq 0, \quad G_j(\bar{x}) \geq 0 \quad \text{for all } j \in \mathcal{J}, \\ \bar{\delta}^{\nu} & \geq & 0, \\ 1 & = & \bar{\delta}^{\nu} + \sum_{j \in J^{\nu}} \bar{\lambda}_j^{\nu} + \sum_{j \in \mathcal{J}} \bar{\Lambda}_j^{\nu}. \end{array} \right.$$

Last equation guarantees that not all Lagrange multipliers vanish.

# INDIVIDUAL CONSTRAINTS

Recall that the positive/negative parts of $a \in \mathbb{R}$ are defined as

$$a^+ = \max\{a, 0\}, \quad a^- = \min\{a, 0\}.$$

Due to the Kojima-trick, we can equivalently represent:

$$\bar{\lambda}_j^\nu \cdot g_j^\nu(\bar{x}) = 0, \quad \bar{\lambda}_j^\nu \geq 0, \quad g_j^\nu(\bar{x}) \geq 0$$

by setting $\bar{\lambda}_j^\nu = \left(\bar{\gamma}_j^\nu\right)^+$ and ensuring that

$$g_j^\nu(\bar{x}) + \left(\bar{\gamma}_j^\nu\right)^- = 0.$$

$$\boxed{\text{INTRODUCTION OF NEW NONSMOOTH VARIABLES } \bar{\gamma}_j^\nu}$$

## SHARED CONSTRAINTS

For $a^1, \dots a^N \in \mathbb{R}$ we define

$$\left(a^1, \dots a^N\right)^- = (-1)^{N-1} \cdot \prod_{\nu=1}^{N} (a^\nu)^- .$$

Analogously to the Kojima-trick, we can equivalently represent:

$$\bar{\Lambda}_j^\nu \cdot G_j(\bar{x}) = 0, \quad \bar{\Lambda}_j^\nu \geq 0, \quad G_j(\bar{x}) \geq 0, \quad \nu = 1, \dots, N,$$

by setting $\bar{\Lambda}_j^\nu = \left(\bar{\Gamma}_j^\nu\right)^+$, $\nu = 1, \dots, N$, and ensuring that

$$G_j(\bar{x}) + \left(\bar{\Gamma}_j^1, \dots, \bar{\Gamma}_j^N\right)^- = 0.$$

$\boxed{\text{LAGRANGE MULTIPLIERS } \bar{\Lambda}_j^\nu,\ \nu = 1, \dots, \text{ ARE LINKED}}$

# NONSMOOTH VARIABLES

We get for all $\nu = 1, \ldots, N$:

$$\begin{cases} \left(\bar{\delta}^{\nu}\right)^{+} \cdot D_{x^{\nu}} f^{\nu}(\bar{x}) & = & \displaystyle\sum_{j \in J^{\nu}} \left(\bar{\gamma}_j^{\nu}\right)^{+} \cdot D_{x^{\nu}} g_j^{\nu}(\bar{x}) \\ & & \displaystyle + \sum_{j \in \mathcal{J}} \left(\bar{\Gamma}_j^{\nu}\right)^{+} D_{x^{\nu}} G_j(\bar{x}) \\ g_j^{\nu}(\bar{x}) + \left(\bar{\gamma}_j^{\nu}\right)^{-} & = & 0 \quad \text{for all } j \in J, \\ 1 & = & \left(\bar{\delta}^{\nu}\right)^{+} + \displaystyle\sum_{j \in J^{\nu}} \left(\bar{\gamma}_j^{\nu}\right)^{+} + \sum_{j \in \mathcal{J}} \left(\bar{\Gamma}_j^{\nu}\right)^{+}, \end{cases}$$

and

$$G_j(\bar{x}) + \left(\bar{\Gamma}_j^{1}, \ldots, \bar{\Gamma}_j^{N}\right)^{-} = 0 \quad \text{for all } j \in \mathcal{J}.$$

# ACTIVE INEQUALITY CONSTRAINTS

Lagrange multipliers for inactive inequality constraints vanish:

$$\bar{\lambda}_j^\nu = 0 \quad \text{for all } j \notin J_0^\nu(\bar{x}), \quad \bar{\Lambda}_j^\nu = 0 \quad \text{for all } j \notin \mathcal{J}_0(\bar{x}),$$

where the index sets of active inequality constraints are

$$J_0^\nu(\bar{x}) = \left\{ j \in J \,\middle|\, g_j^\nu(\bar{x}) = 0 \right\}, \quad \mathcal{J}_0(\bar{x}) = \{ j \in \mathcal{J} \mid G_j(\bar{x}) = 0 \} .$$

Hence, the corresponding nonsmooth variables can be determined:

$$\bar{\gamma}_j^\nu = -g_j^\nu(\bar{x}) \quad \text{for all } j \notin J_0^\nu(\bar{x}), \quad \bar{\Gamma}_j^\nu = -G_j(\bar{x})^{\frac{1}{N}} \quad \text{for all } j \notin \mathcal{J}_0(\bar{x})$$

SKIP INACTIVE INEQUALITY CONSTRAINTS

# LOCAL PERSPECTIVE

Let $\bar{x}$ be a GNE. Due to continuity reasons, all GNE $x$ sufficiently close to $\bar{x}$ fulfil:

$$J_0^\nu(x) \subset J_0^\nu(\bar{x}), \quad \mathcal{J}_0(x) \subset \mathcal{J}_0(\bar{x}),$$

i. e. active inequality constraints cannot get more.
Hence, we may solve the Fritz-John system with respect to:

$$x, \delta^\nu, \gamma_j^\nu, j \in J_0^\nu(\bar{x}), \Gamma_j^\nu, j \in \mathcal{J}_0(\bar{x}), \nu \in \{1, \dots, N\}.$$

$\boxed{\text{INDEX SETS REMAIN FIXED}}$

# FRITZ-JOHN SYSTEM FOR $x \approx \bar{x}$

For all $\nu \in \{1, \ldots, N\}$ we have:

$$\begin{cases} (\delta^\nu)^+ \cdot D_{x^\nu} f^\nu(x) &= \displaystyle\sum_{j \in J_0^\nu(\bar{x})} \left(\gamma_j^\nu\right)^+ \cdot D_{x^\nu} g_j^\nu(x) \\ & \quad + \displaystyle\sum_{j \in \mathcal{J}_0(\bar{x})} \left(\Gamma_j^\nu\right)^+ D_{x^\nu} G_j(x) \\ g_j^\nu(x) + \left(\gamma_j^\nu\right)^- &= 0 \quad \text{for all } j \in J_0(\bar{x}), \\ 1 &= (\delta^\nu)^+ + \displaystyle\sum_{j \in J_0^\nu(\bar{x})} \left(\gamma_j^\nu\right)^+ + \sum_{j \in \mathcal{J}_0(\bar{x})} \left(\Gamma_j^\nu\right)^+, \end{cases}$$

and

$$G_j(\bar{x}) + \left(\Gamma_j^1, \ldots, \Gamma_j^N\right)^- = 0 \quad \text{for all } j \in \mathcal{J}_0(\bar{x}).$$

# NUMBER OF VARIABLES

For player $\nu \in \{1, \ldots, N\}$:

| $x^\nu$ | $\gamma_j^\nu, j \in J_0^\nu(\bar{x})$ | $\delta^\nu$ | $\Gamma_j^\nu, j \in \mathcal{J}_0(\bar{x})$ |
|---|---|---|---|
| $n_\nu$ | $\lvert J_0^\nu(\bar{x})\rvert$ | 1 | $\lvert \mathcal{J}_0(\bar{x})\rvert$ |

In total:

$$V = \sum_{\nu=1}^{N} n_\nu + \sum_{\nu=1}^{N} \left| J_0^\nu(\bar{x}) \right| + N + N \cdot \left| \mathcal{J}_0(\bar{x}) \right| .$$

# NUMBER OF EQUATIONS

For player $\nu \in \{1, \dots, N\}$:

| gradients | active individual constraints | not all vanishing |
|---|---|---|
| $n_\nu$ | $\lvert J_0^\nu(\bar{x}) \rvert$ | 1 |

Active shared constraints: $\lvert \mathcal{J}_0(\bar{x}) \rvert$

In total:

$$E = \sum_{\nu=1}^{N} n_\nu + \sum_{\nu=1}^{N} \lvert J_0^\nu(\bar{x}) \rvert + N + \lvert \mathcal{J}_0(\bar{x}) \rvert \,.$$

# DEFECT

$$V = \sum_{\nu=1}^{N} n_\nu + \sum_{\nu=1}^{N} |J_0^\nu(\bar{x})| + N + N \cdot |\mathcal{J}_0(\bar{x})|$$

$-$

$$E = \sum_{\nu=1}^{N} n_\nu + \sum_{\nu=1}^{N} |J_0^\nu(\bar{x})| + N + |\mathcal{J}_0(\bar{x})|$$

---

$$D = (N-1) \cdot |\mathcal{J}_0(\bar{x})|$$

UNDERDETERMINED SYSTEM:
MORE VARIABLES THAN EQUATIONS

# FRITZ-JOHN SET

We denote nonsmooth variables as

$$\eta = \left(\delta^{\nu}, \gamma_j^{\nu}, j \in J_0^{\nu}(\bar{x}), \Gamma_j^{\nu}, j \in \mathcal{J}_0(\bar{x}), \nu \in \{1, \ldots, N\}\right),$$

and write the Fritz-John system for short as

$$\mathcal{F}(x, \eta) = 0.$$

Its solutions $\mathcal{F}^{-1}(0)$ form the so-called Fritz-John set.

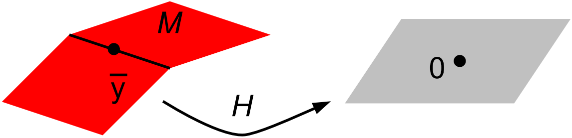
**Theorem**

*For a generic GNEP, locally at any $(\bar{x}, \bar{\eta})$ the Fritz-John set $\mathcal{F}^{-1}(0)$ is a Lipschitz manifold of dimension $(N-1) \cdot |\mathcal{J}_0(\bar{x})|$.*

# LIPSCHITZ MANIFOLD

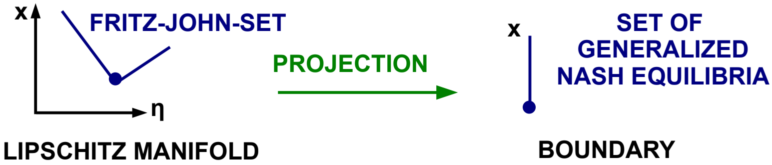
$M \subseteq \mathbb{R}^\ell$ is called a Lipschitz manifold of dimension $k$ if for each $\bar{y} \in M$ there exist open neighborhoods $U \subseteq \mathbb{R}^\ell$ of $\bar{y}$ and $V \subseteq \mathbb{R}^\ell$ of 0 and a Lipschitz homeomorphism $H : U \to V$ (i.e., with $H$ and $H^{-1}$ being Lipschitz continuous), such that

- $H(\bar{y}) = 0$,
- $H(M \cap U) = (\mathbb{R}^k \times \{0_{\ell-k}\}) \cap V$.



# STRUCTURE OF GNE-SET

$$\boxed{\text{FRITZ-JOHN SET } (x,\eta) \iff \text{SOLUTIONS of } \mathcal{F}(x,\eta)=0}$$




GNE-SET

$=$

PROJECTION OF $\mathcal{F}^{-1}(0)$ ON $x$-VARIABLES

# CRUCIAL NUMBER $(N-1)\,|\mathcal{J}_0(\bar{x})|$

$$(N-1)\,|\mathcal{J}_0(\bar{x})| \overset{\textit{encodes}}{=} \begin{cases} \text{DEGENERACIES IN PLAYERS'} \\ \text{PARAMETRIC OPTIMIZATION} \\ + \\ \text{DIMENSION OF GNE-SET} \end{cases}$$

Fritz-John system has $(N-1)\,|\mathcal{J}_0(\bar{x})|$ degrees of freedom.
Violation of nondegeneracy at players' minimizers is possible:

- LICQ $\lightning$: some active gradients may be linearly dependent from the others,
- SC $\lightning$: some Lagrange multipliers may vanish,
- SOSC $\lightning$: some eigenvalues of the Lagrange function's Hessian restricted to the tangent space may vanish.

Remaining degrees of freedom go for the dimension of GNE-set.

# EXERCISES

- Compute GNEs of the following two-player GNEPs.
- Discuss their structure with respect to occuring degeneracies.

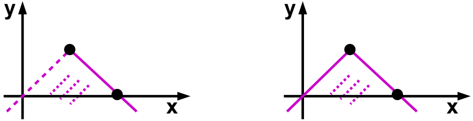
**Example (3)**

$$f^1(x,y) = -x, \quad f^2(x,y) = -y,$$
$$G_1(x,y) = 1 - x - y, \quad G_2(x,y) = x - y.$$

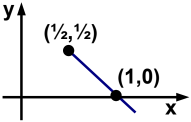
**Example (4)**

$$f^1((x,y),t) = x, \quad f^2((x,y),t) = t,$$
$$G_1((x,y),t) = 1 - (x-t)^2 - (y-(1-2t))^2,$$
$$G_2((x,y),t) = 1 - x^2 - (y+1)^2.$$

# EXAMPLE (3)



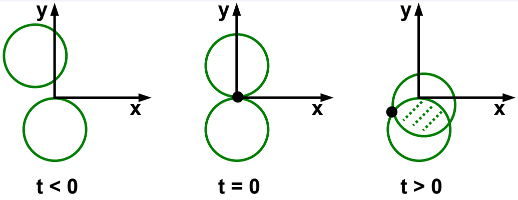

MINIMIZERS OF PLAYERS 1 AND 2




SET OF GENERALIZED NASH EQUILIBRIA

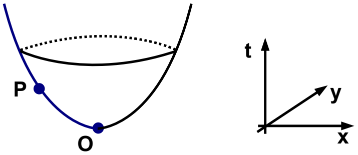
| POINT | $(\mathbf{N}-\mathbf{1})\lvert\mathcal{J}_{\mathbf{0}}\rvert$ | PLAYER 1 | PLAYER 2 | SET OF NE |
|---|---|---|---|---|
| $(1/2, 1/2)$ | 2 | DEG $=1$ | DEG $=1$ | BOUNDARY |
| $(\mathbf{1}, \mathbf{0})$ | 1 | NONDEG | NONDEG | DIM $=1$ |

# EXAMPLE (4)




$f^1((x,y),t) = x$

**FEASIBLE SET OF PLAYER 1**




$f^2((x,y),t) = t$

**SET OF GENERALIZED NASH EQUILIBRIA**

| POINT | $(N-1)\lvert\mathcal{J}_0\rvert$ | PLAYER 1 | PLAYER 2 | SET OF NE |
|---|---|---|---|---|
| O | 2 | DEG $=1$ | DEG $=1$ | BOUNDARY |
| P | 2 | NONDEG | DEG $=1$ | DIM $=1$ |

# SPECIAL CASES FOR $(N-1)\,|\mathcal{J}_0(\bar{x})|$

- ONE PLAYER, i. e. $N = 1$:

  $\boxed{\text{Minimizers are isolated and nondegenarate}}$

- NO SHARED CONSTRAINTS, i. e. $\mathcal{J} = \emptyset$:

  $\boxed{\begin{array}{c}\text{Nash equilibria are isolated and} \\ \text{minimizers are nondegenarate}\end{array}}$

- NONDEGENERATE MINIMIZERS FOR EACH PLAYER:

  $\boxed{\begin{array}{c}\text{Set of generalized Nash equilibria is} \\ \text{a smooth manifold of dimension } (N-1)\,|\mathcal{J}_0(\bar{x})|\end{array}}$

# NORMALIZED NASH EQUILIBRIA

Fritz-John system has $(N-1)\,|\mathcal{J}_0(\bar{x})|$ degrees of freedom. This is due to the fact that the players' Lagrange multipliers for shared constraints are different. Rosen considers those GNEs $\bar{x}$ – called normalized Nash equilibria – with equal Lagrange multipliers, i. e. for $j \in \mathcal{J}_0(\bar{x})$ he sets:

$$\bar{\Lambda}_j^1 = \ldots = \bar{\Lambda}_j^N.$$

This produces additional $(N-1)\,|\mathcal{J}_0(\bar{x})|$ equations and makes the Fritz-John system determined, i. e. $\#$ Variables $= \#$ Equations.

**Theorem**

*Generically, all normalized Nash equilibria are isolated. Moreover, the corresponding players' minimizers are nondegenerate.*

# SHADOW PRICES

- Normalized Nash equilibria assume that Lagrange multipliers for shared constraints coincide. Note that the latter can be viewed as shadow prices. The shadow price is the change in the optimal value of the player's objective function obtained by infinitesimally relaxing the shared constraint.

- Coinciding shadow prices of different players become in some sense public. This observation motivates to relate GNEPs to the markets of common resources. Their prices have, thus, to be modelled explicitly. We go into this direction in Lecture 3.

# LITERATURE


- Rosen J. B. (1965). Existence and uniqueness of equilibrium points for concave $N$-person games. Econometrica, Vol. 33, no. 3, p. 520-534.

- Dorsch, D., Jongen, H. Th., Shikhman, V. (2013). On structure and computation of generalized Nash equilibria. SIAM Journal on Optimization, Vol. 23, no. 1, p. 452-474.

- Dorsch, D., Jongen, H. Th., Shikhman, V. (2013). On intrinsic complexity of Nash equilibrium problems and bilevel optimization. Journal of Optimization Theory and Applications, Vol. 159, no. 3, p. 606-634.

# INTRODUCTION TO GENERALIZED NASH EQUILIBRIUM PROBLEMS

Vladimir Shikhman

Chemnitz University of Technology, Germany

Lecture 3

# ALLOCATION PROBLEM

Players $\{1, \ldots, N\}$

$\boxed{\text{MAXIMIZE WELFARE SUBJECT TO SHARED CONSTRAINTS}}$

$$\max_{\substack{x^\nu \in X^\nu \\ \nu = 1, \ldots, N}} \sum_{\nu=1}^{N} f^\nu(x^\nu) \quad \text{s.t.} \quad \sum_{\nu=1}^{N} A^\nu x^\nu \leq b.$$

- $X^\nu \subset \mathbb{R}^{n_\nu}$ are convex and compact production sets,
- $f^\nu : \mathbb{R}^{n_\nu} \to \mathbb{R}$ are concave profit functions,
- $A^\nu \in \mathbb{R}^{m \times n_\nu}$ are transformation matrices,
- $b \in \mathbb{R}^m$ is the vector of available resources.

# DUAL APPROACH

$$\max_{\substack{x^\nu \in X^\nu \\ \nu=1,\dots,N}} \sum_{\nu=1}^{N} f^\nu(x^\nu) \quad \text{s.t.} \quad \underbrace{\sum_{\nu=1}^{N} A^\nu x^\nu \le b}_{\textit{dual multipliers } p\in\mathbb{R}^m}$$

$$=$$

$$\max_{\substack{x^\nu \in X^\nu \\ \nu=1,\dots,N}} \inf_{p\in\mathbb{R}^m_+} \sum_{\nu=1}^{N} f^\nu(x^\nu) + \underbrace{\left\langle p, b - \sum_{\nu=1}^{N} A^\nu x^\nu \right\rangle}_{\textit{penalty}}$$

Let production $x = \left(x^1, \dots, x^N\right)$ be optimal. If the $j$-th shared constraint is violated, then taking the price $p_j \to \infty$ drives the penalty to $-\infty$, a contradiction to the optimality of $x$. Vice versa, inactive $j$-th shared constraint corresponds to $p_j = 0$ in view of minimization. Thus, the penalty vanishes at the optimal $x$ and $p$.

# SION'S MINMAX THEOREM

**Theorem**

*Let $F : X \times P :\to \mathbb{R}$ satisfy:*

- $X$ *is compact and convex subset of* $\mathbb{R}^N$,
- $F(\cdot, p)$ *is concave on* $X$ *for all* $p \in P$,
- $P$ *is convex subset of* $\mathbb{R}^m$,
- $F(u, \cdot)$ *convex on* $P$ *for all* $x \in X$.

*Then it holds:*

$$\max_{x \in X} \inf_{p \in P} F(x, p) = \inf_{p \in P} \max_{x \in X} F(x, p).$$

The compactness assumption is essential.

# EXERCISE

## Example (5)

Let us denote by $S$ the set of probability measures on $\mathbb{N}$, i. e.

$$S = \left\{ (z_i)_{i \in \mathbb{N}} \subset \mathbb{R} \;\middle|\; \sum_{i=1}^{\infty} z_i = 1, z_i \geq 0, i \in \mathbb{N} \right\}.$$

Consider the following function defined for $x, y \in S$:

$$G(x,y) = \sum_{i,j=1}^{\infty} u(i,j) x_i y_i, \quad \text{where } u_{i,j} = \begin{cases} 1 & i > j, \\ 0 & i = j, \\ -1 & i < j. \end{cases}$$

Show that $\sup\limits_{x \in S} \inf\limits_{y \in S} G(x,y) = -1$ and $\inf\limits_{x \in S} \sup\limits_{y \in S} G(x,y) = 1$.

# LAGRANGE DUALITY

Sion's minmax theorem allows to exchange "inf" and "max":

$$\max_{\substack{x^\nu \in X^\nu \\ \nu = 1,\ldots,N}} \inf_{p \in \mathbb{R}^m_+} \underbrace{\sum_{\nu=1}^N f^\nu(x^\nu) + \left\langle p, b - \sum_{\nu=1}^N A^\nu x^\nu \right\rangle}_{=F(x,p)}$$

$$=$$

$$\inf_{p \in \mathbb{R}^m_+} \max_{\substack{x^\nu \in X^\nu \\ \nu = 1,\ldots,N}} \sum_{\nu=1}^N f^\nu(x^\nu) + \left\langle p, b - \sum_{\nu=1}^N A^\nu x^\nu \right\rangle$$

- $F(\cdot, p)$ is concave on the compact and convex set $\prod_{\nu=1}^N X^\nu$ for all $p \in \mathbb{R}^m_+$,
- $F(x, \cdot)$ is convex on the convex set $\mathbb{R}^m_+$ for all $x \in \prod_{\nu=1}^N X^\nu$.

# PRICING PROBLEM

$$\inf_{p \in \mathbb{R}^m_+} \max_{\substack{x^\nu \in X^\nu \\ \nu = 1, \ldots, N}} \sum_{\nu=1}^{N} f^\nu(x^\nu) + \left\langle p, b - \sum_{\nu=1}^{N} A^\nu x^\nu \right\rangle$$

$$=$$

$$\inf_{p \in \mathbb{R}^m_+} \underbrace{\sum_{\nu=1}^{N} \max_{x^\nu \in X^\nu} \left[ \underbrace{f^\nu(x^\nu)}_{profit} - \underbrace{\langle p, A^\nu x^\nu \rangle}_{cost} \right] + \langle p, b \rangle}_{=\Psi(p)}$$

$$\boxed{\text{MINIMIZE REVENUE WITH RESPECT TO THE PRICES}}$$

# MARKET EQUILIBRIUM

$$\underbrace{\inf_{p \in \mathbb{R}^m_+}}_{\textit{pricing}} \sum_{\nu=1}^{N} \underbrace{\max_{x^\nu \in X^\nu}}_{\textit{production}} \left[f^\nu(x^\nu) - \langle p, A^\nu x^\nu \rangle\right] + \langle p, b \rangle$$

A vector $(\bar{x}_1, \ldots, \bar{x}_N, \bar{p})$ is called market equilibrium if it satisfies:

- $\nu$-th producer optimally adjusts production, i. e.

$$\bar{x}^\nu \in \arg\max_{x^\nu \in X^\nu} \left[f^\nu(x^\nu) - \langle \bar{p}, A^\nu x^\nu \rangle\right], \quad \nu = 1, \ldots, N,$$

- the market of resources is cleared, i. e.

$$\bar{p} \geq 0, \quad \sum_{\nu=1}^{N} A^\nu \bar{x}^\nu \geq b, \quad \left\langle \bar{p}, b - \sum_{\nu=1}^{N} A^\nu \bar{x}^\nu \right\rangle = 0.$$

# CHARACTERIZATION

**Theorem (Strong duality)**

*The vector* $(\bar{x}_1, \ldots, \bar{x}_N, \bar{p})$ *is a market equilibrium if one one the following equivalent assertions holds:*

- $(\bar{x}_1, \ldots, \bar{x}_N)$ *and* $\bar{p}$ *solve the allocation and pricing problem,*
- *production* $(\bar{x}_1, \ldots, \bar{x}_N)$ *and prices* $\bar{p}$ *are feasible, i. e.*

$$\bar{x}^\nu \in X^\nu, \nu = 1, \ldots, N, \quad \sum_{\nu=1}^{N} A^\nu \bar{x}^\nu \leq b, \quad \bar{p} \in \mathbb{R}^m_+,$$

*and they close the primal-dual gap:*

$$\sum_{\nu=1}^{N} f^\nu(\bar{x}^\nu) = \Psi(\bar{p}).$$

# SLATER CONSTRAINT QUALIFICATION

Is the revenue minimization problem solvable?

$$\inf_{p \in \mathbb{R}^m_+} \underbrace{\sum_{\nu=1}^{N} \max_{x^\nu \in X^\nu} \left[ f^\nu(x^\nu) - \langle p, A^\nu x^\nu \rangle \right] + \langle p, b \rangle}_{=\Psi(p)}$$

For that, we need to assume the Slater constraint qualification. The latter is said to hold if a strict feasible production regime can be implemented, i. e.

$$\text{there exist } \widetilde{x}^\nu \in X^\nu, \nu = 1, \dots, N, \text{ with } \sum_{\nu=1}^{N} A^\nu \widetilde{x}^\nu < b.$$

# LOWER LEVEL SETS OF Ψ

Due to the Slater constraint qualification, we have:

$$\begin{aligned}
\Psi(p) &= \sum_{\nu=1}^{N} \max_{x^\nu \in X^\nu} \left[ f^\nu(x^\nu) - \langle p, A^\nu x^\nu \rangle \right] + \langle p, b \rangle \\
&\geq \sum_{\nu=1}^{N} \left[ f^\nu(\widetilde{x}^\nu) - \langle p, A^\nu \widetilde{x}^\nu \rangle \right] + \langle p, b \rangle \\
&= \underbrace{\sum_{\nu=1}^{N} f^\nu(\widetilde{x}^\nu)}_{=\gamma} + \left\langle p, \underbrace{b - \sum_{\nu=1}^{N} A^\nu \widetilde{x}^\nu}_{=z>0} \right\rangle
\end{aligned}$$

# EQUILIBRIUM PRICES

Hence, for any $\varrho \in \mathbb{R}$ it holds:

$$\Psi(p) \leq \varrho \quad \Longrightarrow \quad \langle p, z \rangle + \gamma \leq \varrho.$$

Since $z > 0$, it follows:

**Theorem (Existence of equilibrium prices)**

*Under Slater constraint qualification, the lower level sets of $\Psi$ are bounded. Moreover, the set of minimizers of $\Psi$ is nonempty and compact.*

Up to now, we assume that the Slater constraint qualification is fulfilled. This correspond to the strict feasible production regime, a quite reasonable assumption from the economic point of view.

# PRICE ADJUSTMENT

How to solve the revenue minimization problem?

$$\min_{p \in \mathbb{R}^m_+} \quad \Psi(p),$$

where

$$\Psi(p) = \sum_{\nu=1}^{N} \underbrace{\max_{x^\nu \in X^\nu} \quad [f^\nu(x^\nu) - \langle p, A^\nu x^\nu \rangle]}_{=\Psi^\nu(p)} + \langle p, b \rangle .$$

$$\boxed{\text{SUBGRADIENT METHODS}}$$

# SUBDIFFERENTIAL OF $\Psi^\nu$

$$\Psi^\nu(p) = \max_{x^\nu \in X^\nu} \left[ f^\nu(x^\nu) - \langle p, A^\nu x^\nu \rangle \right]$$

- $\Psi^\nu$ is convex as the maximum of convex functions,
- its convex sudifferential is

$$\partial \Psi^\nu(p) = \{ - A^\nu x^\nu \mid x^\nu \in X_*^\nu(p) \},$$

where

$$X_*^\nu(p) = \arg\max_{x^\nu \in X^\nu} \left[ f^\nu(x^\nu) - \langle p, A^\nu x^\nu \rangle \right]$$

describes the optimal production given the prices $p$.

"DIFFERENTIATE INSIDE THE BRACKETS"

# SPECIAL CASE

Let $\bar{x}^\nu \in X_*^\nu(p)$ be the unique nondegenerate maximizer of $f^\nu(\cdot) - \langle \bar{p}, A^\nu \cdot \rangle$ lying in the interior of $X^\nu$, i. e.

$$Df^\nu(\bar{x}^\nu) - (A^\nu)^T \bar{p} = 0, \quad D^2 f^\nu(\bar{x}^\nu) \prec 0.$$

Implicit function theorem provides the existence of $x^\nu(p)$, $p \approx \bar{p}$, satisfying:

$$Df^\nu(x^\nu(p)) - (A^\nu)^T p = 0, \quad x^\nu(\bar{p}) = \bar{x}^\nu.$$

Due to continuity, $D^2 f^\nu(x^\nu(p)) \prec 0$ for $p$ sufficiently close to $\bar{p}$. Hence, second-order sufficient condition ensures that $x^\nu(p)$ maximizes $f^\nu(\cdot) - \langle p, A^\nu \cdot \rangle$ on $X^\nu$.

# SPECIAL CASE (CONTINUED)

Then, we obtain for the optimal value function:

$$\Psi^\nu(p) = f^\nu\left(x^\nu(p)\right) - \langle p, A^\nu x^\nu(p)\rangle .$$

Let us compute its gradient by using the chain rule:

$$\begin{aligned} D\Psi^\nu(p) &= D_x f^\nu\left(x^\nu(p)\right) Dx^\nu(p) - A^\nu x^\nu(p) - (A^\nu)^T p\, Dx^\nu(p) \\ &= \underbrace{\left(D_x f^\nu\left(x^\nu(p)\right) - (A^\nu)^T p\right)}_{=0,\ \textit{due to optimality}} Dx^\nu(p) - A^\nu x^\nu(p) \end{aligned}$$

Evaluating at $\bar{p}$ finally gives the desired formula:

$$D\Psi^\nu(\bar{p}) = -A^\nu \underbrace{x^\nu(\bar{p})}_{=\bar{x}^\nu} = -A^\nu \bar{x}^\nu .$$

# EXERCISE

## Example (6)

Consider the following max-type function:

$$\psi(a, b) = \max_x \left[ -\frac{1}{2} a x^2 + b x \right]$$

for $a > 0$ and $b \approx 0$. Discuss the sensitivity of $\psi$ with respect to the parameters $a$ and $b$ both by "differentiating inside the brackets" and by a direct calculation. Is $\psi$ convex?

# SUBDIFFERENTIAL OF $\Psi$

$$\Psi(p) = \sum_{\nu=1}^{N} \underbrace{\max_{x^\nu \in X^\nu} \left[ f^\nu(x^\nu) - \langle p, A^\nu x^\nu \rangle \right]}_{=\Psi^\nu(p)} + \langle p, b \rangle .$$

- $\Psi$ is convex as the sum of convex functions,
- its convex sudifferential is

$$\partial\Psi(p) \;=\; \sum_{\nu=1}^{N} \partial\Psi^\nu(p) + b = \left\{ \underbrace{b - \sum_{\nu=1}^{N} A^\nu x^\nu}_{supply\ -\ demand} \;\middle|\; x^\nu \in X_*^\nu(p) \right\} .$$

# TÂTONNEMENT PROCESS

$$\text{Excess demand: } z(p) = \underbrace{\sum_{\nu=1}^{N} A^{\nu} x^{\nu}(p) - b,}_{\textit{demand} \; - \; \textit{supply}}$$

where $x^{\nu}(p) \in X_*^{\nu}(p)$ is an optimal production of $\nu$-th player.

$$\text{WALRASIAN TÂTONNEMENT: } \boxed{\frac{dp}{dt} = z(p)}$$

"demand $>$ supply $\Longrightarrow$ price rises"
"demand $<$ supply $\Longrightarrow$ price falls"

$\boxed{\text{COUPLE THE PRICE UPDATE ON THE EXCESS DEMAND}}$

# PROJECTED SUBGRADIENT METHOD

By discretizing the Walrasian tâtonnement with respect to the time variable, we get:

$$\frac{dp}{dt} = z(p) \quad \rightsquigarrow \quad \frac{p(t+\gamma) - p(t)}{\gamma} = z(p(t)).$$

By taking into account that prices should be nonnegative, we obtain the projected subgradient method:

$$p_{t+1} = \left[p_t + \gamma_t \cdot z\left(p_t\right)\right]_+ ,$$

where $\gamma_t$ is a stepsize to be chosen and $z\left(p_t\right) \in -\partial \Psi\left(p_t\right)$.

# RESOURCE MANAGEMENT

**1.** $\nu$-th producer chooses optimal production $x^\nu(p_t) \in X_*^\nu(p_t)$.

**2.** Manager observes current excess demand

and updates prices $p_{t+1} = \left[ p_t + \gamma_t \cdot \left( \sum_{\nu=1}^{N} A^\nu x^\nu(p_t) - b \right) \right]_+$

Does the vector of productions and prices

$$\left( x^1(p_t), \dots, x^N(p_t), p_t \right)$$

approach the set of market equilibria for $t \to \infty$, how fast?

USE AVERAGING FOR STABILIZING THE SEQUENCE

# RESOURCE MANAGEMENT VIA AVERAGING

**1.** $\nu$-th producer maximizes revenue by $x^\nu(p[t]) \in X_*^\nu(p[t])$

and implements average production $x^\nu[t] = \sum_{r=0}^{t} x^\nu(p[r])$.

**2.** Manager observes current excess demand

and forecasts prices $p^+[t+1] = \frac{1}{\Gamma[t]} \circ \left[\sum_{\nu=1}^{N} A^\nu x^\nu[t] - b\right]_+$

by using average parameters $\Gamma[t] = \frac{1}{t+1} \sum_{r=0}^{t} \gamma[r]$.

**3.** Manager averages price forecasts $p[t+1] = \frac{1}{t+2} \sum_{r=0}^{t+1} p^+[r]$

# PRIMAL-DUAL ITERATION

$$\underbrace{x^{\nu}[t]}_{\substack{\textit{next}\\ \textit{production}}} = \frac{t}{t+1} \underbrace{x^{\nu}[t-1]}_{\substack{\textit{previous}\\ \textit{production}}} + \frac{1}{t+1} \underbrace{x^{\nu}(p[t])}_{\substack{\textit{production}\\ \textit{forecast}}}$$

$$\underbrace{p[t+1]}_{\substack{\textit{next}\\ \textit{price}}} = \frac{t+1}{t+2} \underbrace{p[t]}_{\substack{\textit{previous}\\ \textit{price}}} + \frac{1}{t+2} \underbrace{p^{+}[t+1]}_{\substack{\textit{price}\\ \textit{forecast}}}$$

$\boxed{\text{COUPLING VIA FORECAST}}$

$$\underbrace{p^{+}[t+1]}_{\substack{\textit{price}\\ \textit{forecast}}} = \frac{\mathbf{1}}{\Gamma[t]} \circ \underbrace{\left[\sum_{\nu=1}^{N} A^{\nu} x^{\nu}[t] - b\right]_{+}}_{\textit{excess demand}}$$

# ADMISSIBLE PARAMETERS

$$p^{+}[t+1] = \frac{\mathbf{1}}{\Gamma[t]} \circ \left[ \sum_{\nu=1}^{N} A^{\nu} x^{\nu}[t] - b \right]_{+} \quad \text{with} \quad \Gamma[t] = \frac{1}{t+1} \sum_{r=0}^{t} \gamma[r]$$

**Theorem (Convergence to market equilibrium)**

*Choose* $\gamma[t] \to \mathbf{0}$ *and* $\sum_{r=0}^{t} \gamma[r] \to \infty$, *then it holds for* $t \to \infty$*:*

$$\underbrace{\Psi(p[t]) - \sum_{\nu=1}^{N} f^{\nu}(x^{\nu}[t])}_{\textit{primal} - \textit{dual gap}} \to 0, \quad \underbrace{\left\| \left[ \sum_{\nu=1}^{N} A^{\nu} x^{\nu}[t] - b \right]_{+} \right\|_{2}}_{\textit{infeasibility gap}} \to 0.$$

# OPTIMAL PARAMETERS

**Theorem (Convergence rate)**

*Choose* $\gamma[t] \sim \frac{1}{\sqrt{t+1}}$, *then it holds for* $t \to \infty$*:*

$$\underbrace{\Psi(p[t]) - \sum_{\nu=1}^{N} f^{\nu}(x^{\nu}[t])}_{\textit{primal} - \textit{dual gap}} \sim O\left(\frac{1}{\sqrt{t+1}}\right),$$

$$\underbrace{\left\| \left[ \sum_{\nu=1}^{N} A^{\nu} x^{\nu}[t] - b \right]_{+} \right\|_{2}}_{\textit{infeasibility gap}} \sim O\left(\frac{1}{\sqrt{t+1}}\right).$$

# LITERATURE


- Nesterov, Yu. (2004). Introductory Lectures on Convex Optimization. Kluwer Academic Publishers, Boston.

- Anstreicher, K.M. ; Wolsey, L.A. (2009). Two "well-known" properties of subgradient optimization. Mathematical Programming, Vol. 120, p. 213-220.

- Nesterov, Yu. ; Shikhman, V. (2018) Dual subgradient method with averaging for optimal resource allocation. European Journal of Operational Research, Vol. 270, no. 3, p. 907-916.